Modeling Social Participation as Predictive of Life Satisfaction and Social Support: Scale or Index?

Anne T. Zelenka
University of Denver

MODELING SOCIAL PARTICIPATION AS

PREDICTIVE OF LIFE SATISFACTION AND SOCIAL CONNECTEDNESS:

SCALE OR INDEX?

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Anne T. Zelenka

November 2011

Advisor: Dr. Duan Zhang

Author: Anne T. Zelenka
Title: MODELING SOCIAL PARTICIPATION AS PREDICTIVE OF LIFE SATISFACTION AND SOCIAL SUPPORT: SCALE OR INDEX?
Advisor: Dr. Duan Zhang
Degree Date: November 2011

Abstract

Social participation in late adulthood through activities such as volunteering with charities, playing sports, and joining clubs can increase life satisfaction, directly by providing enjoyable engagement and indirectly by increasing a person's sense of social connectedness. When reported levels of different types of activities are used to measure social participation, conventional measure development procedures based on classical test theory lead to a proliferation of small participation subscales that don't show good reliability, don't have theoretical power, and don't match researchers' conceptions of the dimensions of participation. Based on the poor performance of conventional approaches, some researchers have suggested that social participation should be modeled as an *index* composed of its indicators rather than as a *scale* in which indicators reflect an underlying latent factor. Typical approaches in psychosocial research rely on reflective-indicator models, which correspond to scale development, rather than incorporating composite variables with causal indicators. The latter approach, where manifest indicators are specified as causing the unobserved construct, is sometimes known as *formative measurement*, since the construct of interest is formed by its indicators. This study compared a scale model of social participation based on reflective measurement to an index model based on formative measurement.

Using a sample representative of community-dwelling U.S. adults over age 65 from the Health and Retirement Study's 2008 wave of data collection, two alternative

measurement models of participation were constructed using sixteen items that recorded frequency of participation in different activities. Because patterns of participation differed for males and females, gender-specific models were developed. The scale models assigned participation items to subscales based on item intercorrelations. The index model assigned items to participation composites based on predictive associations with the outcomes of social connectedness and life satisfaction.

The index construction process led to a unidimensional representation of participation, composed of six of the original sixteen participation activity items. The initial attempts to build a scale model led to structures with many small factors and poor predictive validity. Based on the findings of unidimensionality for the index model, a single-factor scale model was explored for female respondents only. Results showed that both index and scale approaches have the potential to produce participation models that are parsimonious, well-fitting, and externally valid even though conventional scale development rules-of-thumb and current conceptions of the domains of participation lead the researcher to non-parsimonious, poorly-fitting solutions lacking predictive capability.

Participation measurement instrument developers often theorize the existence of three or more dimensions of participation. Whether they use conventional (reflective indicator) or more radical (formative indicator) models, they are advised, based on this study’s results, to evaluate a single-dimensional structure among their candidate models.

Acknowledgements

I would like to thank my dissertation advisor, Dr. Duan Zhang, for her timely and expert feedback on my work, whether for coursework, conference submissions, or this final project. Thank you to committee member and department head Dr. Kathy Green for mentoring and inspiration. Thank you to Dr. Julia Dmitrieva for providing insightful suggestions on my dissertation and for agreeing to serve on my committee during a busy period in her life. Thank you to Dr. Walter LaMendola for serving as my external chair and for bringing insight and expertise to my dissertation defense.

My parents Leigh Truitt and Christie Cave and their respective partners Helga Watt and George Cave inspired this project through the example they set of how to live the good life in retirement. They have shown me the range of ways that a person can meaningfully engage with life once paid work no longer fills up the hours and days.

Finally, a special thank you to my husband, Rick for his unflagging support and encouragement.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction and Literature Review

As adults in Western, developed countries enter their sixties and seventies, they transition away from paid work into other social roles and activities such as volunteering with community organizations, joining social clubs, pursuing hobbies with like-minded friends, competing in sports events, or engaging in non-competitive physical activity (Harlow & Cantor, 1996). This *social participation* predicts higher life satisfaction in older adults and is therefore an active area of gerontological study (Harlow & Cantor, 1996; Wahrendorf, Ribet, Zins, & Siegrist, 2008; Warr, Butcher, & Robertson, 2004). In addition to its importance in aging research, social participation serves as a key outcome in the area of disability and rehabilitation (Dijkers, 2010). Reflecting its significance in that setting, a number of participation measurement instruments have been developed for use with healthy and disabled respondents (e.g., M. Brown, et al., 2004; Mars et al., 2009; Reistetter, Spencer, Trujillo, & Abreu, 2005; Sander et al., 1999; Schuling, de Haan, Limburg, & Groenier, 1993). Of course, aging research and disability research overlap, since as people age, they are more likely to suffer health problems that lead to disability (Brault, 2008). Social participation plays a pivotal role in studies of both healthy aging and aging with disability, and has justifiably received ample research attention.

Though formal instruments and ad hoc approaches exist for measuring social participation, researchers have not reached agreement about its definition or how to model it (Whiteneck & Dijkers, 2009). This research study considered how social participation might best be modeled in the context of its predictive relationship with social connectedness and life satisfaction in U.S. adults age 65 and over. The study asked: Should social participation, measured with reported levels of different activities, be modeled as an index (with causal indicators) or as a scale (with reflective indicators)? It answered the question using structural equation modeling (SEM) to compare measurement models developed using the two approaches. The study demonstrated a set of procedures for testing and refining index measures based on the incorporation of composite variables into a structural equation model that includes predicted outcomes. The results were compared to scale models constructed using conventional exploratory and confirmatory factor analysis techniques.

The construct of social participation, measured by reported levels of different activities, can be modeled using a reflective or formative approach. In conventional measurement modeling based on classical test theory (CTT), observed indicators of an unobserved construct are modeled as reflecting an underlying common factor (Kline, 2005). This *reflective measurement model* of participation is shown in Figure 1, in which the levels of different kinds of participatory activities are determined by an underlying latent factor, perhaps a drive to participate socially. Typically, measures developed using CTT are called *scales*. Some researchers, however, have argued that participation is better modeled as an *index*, where observed levels of participation in different activities compose or in some sense cause the unobserved overall participation level (Dijkers, 2010; Mars et al., 2009). This model, as shown in Figure 2, is known as a *formative measurement model* (Diamantopoulos, Riefler, & Roth, 2008).

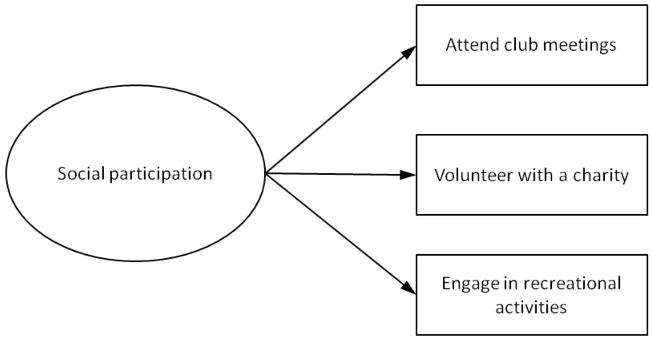


Figure 1. Reflective model of social participation

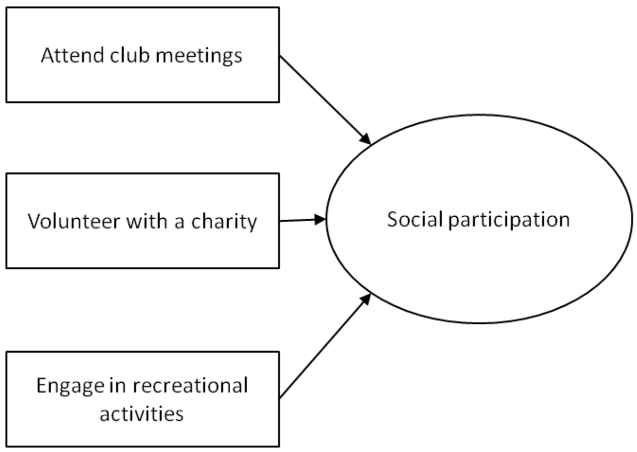


Figure 2. Formative model of social participation

Structural equation modeling (SEM), the analytical approach used in this study to compare the reflective and formative approaches, is a family of techniques that allows for estimating the complex inter-relationships that hold among psychosocial variables, correcting for measurement error through the use of multiple indicators of each underlying, unobserved variable (Kline, 2005). In contrast to more traditional statistical methods such as analysis of variance (ANOVA) and multiple regression, SEM usually begins with the researcher's developing a theoretically-grounded, a priori model of the research situation of interest. In SEM, the focus is not on granular null hypothesis significance tests of individual mean comparisons or regression coefficients but rather on the overall model and how it fits the empirical data (Kline, 2005; Rodgers, 2010). SEM is thus primarily a modeling technique rather than a technique for testing individual, fine-grained hypotheses, though it supports that as well. In SEM, statistical inferences are made, but in support of developing a model that helps us understand the world better. SEM allows researchers to pit alternative models against each other to see which one fits the empirical data better. It is a good match for this study, which evaluated whether an index or scale model of social participation is most appropriate based on overall model fit, parsimony, and predictive validity.

In CTT-based structural equation modeling using reflective measurement models, measurement portions of the model can and usually are tested in isolation from the structural model which specifies directional relationships between latent constructs (Anderson & Gerbing, 1988). But with a formative model the consequences of formative constructs play a key role in estimation and validation; the measurement model cannot be tested in isolation from the predictive structural model (Jarvis, MacKenzie, & Podsakoff, 2003). In order to test the proposed formative model, the formative constructs must be embedded into a meaningful structural model—that is, a nomological network—that defines the formative construct. Testing the index model of social participation therefore requires consideration of the nomological network in which participation constructs derive their meaning.

Based on the results of prior research and theoretical considerations, a mediational model where social participation influences life satisfaction both directly and via the mediator of perceived social connectedness was hypothesized (Figure 3). Life satisfaction, a global cognitive assessment of one's life quality, represents one important component of subjective well-being, along with positive and negative affect, where affect comprises moods and emotions (Diener, Suh, Lucas, & Smith, 1999). While subjective well-being has been shown to be related to stable personality traits and is itself somewhat stable over time, there is evidence that life events predict subjective well-being, including life satisfaction, beyond the effects of personality (Headey & Wearing, 1989). Herero and Extremera (2010) proposed that social activities mediate the relationship between personality and subjective well-being among older adults and found evidence consistent with such a model. This suggests that social participation has a causal influence on life satisfaction and other components of subjective well-being. Social participation may act partially via an increased sense of social connectedness; activities such as volunteering, club participation, sports, and domestic hobbies bring people into contact with others who can provide a sense of belonging (Aquino, Russell, Cutrona, & Altmaier, 1996; Newsom & Schulz, 1996).

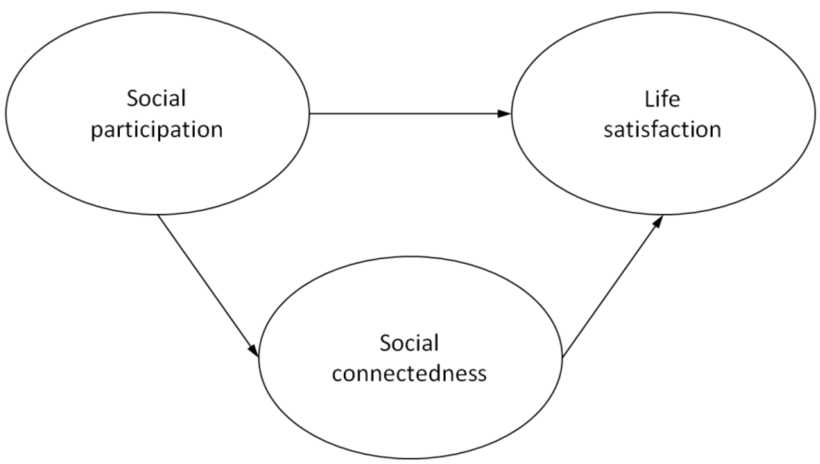


Figure 3. Mediational model of relationship between social participation and life satisfaction

While the actual relationships between life satisfaction, social participation, and perceived social connectedness (controlling for personality, demographic, and other important covariates) are almost certainly much more complex than the simple mediational model proposed in Figure 3, this model offered a theoretically and empirically-defensible entry point for thinking about social participation insofar as it improves someone's well-being and for comparing reflective measurement of participation to formative measurement. It provided a means of estimating the formative model, which requires outcome constructs for identification purposes. It also made it possible to compare the predictive validity of both models.

Formative measurement models have not been widely used in psychosocial research, given its emphasis on measuring latent psychological constructs which serve as the prototypical setting for deploying the common factor model. They have, however, been used with some success in marketing research (e.g., Brock & Zhou, 2005; Bruhn, Georgi, & Hadwich, 2008; Collier, 2006; Johnson, Bruner, & Kumar, 2006; Lin, Sher, & Shih, 2005; Pavlou & Gefen, 2005). Researchers have debated the usefulness and validity of formative measurement models due to practical and philosophical concerns (Bagozzi, 2007; Bollen, 2007; Borsboom, Mellenbergh, & van Heerden, 2003; Edwards & Bagozzi, 2000; Howell, Breivik, & Wilcox, 2007; MacKenzie, Podsakoff, & Jarvis, 2005). One important practical concern is that of ensuring that formative models are identified, that is, that their parameters can be uniquely estimated (Bollen & Davis, 1994/2009). It is not within the scope of this project to address the philosophical critiques of formative measurement.

Statement of the Problem

To understand the aging experience, researchers and policymakers need to understand and usefully measure social participation, then incorporate social participation measures into theoretically meaningful and empirically grounded models. Social participation is a good of its own, bringing enjoyment, engagement, and meaning to a person's life. It is associated with increased life satisfaction and also with an increase in social connectedness, itself an important predictor of well-being, both physical and mental. Unfortunately, as people age, the possibility of meaningful participation may be restricted due to health problems, age-related disability, or social expectations around what is and isn't acceptable during a certain life stage. Social participation is therefore a key outcome to consider in interventional or observational studies that investigate successful aging, whether in the presence of disability or not.

A number of instruments exist to measure participation but researchers have not agreed on what activities should be included or how the measures should be developed and refined. Typically, researchers consider levels of participation in different types of activities as composing the level of overall participation a person achieves (as in a formative approach), but such an approach does not conform to CTT-based scale development techniques, which assume reflective measurement. Formative approaches to measurement models that might be used to develop index measures suffer from identification problems and questions about their epistemological grounding. Can formative measurement modeling be made useful in constructing an empirically and nomologically valid measure of social participation?

Purpose of the Study

The purpose of this study was to develop two models of social participation in community-dwelling U.S. adults over age 65 from self-reported levels of participation in different activities using the reflective (scale) and formative (index) approaches, then to compare the results based on empirical fit, predictive validity, and parsimony. The study demonstrated the use of techniques for constructing index measures using formatively-defined composite variables in structural equation modeling and compared the performance of the index model to the scale model. Additionally, it provided evidence about whether social participation is best modeled unidimensionally or multidimensionally and offered guidance as to what activities should be included in the definition of social participation.

Research Questions

The research questions to be answered by this study were:

(1) What measurement model for social participation has stronger validity: reflective (scale) or formative (index)? There are at least two types of validity at issue: first, how well does the model capture a researcher's conception of social participation (content validity) and second, how well does the model predict outcomes (predictive validity)?

(2) What dimensions of social participation should be modeled? Is social participation unidimensional or multi-dimensional, and if multi-dimensional, how many dimensions must be used to capture the full range of the construct?

(3) What activities should be included in the definition of social participation? Is a broad and deep set of activities required to fully capture it, or can just a few key activities be used?

Literature Review

The following review includes eight major sections. The first section reviews the concept of social participation and its importance in aging and disability research. The second section reviews research about the relationships among social participation, life satisfaction, and social connectedness. The third section discusses varying definitions of social participation. The fourth section presents approaches to modeling and measuring social participation, including ad hoc measures and formal participation measurement instruments. The fifth and sixth sections cover relevant data analytic techniques; first, there is a discussion of conventional two-step structural equation modeling and then the formative measurement approach is presented, including challenges and critiques. The seventh section discusses criteria for choosing between reflective and formative approaches, using social participation as an example. Finally, the eighth section describes how traditional techniques such as canonical correlation might be used in the development of index models and discusses the advantages and disadvantages of using SEM instead.

The importance of social participation. A key focus of older adults in the United States is how they will engage meaningfully with life, whether through paid work or other avenues such as volunteering with nonprofits, joining social or other clubs, pursuing hobbies and recreation, or connecting with people online. I will call this meaningful engagement *social participation*. While people of all ages seek meaningful social activities, this becomes of increasing concern as people age as they transition out of the workforce and may no longer have ongoing family responsibilities (Harlow & Cantor, 1996). At the same time, financial or health constraints may close off certain avenues of participation, leaving older adults with fewer ways to feel a sense of purpose, meaning, and engagement. Some researchers have suggested that the U.S. lags in providing opportunities for productive work after retirement (Veroff & Veroff, 1980). To the extent that older adults can replace important social roles no longer available to them (paid work, active caregiving of their own children) with other meaningful and social activities (for example, volunteering with nonprofits or pursuing hobbies that bring them into connection with like-minded people), they may be able to better maintain well-being even in the face of aging-related social and health changes (Baker, Cahalin, Gerst, & Burr, 2005).

A person's ability to participate, their level of participation, and their subjective experience of participating are all considered important outcomes in disability and rehabilitation research (Dijkers, 2010; Noreau et al, 2004; Reistetter et al., 2005; Whiteneck, 2010). Rehabilitation researchers Noreau et al. (2004) called for more attention to social participation, saying it "is one of the areas that deserves much more attention as it is increasingly considered a pivotal outcome of a successful rehabilitation" (p. 346). Reistetter et al. (2005) called *community integration*, which relies upon a foundation of social participation, "the premier goal of rehabilitation following brain injury" (p. 139). Dijkers (2010) noted that "participation is a key outcome of rehabilitation and of other medical and social service programs supporting persons who, because of impairments resulting from injury, birth defect, disorder, or aging, are involved in family, household, community, and society to a lesser degree than they, their service providers, or society may desire" (p. S5). Participation is seen as an important consequential outcome to consider in the presence of activity limitations; Whiteneck (2010) suggested using a participation measure as a secondary outcome in trials of interventions targeted at reducing activity limitations (p. S57).

Gerontologists concern themselves not just with a lack of disease and disability in older people, but also with the question of what constitutes *successful aging*. Social participation may be a key component of successful aging, which Rowe and Kahn (1997) defined in this way: "Successful aging is multidimensional, encompassing the avoidance of disease and disability, the maintenance of high physical and cognitive function, and *sustained engagement in social and productive activities*" (p. 433, emphasis added). Accurately defining, measuring, and modeling social participation is therefore key to understanding older adults' experience of aging and to ensuring that it is positive, whether a person is challenged by disability or not.

Participation, life satisfaction, and social connectedness.

Participation and subjective well-being. Prior research has suggested that a variety of activities predict well-being in older adulthood (e.g., Baker, et al., 2005; Hao, 2008; Harlow & Cantor, 1996; Herero & Extremera, 2010; Hinterlong, Morrow-Howell, & Rozario, 2007; Wahrendorf, et al., 2008; Warr et al., 2004). Social activities appear to be among the most important in predicting well-being. Harlow and Cantor (1996), considering eight clusters of types of participation, found that life satisfaction in late adulthood was predicted by social activities, mass communication use, and community service activity. Herero and Extremera (2010) proposed a mediator model of the relationship between personality and subjective well-being among older adults, finding that among a range of activities including social activities, mass communication use, and home hobbies, only social activities partially mediated the relationship between personality variables and well-being. In a study considering British adults between 50 and 74 years of age, family and social activities were the most significant predictors of well-being across a range of different types of activities (Warr et al., 2004).

The findings of a relationship between social participation and well-being do not hold without qualification. Levasseur, Desrosiers, and Noreau (2004) called social participation a "restricted determinant of quality of life" (p. 1211), because it was positively but only weakly correlated with quality of life in older adults with physical disabilities. Wahrendorf et al. (2008) found that socially productive activities are associated with higher health and well-being only when people experience autonomy and perceived control in those activities. Caregiving in certain situations has been shown to be associated with poorer health and lower life satisfaction (Hinterlong et al., 2007), but volunteering, a role with higher status than informal caregiving, has been shown to be associated with less anxiety and higher life satisfaction (Hao, 2008).

Social connectedness as a mediator. Social activities may act on well-being by increasing a person's experience of social connectedness and social support, which have been shown to be associated with various measures of subjective well-being (see, for example, Aquino et al., 1996; Baker et al., 2005; Lubben & Gironda, 2003; Morrow-Howell, 2010; Newall et al., 2009; Newsom & Schulz, 1996; Rook, 1987). Improved social ties can be conceptualized and measured in a variety of ways; the current study uses the opposite of loneliness as the social outcome construct of interest. Peplau and Perlman (1982) defined loneliness as "a deficiency in one's social relationships that is subjectively experienced as unpleasant" (cited in Newcomb & Bentler, 1986, p. 520). The opposite of loneliness has been defined as "embeddedness," (de Jong Gierveld & van Tilburg, 2006, p. 582) and as "feelings of belongingness" (de Jong Gierveld & Kamphuls, 1985); here I call it "perceived social connectedness" or just "social connectedness." This affective sense of being embedded in a network of social ties is correlated with social support, which has been more heavily studied in its relation to well-being. Social support can be defined as "the existence or availability of people on whom we can rely, people who let us know that they care about, value and love us" (Sarason, Levine, Basham, & Sarason, 1983, p. 127). Newcomb & Bentler's (1986) confirmatory factor analysis of loneliness and social support scales found evidence consistent with a higher-order factor of “general social attachment” giving rise to both loneliness and social support factors. Especially when measured subjectively, a person’s experience of social support may overlap considerably with their experiences of loneliness; either of these constructs may thus serve in the proposed model as a mediator of the relationship between social participation and life satisfaction.

Activity theory suggests that social activity can help older adults maintain health and happiness as they age (Lemon, Bengtson, & Peterson, 1972); research evidence supports this hypothesis. Aquino et al. (1996) modeled the relationship between volunteer work and life satisfaction and found evidence for social support as a mediator between the two. They did not find that social support mediated the relationship between work status and life satisfaction; paid employment appeared to work directly (or at least not via social support) to increase life satisfaction. Social support and related social attachment constructs have also been shown to mediate the relationship between functional status and quality of life; higher social support was associated with decreased effects of physical disability on quality of life (Newsom & Schulz, 1996). Despite the variety of ways of conceptualizing and measuring social support, researchers have consistently found a statistically significant relationship between social support and well-being (Lubben & Gironda, 2003). To the extent that adults in late life participate in social activities, they are likely to increase their social connectedness, and thus improve their well-being (Cornwell, Laumann, & Schumm, 2008). These varied findings suggest that the hypothesized model of the relationship between life satisfaction and social participation as mediated by perceived social connectedness has empirical grounding.

This model therefore can serve as a useful and valid tool for identifying the formative model of social participation and for evaluating the predictive performance of both the reflective and formative models.

Defining participation. As with many constructs considered in psychosocial and health research, there is not widespread agreement on the definition of social participation as an outcome or predictor (Levasseur et al., 2004). The question of what it means and how to measure it has been debated most extensively within the disability and rehabilitation research communities, partly due to the evolution of conceptual disability models that has occurred in the last forty to fifty years in that arena. During the 1960s, Saad Nagi proposed a social model of disability that identified the importance of social integration and social inequality in understanding limitations on the disabled (Noreau et al., 2005). In this model, *disability* "refers to social rather than to organismic functioning" (Nagi, 1991, p. 315) and is defined as the "limitation in performance of socially defined roles and tasks within a sociocultural and physical environment" (Nagi, 1991, p. 322). The key to full participation in this model is the acting out of social roles (Jette, Haley, & Koohoomjian, 2003).

In psychosocial research, the distinction between simple activity and social participation has also been recognized. In studying the relationship between psychological well-being and activity in older people, Warr et al. (2004) excluded "routine maintenance activities," instead focusing on more voluntary behaviors "that might be expected to yield rewards that can bear upon psychological well-being" (p. 172). Harlow and Cantor (1996) used cluster analysis to reduce level of participation in 33 activities into eight domains: social activities, mass communication use, building knowledge, home activities and hobbies, creative activities, activities outside of the home, community service activities, and games. Most or all of the activities considered by Harlow and Cantor would seem to qualify as activities that might “yield rewards that can bear upon psychological well-being,” to borrow Warr et al.’s (2004) terminology. Their focus was on participation rather than on simple activity.

Formal participation instruments use varying sets of activities to capture a person’s level of participation. The Frenchay Activities Index (FAI) includes items such as gardening, gainful work, and outings/car rides but does not include items about volunteering or religious attendance (Schuling et al., 1993) while the Participation Objective, Participation Subjective (POPS) instrument includes volunteering and religious activities in addition to doing yard work, working for pay, and driving or riding in a car (M. Brown, 2006). The Maastricht Social Participation Profile (MSPP) includes organized volunteer work, eating out, organized day trips, and offering practical help to acquaintances, but does not include paid employment or gardening (Mars et al., 2009). The Community Integration Questionnaire (CIQ), which covers a broader construct than social participation, asks respondents about work, school, and volunteer work but not informal caregiving, religious attendance, or gardening (Sander et al., 1999). None of these four instruments includes items capturing participation in educational activities. Since lifelong learning has been shown to be associated with improved self-confidence, self-esteem, and satisfaction with life in older adults (Dench & Regan, 2000), these instruments may not fully represent the breadth of participation possibilities.

Many psychosocial studies of the relationship between participation and well-being among older adults have limited their study to *productive participation* (e.g., Baker, et al., 2005; Hinterlong et al., 2007; Wahrendorf et al., 2008), defined by Baker et al. (2005) as those activities that benefit others, include a social component, and are perceived as meaningful by the individuals that engage in them. Hinterlong et al. (2007) examined associations between "productive engagement" and physical and mental health using a nationally representative sample of adults over age 60. They considered five productive roles: paid worker, irregular paid worker, unpaid volunteer, caregiver, and provider of informal social assistance. Baker et al. (2005) included paid work, formal volunteering, caregiving, informal helping, and do-it-yourself activities in their definition of productive participation, but did not include religious attendance. Hao (2008) considered the relationships of paid work and volunteering to maintenance of mental health in later life, using only those two types of participation in his definition of productive activities. Jung, Gruenewald, Seeman, and Sarkisian (2009) assessed the relationship of engagement in productive activities to the development of frailty in older adults using level of involvement in volunteering, paid work, and child care. Wahrendorf et al. (2008) considered three types of "socially productive activities," volunteer work, informal help, and caregiving, in a study of how autonomy and perceived control in such activities influenced their relationship with well-being.

Measuring participation. Whiteneck (2010) noted that "interest in the measurement of participation has increased exponentially over the last three decades, with over 30 instruments purporting to measure participation now appearing in the literature, but without any agreement on the most appropriate method of measurement, let alone consensus on a widely applicable psychometrically sound specific assessment tool" (p. S54). Indeed, there exist a wide variety of approaches and instruments used in measuring participation, for a variety of purposes. Psychosocial researchers tend to use ad hoc measures rather than validated instruments, while the disability and rehabilitation research community has developed a variety of instruments that measure participation or participation limitations in various ways.

Perhaps the simplest way to include participation in a quantitative analysis as a predictor or outcome variable is to use single indicator variables of different types of participation. An example of the single indicator approach is found in Hao (2008), where the relationship between productive activities and psychological well-being among older adults was modeled longitudinally with binary time-varying indicators of paid work and volunteer status. The benefit of using such a model is that it allows independent investigation of the consequences of different kinds of participation. The drawback is that it doesn't allow investigation of whether there may be some common mechanism through which participation, broadly conceived, influences well-being; this makes for a less parsimonious model. Morrow-Howell (2010) noted the importance of analyzing interactions and overall patterns of participation: "To date, most studies of co occurring activities have focused on productive activities, excluding leisure, religious, or social activities. Yet these activities are likely important in the balance that maximizes outcomes for the individual. The empirical issue of how to assess and analyze multiple activities and patterns remains a challenge" (p. 464).

Ad hoc measurement of participation. Psychosocial researchers have tended to use ad hoc participation measures developed on a one-off basis for their individual studies that summarize the level (and possibly also diversity) of participation across multiple types. For example, Hinterlong et al. (2007) recorded number of roles among five types of roles including working, volunteering, and caregiver roles as well as amount of time spent in each role, then used those as predictors in their longitudinal model of how productive engagement is associated with physical and mental health. This approach of counting number of roles is fairly common in the literature and reflects the hypothesis of role theory that taking on multiple roles may improve well-being (Morrow-Howell, 2010). Harlow and Cantor (1996) reduced level of participation in 33 activities into eight domains using cluster analysis, then computed eight cluster index scores for domains such as “social activities” and “community service activities” and used these index scores in regression models. In a study of how social participation relates to loneliness among adults over age 72, Newall et al. (2009) tallied the number of social activities respondents had participated in during the past week. The activities include items such as church-related activities, doing community volunteer work, or visiting friends. Then they simply summed the number of activities the respondents engaged, creating one social participation score for each person.

Participation instruments. A variety of formal participation measurement instruments have been developed within the context of rehabilitation and disability research. Some instruments that call themselves participation instruments actually measure participation limitations; instruments in this category include the Keele Assessment of Participation (Wilkie, Peat, Thomas, Hooper, & Croft, 2005), the ICF Measure of Participation and ACTivities (IMPACT; Post et al., 2008), and the Assessment of Life Habits (LIFE-H; Noreau et al., 2004). While these are appropriate as measures of outcomes in disability research, measures that address overall participation levels rather than limitations will be more useful in the present psychosocial research context.

The Maastricht Social Participation Profile (MSPP; Mars et al., 2009) measures actual social participation according to a definition developed by older adults with chronic physical illness. It includes four sub-indexes: consumptive participation, formal social participation, informal social participation with acquaintances, and informal social participation with family. The MSPP provides diversity and frequency scores for each sub-index, with the diversity score representing the “number of items on which a respondent had a score of at least one” (Mars et al., 2009, p. 1209). The developers of the MSPP did not compute scale reliabilities or use exploratory or confirmatory factor analysis since they considered participation to be composed of observed levels of participation rather than reflecting an underlying latent participation construct (Mars et al., 2009).

The Participation Objective, Participation Subjective scale (POPS) is similar to the MSPP in the approach taken to its development. It was also developed assuming a causal indicator model rather than a typical CTT-based reflective indicator model (M. Brown et al., 2004). In addition to gathering information about activity level in a number of different areas, the POPS also gathers a person’s subjective opinions of their participation, so provides both an objective and subjective measure of participation. The questionnaire asks the respondent about their desired activity level as well as how important that particular activity or participation type is to them (M. Brown, 2006). Typically, participation has been measured objectively, using some sort of inventory of a person’s engagement in different types of activities conceptualized as representing participation, but there is a recognition that participation may also be measured subjectively (Noreau et al, 2005). Because choice of what activities to participate in and what level to engage in depends on personal preference, participation instruments may be enhanced by the collection of subjective indicators. Subjective measures "assess a person's satisfaction with participation rather than actual performance" (Whiteneck, 2010, p. S55).

In contrast to the MSPP and the POPS, the Frenchay Activities Index (FAI) was developed using CTT (Mars et al, 2009). It is composed of 15 items, with three subscales: domestic, leisure/work, and outdoors (Schuling et al., 1993). It does not specifically ask about volunteering or informal caregiving, but includes the item "gainful work." A principal-components analysis of validation data gathered from stroke patients and control respondents age 65 and older found two factors in the data, suggesting that the FAI could be refined by using two subscales, the original domestic subscale and an outdoor subscale, eliminating the gainful work item (Schuling et al., 1993). The "gainful work" item did not load on either of these factors. This instrument, should it be refined as suggested, would not capture an important kind of participation—that achieved in productive activities such as employment, volunteering, and caregiving—but it does capture participation at a higher level than in-home activities of daily life.

The FAI is a good example of the misfit of the common factor model to participation data, because internal consistency reliability as measured by Cronbach's alpha was generally inadequate in Schuling et al.'s (1993) study of stroke patients compared to a control group. Reliabilities for the domestic scale were acceptable, ranging from .82 for the control group to .88 for the stroke group measured poststroke. The activities on the domestic scale would not generally be called participatory, however, but rather represented what health researchers refer to as instrumental activities of daily living: washing up, washing clothes, and light housework were among them. For the leisure/work domain and outdoors domain, alphas were quite a bit lower, not reaching the .80 cutoff generally used by researchers as representing adequate reliability as measured by alpha (Brown, T., 2006). The leisure/work domain showed reliabilities ranging from .58 to .63 before the “gainful work” item was deleted; eliminating this item raised reliability to .61 for the stroke group prestroke, .65 poststroke, and .69 for the control group. The outdoors domain showed reliability of .55 to .67 before the “reading books” item was deleted; deletion of this item raised the reliability to between .66 and .73, still inadequate.

The concept of *community integration* overlaps substantially with the concept of social participation, and there are two important instruments that specifically measure community integration: the Community Integration Questionnaire (CIQ) and the Community Integration Measure (CIM) (Reistetter et al., 2005). The CIQ, considered an industry standard within brain injury research (Reistetter et al., 2005), takes the objective approach to measuring community integration, with items such as “approximately how many times a month do you usually visit your friends or relatives?” and “in the past month, how often did you engage in volunteer activities?” (Willer, Rosenthal, Kreutzer, Gordon, & Rempel, 1993). These CIQ items are organized into three dimensions: home integration, social integration, and productive activity (Willer et al., 1993). The CIM takes a subjective approach with ten items asking about respondents’ perceived connections within their communities (Reistetter et al., 2005). The CIM includes items that seem related to participation, such as “I have something to do in this community during the main part of my day that is useful and productive,” and others that address community integration in a broader sense, such as “I feel that I am accepted in this community” (Reistetter et al., 2005). In a validation study with 91 participants (51 with brain injury and 40 without), the CIM showed a statistically significant correlation of .34 with the CIQ, indicating they shared 12% of variance (Reistetter et al., 2005). This seems rather low and suggests that the objective approach and subjective may not be targeting the same underlying construct.

Structural equation modeling. This study used conventional structural equation modeling techniques based on reflective measurement to test and refine a set of participation measurement scales. It also explored the use of a *formative measurement model* for participation, in which observed levels of participation in different activities were modeled as composing the participation construct rather than reflecting the construct. SEM is a family of statistical techniques that allows for confirmatory and exploratory modeling of the complex inter-relationships among latent variables, their observed indicators, and additional observed variables (Bollen, 1989). SEM has its roots in early twentieth century work by Charles Spearman, who developed techniques now known as exploratory factor analysis (EFA), and by Sewall Wright, who developed the basics of path analysis (Kline, 2005). In the early 1970s, the measurement techniques of factor analysis and the structural techniques of path analysis were brought together by K.G. Jöreskog, J.W. Keesling, and D.E. Wiley into a framework known at one time as the JKW framework and now called structural equation modeling (Kline, 2005).

Kline (2005) identified seven characteristics commonly seen in structural equation models: (1) they are a priori (not generated from the data but generated before looking at the data), (2) they distinguish latent (unobserved) from observed variables, (3) their basic statistic is the covariance, not the mean, (4) they can be used with both non-experimental and experimental data, (5) they can be used to represent many standard statistical procedures such as multiple regression and ANOVA, (6) their estimation and related inference tests are based on large-sample asymptotic theory, and (7) the role of statistical tests is less important in SEM than in more traditional techniques.

To some researchers, SEM represents the vanguard of an epistemological revolution taking place in data analysis: the transition from an overarching focus on null hypothesis significance testing (NHST) to a focus on modeling the world mathematically and probabilistically (Rodgers, 2010). The typical use of SEM requires the researcher to develop a fully-conceived and theoretically grounded view of the phenomenon at issue (as Kline, 2005, said, the models are a priori), and then provides the researcher with statistics (in the form of chi-square tests and goodness-of-fit indexes) that evaluate the model as a whole. Individual statistical significance tests such as might be undertaken with statistical techniques such as post-hoc mean comparisons in ANOVA are of less concern than overall model fit in an SEM study (Kline, 2005). The purpose is to develop a theoretically defensible and empirically validated model, not to reject one null hypothesis or another. In this tradition, the researcher asks if the model "works to achieve its goals… compared with other models that are reasonable competitors" (Rodgers, 2010, p. 4). SEM, more than other techniques, encourages the researcher to consider a range of competing models, all based on a comprehensive theoretical and practical understanding of the world.

The two-step approach to SEM. A structural equation model includes both a measurement model (the factor model) and a structural portion (akin to a path model, but with causal relationships specified among latent rather than observed variables). In the commonly-used two-step modeling approach as formulated by Anderson and Gerbing (1988), the full structural regression model is first transformed into a confirmatory factor analysis (CFA) model, with all directional relationships specified as unanalyzed (i.e., non-directional) associations. This measurement model is estimated to see if good fit is achieved and if so, the structural model is estimated and evaluated. Extensions and alternatives to this approach have been suggested (see, for example, Mulaik and Millsap, 2000, for a four-step approach), but this distinct separation of measurement model testing from structural model testing is quite common and well-established. When formative constructs are introduced, the two-step process is no longer feasible, so alternative approaches must be devised.

Identification of structural equation models. A key task in structural equation modeling is ensuring that the specified model is *identified*, and ideally, overidentified. More detailed discussion of identification rules and issues can be found in any structural equation modeling textbook such as Bollen (1989) or Kline (2005) as well as in chapter two of this study; the topic is covered only briefly here so as to set the context for the identification issues that arise in formative measurement modeling. A model is identified if it is in theory possible to arrive at a unique estimate for each free parameter (regression weight, variance, and covariance) in the model (Kline, 2005). Structural equation models comprise a set of linear equations to be estimated using observed variances and covariances. The first requirement for identification is that there must be more observations than parameters to be estimated (Kline, 2005). This just represents the mathematical requirement in solving a system of linear equations that there must be more knowns than unknowns; when there are not, there will be many solutions to the system of equations. No unique solution will be available. The second requirement for identification is that each latent variable must be assigned a scale so that estimates of effects involving latent variables can be calculated (Kline, 2005). There are additional concerns as well in achieving theoretical model identification, including restrictions on patterns of reciprocal causality and correlated disturbances; these are discussed in chapter two.

An identified model may be *just-identified* or *overidentified*. A just-identified model has the same number of parameters as observations and, assuming it meets other identification requirements such as having a scale assigned to each latent variable, can be estimated, but does not allow for model fit evaluation, discussed below (Kline, 2005). An overidentified model has more observations than parameters and provides for the calculation of fit statistics (Kline, 2005). Generally, researchers evaluate overidentified models.

Confirmatory factor analysis. The purpose of CFA is to account for the variance and covariance of a set of observed indicators using a set of common, unobserved factors (T. Brown, 2006). This is known as the *common factor model* (T. Brown, 2006). The common factors are considered *latent variables*, which are defined by Bollen (2002) as random or nonrandom variables "for which there is no sample realization for at least some observations in a given sample" (p. 612). Of course, in the case of factors, there are no sample realizations for any observations in a given sample; the common factors are always unobserved. CFA, like its cousin *exploratory factor analysis* (EFA), is based on Thurstone's (1947, as cited in T. Brown, 2006) common factor model, which says that each observed indicator is a linear function of one or more common factors and a factor unique to that indicator, representing measurement error or other unique influences on the indicator. The variance is partitioned into *common variance* (that variance due to the latent factor, estimated by considering variance shared with other indicators specified to load on that factor) and *unique variance* (variance specific to that single indicator, whether systematic or random) (T. Brown, 2006). An example of a single-factor CFA model is shown in Figure 4. This can be expressed mathematically as:

$$y_i = \lambda_{i1}\eta_1 + \epsilon_i$$

where η_1 is the latent factor, y_i is the *i*th observed indicator, e_i is the unique error associated with the *i*th indicator, and λ_{i1} is the coefficient expressing the latent variable's expected effect on the indicator (Bollen & Lennox, 1991). In this model, the indicators are called *reflective*, because they reflect the underlying value of the latent factor (Diamantopoulos, Riefler, & Roth, 2008). They are also sometimes known as *effect indicators*, because they are specified as effects of the latent factor η_1 (Bollen & Davis, 1994/2009).

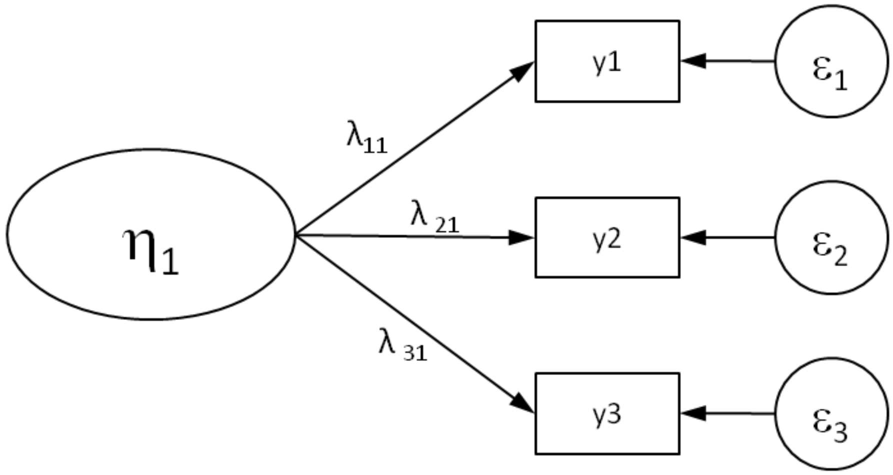


Figure 4. Single-factor CFA model (reflective measurement model)

CFA models often include more than one factor, as when a researcher uses CFA to test the measurement model for a structural equation model using the two-step model evaluation approach suggested by Anderson and Gerbing (1988). In this case, the researcher includes all latent variables that appear in the overall model. If each indicator depends on just one factor and the error terms for each indicator are independent of one another, then this is considered a unidimensional measurement model (Anderson & Gerbing, 1988). Typically, researchers will start with a unidimensional model, and then consider adding correlated errors or allowing indicators to load on more than one factor if adequate fit is not achieved (Kline, 2005). This *model respecification* may be guided by theoretical considerations, as when indicators for items with negative wording are correlated on the assumption of some shared method variance, but often researchers inspect empirical *modification indexes* or *correlation residuals* to determine how to change the model. Modification indexes identify which changes will bring the largest improvement in model fit while correlation residuals identify areas of poor fit in the model (Kline, 2005).

Structural regression. After an adequate measurement model is developed, the researcher specifies directional relationships between the latent factors, and may also include single indicator variables in the model, such as when controlling for demographic covariates. If there is evidence of adequate model fit, this suggests that the specified structural regression model is consistent with the empirical data, but it does not prove that the relationships as specified represent causal relationships. The researcher should also consider *equivalent models*, which have different paths among the variables but imply the same predicted covariances (Kline, 2005). As in CFA, the researcher may respecify the structural regression model, using modification indexes or residual correlations to guide the changes.

If an acceptable measurement model was established, then the factor loadings for indicators in the structural regression model should show only slight changes when various SR models are tested. If the factor loadings change noticeably under different SR models, this suggests that the measurement model lacks stability, which may be evidence of *interpretational confounding* (Kline, 2005). Burt (1976) defined the problem of interpretational confounding as occurring when the "assignment of empirical meaning to an unobserved variable … is other than the meaning assigned to it by an individual a priori to estimating unknown parameters" (p. 4). In such a situation, Burt said, "Inferences based on the unobserved variable then become ambiguous and need not be consistent across separate models" (p. 4).

Assessing model fit. In both CFA and SR, a variety of fit statistics and indexes are used to assess whether adequate model fit has been achieved. A few commonly-used ones are described here, but there exist many more that may be used in SEM.

Chi-square goodness-of-fit test. The model chi-square statistic tests the researcher's overidentified model against a comparison just-identified model in which all observed variables are considered correlated (Kline, 2005; Bentler & Bonett, 1980). A significant chi-square value rejects the null hypothesis that the researcher's model fits perfectly in the population (Kline, 2005), so the researcher is looking for a nonsignificant chi-square value. The chi-square fit statistic is problematic because the null hypothesis of perfect fit in the population is unrealistic and unlikely to ever actually hold (Kline, 2005); the modeling approach to statistics says that our models match reality in some important ways but are also simpler than reality (Rodgers, 2010). The chi-square test statistic also suffers from a second weakness: it is directly dependent on sample size. This means that the probability of rejecting the researcher's model increases with the number of cases considered, even when the model is "minimally false" (Bentler & Bonett, 1980, p. 591). In order to reject any overidentified model based on the model chi-square, all the researcher must do is gather a large enough number of cases (Kline, 2005).

The problems with the chi-square goodness-of-fit test have led SEM methodologists to develop a variety of approximate fit indexes that can guide applied researchers towards models that represent interesting simplifications of reality but still reflect empirical reality. Fit indexes may be categorized as absolute or incremental; *absolute fit indexes* quantify how well the researcher's priori model reproduces sample data while *incremental fit indexes* compare the improvement in fit achieved by comparing the researcher's model with some more restricted and nested baseline model, typically an independence model that specifies no relationship among observed variables (Hu & Bentler, 1999). Some of the more popular fit indexes are known as *parsimony-adjusted*;

this means they take into account how many parameters are estimated (Kline, 2005). Generally, more parsimonious models are favored, assuming they explain the data as well as models with more parameters, on the basis that they are more elegant and simple.

Root mean square error of approximation (RMSEA). The RMSEA is a parsimony-adjusted absolute fit index based on the noncentral chi-square distribution, implying that there is no assumption that the researcher's model fits perfectly in the population (Kline, 2005). For RMSEA, smaller is better, with higher values representing poorer fit. Hu and Bentler (1999) proposed that an RMSEA cutoff around .06 is appropriate, based on a simulation study that considered CFA models only. Another common rule of thumb is to consider RMSEA less than or equal to .05 as good, RMSEA between .05 and .08 as adequate, and values greater than or equal to .10 as poor or inadequate fit (Browne & Cudeck, 1993).

Comparative fit index (CFI). The CFI is an incremental fit index that is scaled to have a value between zero and one, with higher values representing better fit (Kline, 2005). Like the RMSEA, the CFI is noncentrality-based and therefore does not assume a perfect population fit for the researcher's model (Hu & Bentler, 1999). Hu & Bentler's (1999) simulation study suggested a cutoff value close to .95 for the CFI.

Standardized root mean square residual (SRMR). The SRMR measures the mean absolute correlation residual, "the overall difference between the observed and predicted correlations" (Kline, 2005, p. 141). Like the RMSEA, it is a badness-of-fit index, with higher values representing poorer fit. Researchers generally look for SRMR less than .10 (Kline, 2005); Hu & Bentler (1999) suggested a cutoff of .08 for this index.

Formative measurement modeling. The vast majority of measurement in the social sciences assumes effect indicators, where observed variables reflect the level of an underlying latent variable (Bollen, 2002). Some constructs may be better measured with *causal indicators*, that is, observed variables that directly affect, or in some sense cause, the levels of the latent construct, as shown in Figure 5. A variable measured with causal indicators may be called a *composite variable*, since it is made up of observed variables. This type of measurement is sometimes known as *formative measurement*; the latent variable of concern is in some sense formed by other variables.

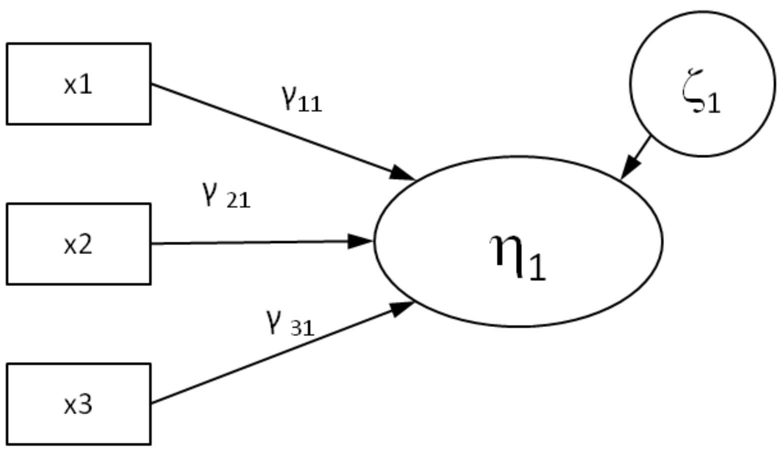


Figure 5. Formative measurement model

The idea of causal indicators is not new; Blalock (1964, as cited in Bollen & Davis, 1994/2009) offered a variety of examples of constructs that might be measured with causal indicators, including exposure to discrimination (caused by observed variables gender and race), socio-economic status (caused by observed variables education, income, and occupational prestige), and frustration in an experimental setting (caused by withholding of food and sleep). Formative measures are not used very often because they are not easy to incorporate into structural equation models and, perhaps because of the difficulty of using them, procedures for their use are not taught in many social science disciplines (Bollen & Davis, 1994/2009). Furthermore, formative measurement has come under philosophical attack for its lack of realist underpinnings; without epistemological justification for its use, researchers may feel reluctant to use it.

The formative model shown in Figure 5 can be expressed mathematically as:

$$\eta_1 = \gamma_{11}x_1 + \gamma_{21}x_2 + \gamma_{31}x_3 + \zeta_1$$

where η_1 is the composite variable, x_i is the *i*th observed variable thought to influence the composite variable, ζ_i represents all unobserved factors influencing the formative variable, and γ_{i1} is the coefficient expressing each observed variable's expected contribution to the composite variable (Bollen & Lennox, 1991). In this model, the indicators are sometimes called *causal*, because they determine the level of the unobserved, composite variable (Diamantopoulos, Riefler, & Roth, 2008). They may also be known as *formative* or *composite indicators* (Bollen & Lennox, 1991). They are distinguished from *effect* or *reflective indicators* used in the common factor model, in which manifest variables are caused by or otherwise reflect the underlying latent factors (Diamantopoulos, Riefler, & Roth, 2008).

One critical aspect of the formative model as opposed to the reflective model is the specification of error. In the reflective model shown in Figure 4, unique errors are specified for each indicator variable; these are labeled ε_1, ε_2, and ε_3. These represent all unmeasured causes of variance in the indicators, including measurement error (Kline, 2005). In the formative model of Figure 5, the error, ζ_1, is specified only at the level of the formative construct. This does not represent measurement error but rather models all unmeasured causes of the formative construct. For example, if one were measuring socio-economic status formatively and included only income, wealth, and where a person lived, but not level of education, education's contribution to socio-economic status would be part of the error term. In this model, the indicators themselves are assumed to be error-free (Edwards & Bagozzi, 2000). Measurement error at the indicator level could be modeled with the use of multiple indicators per conceptual construct contributing to the formative construct, a tactic recommended by Edwards (2010). As with other single-indicator constructs, it is possible and perhaps desirable to specify a fixed amount of measurement error for each formative indicator if multiple indicators per item contributing to the formative construct are not available. Grace and Bollen (2006, 2008) use this approach in estimating a model that includes formative constructs.

Identification of formative models. One of the most difficult problems in formative measurement is achieving identification of the model. A formative construct in isolation (with no reflective indicators or constructs deemed causally posterior to it) does not constitute an identified model. Bollen and Davis (1994/2009) offered a set of rules for judging whether a model including formative constructs is identified. As with all structural equation models, the researcher must ensure that latent variables are scaled and that the number of estimated parameters in a model is less than or equal to the number of elements in the observed covariance matrix. Also, any formative constructs must emit at least two paths, either to observed or latent variables; this is known as the *2+ emitted paths rule* and is necessary but not sufficient in guaranteeing identification (Bollen & Davis, 1994/2009).

One example where the 2+ emitted paths rule is met but the model is underidentified occurs if the formative construct emits paths to two latent variables which are themselves related, either directionally or with correlated disturbances

(Diamantopoulos, Riefler, & Roth, 2008). This particular situation, shown in Figure 6, might be expected to occur regularly in secondary data analysis projects, where a researcher is faced with a construct measured with only formative indicators that is theorized to causally influence two other latent variables, which are themselves related. This is, in fact, the situation confronted in this study: social participation is hypothesized to emit two paths, to social connectedness and life satisfaction, but the outcome constructs are certainly themselves related. Therefore when social participation is specified formatively, the proposed model is underidentified, even though the 2+ emitted paths rule has been met using social connectedness and life satisfaction as outcomes.

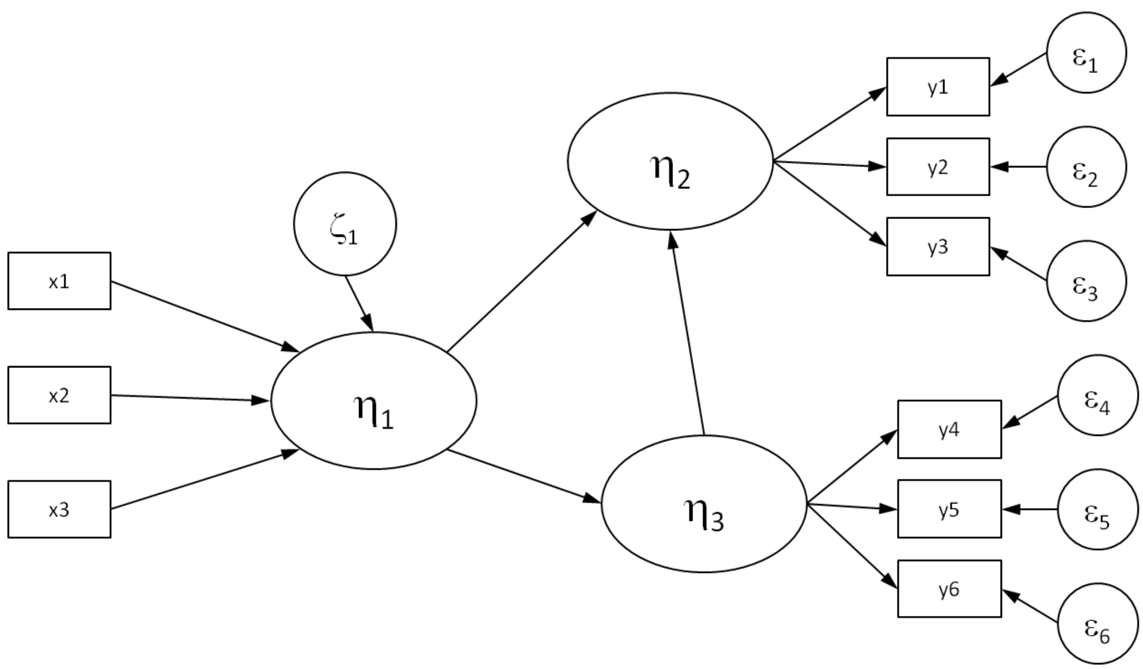


Figure 6. Underidentified model with a formative construct

Franke, Preacher, and Rigdon (2008) note that "unlike reflective constructs, formative constructs mediate the effects of their indicators on other variables, constraining their indicators to have the same proportional influence on the outcome variables" (p. 1230). Indeed, this is the essence of the formative model: that the indicators making up a formative construct work through the same mechanism to influence outcome constructs of interest. The existence of a formative construct in a model thus implies certain proportionality relationships that can be tested even in underidentified models such as the one under consideration in this project (Bollen & Davis, 1994/2009; Franke et al., 2008). Bollen and Davis (1994/2009) suggested two approaches to investigating whether presumed proportionality relationships hold in an underidentified model with formative constructs: setting the variance of the formative construct to zero or imposing proportionality constraints on a transformed version of the model (called the *partially reduced form model*) to check whether they are reasonable or not. The first approach is the one used in this study to develop an index model of participation; the second (use of proportionality constraints with the partially reduced form model) is not described here but details may be found in Bollen and Davis (1994/2009).

Zero error-variance formative models. When the error variance of the formative construct is specified as zero, Grace and Bollen (2008) labeled it a *composite variable* and used a hexagon to represent it (Figure 7), distinguishing it from latent variables measured reflectively or formatively. Conceptually, fixing the disturbance to zero means that the model incorporates all possible causes of the conceptual construct that the researcher would like to measure; this may be unrealistic in many situations (Diamantopoulos, Riefler, & Roth, 2008). However, such an approach does allow consideration of collections of disparate causes thought to have common patterns of influences on outcomes. As such, this zero error-variance formative model approach is consistent with an index model of measurement, which seeks useful combinations of indicators representing multiple facets of a construct. Grace and Bollen (2008) recognized that while composites represent a compromise, they have the advantage of providing increased generality, interpretability, and simplicity in certain modeling situations:

> Without composites, models that consider substantial complexity and/or seek to address relations among a large number of variables, will have the tendency to be highly specific, possessing a separate set of effects from each of a suite of intercorrelated indicators. (p. 210)

Procedures for testing models that include composites have not been well developed like those for specifying and evaluating models with reflectively-measured constructs. Grace and Bollen (2006; 2008) provided a detailed discussion and empirical examples of their use; specific procedures for evaluating such models are discussed in chapter two.

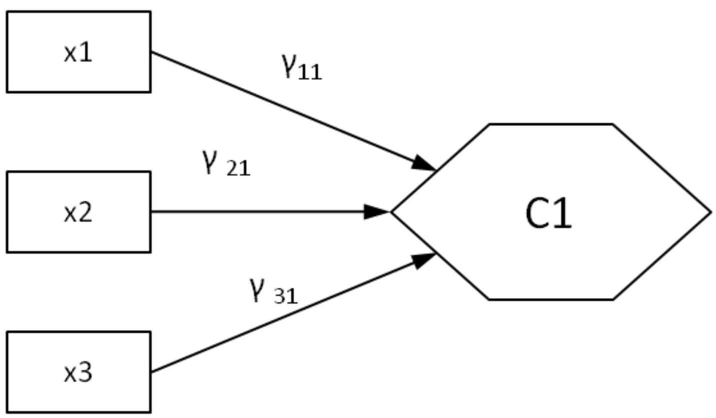


Figure 7. Composite variable

Use of a MIMIC model. An alternative to setting the disturbance of a formatively-defined construct to zero that may be available in some research situations is to add reflective indicators to the formatively-measured construct. The difficulty in some situations is that no reflective indicators may be available. However, this approach offers a way to more fully define a construct and to achieve an identified model that still includes formative measures, which some researchers consider to most accurately reflect certain modeling situations. This is akin to the multiple indicators and multiple causes

(MIMIC) approach described by Jöreskog and Goldberger (1975) and is shown in Figure 8. Diamantopoulos and Winklhofer (2001) suggested that a researcher developing an index follow a four-step process: specifying content scope, specifying indicators to capture the full breadth of the context, screening for indicator collinearity, and then establishing external validity by specifying two or more outgoing paths to outcome constructs or to reflective indicators. Here, external validity is similar to but not synonymous with criterion-oriented validity. If outcome constructs are used to establish external validity, that validity is synonymous with criterion-oriented validity. If reflective indicators are used, this validity seems closer to content validity in nature. The first two steps are conceptual, the second mechanical (looking at correlations and other indicators of collinearity), while the fourth may be accomplished with use of a MIMIC model (Diamantopoulos & Winklhofer, 2001) or by embedding the index model into a broader structural model that includes predicted outcomes.

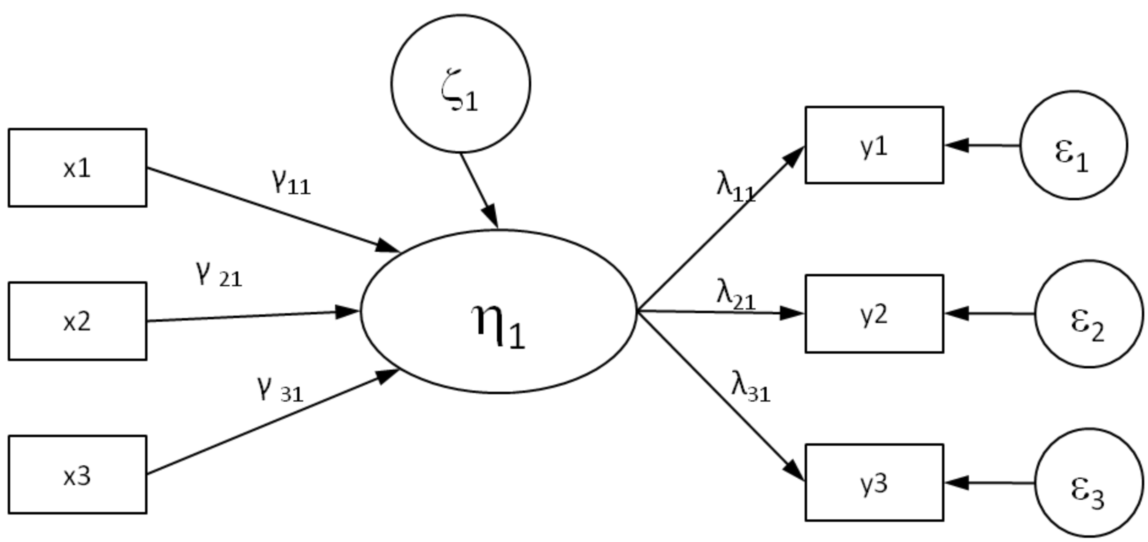


Figure 8. MIMIC-style formative construct

Critiques of formative measurement. Formative measurement has been subject to a number of critiques, on both epistemological and practical grounds. Howell et al. (2007) argued that formative measures are inherently susceptible to interpretational confounding, that is, that causal indicators cannot fully define a construct so that the resultant empirically-defined construct matches the researcher's intended theoretical meaning. In response to this critique, Bollen (2007) replied that interpretational confounding can arise in formative or reflective models and is due to misspecification in underidentified models, not the type of indicator used.

Edwards (2010), comparing formative and reflective approaches, concluded that formative measurement is not viable and recommends that researchers use alternative models based on reflective measurement that achieve the same objectives. He argued that formative measurement is fallacious because a formative construct cannot be a real entity that exists in the world. Edwards' critique is grounded in a *critical realist* view of epistemology, which proposes that there exist real, objective entities out there in the world and further suggests that we can know about those things. An alternative epistemological view, promoted by philosophers such as Charles Sanders Peirce, William James, John Dewey, and recently, Richard Rorty, is *pragmatism* (Hookway, 2010). This philosophy is captured in Peirce's pragmatic maxim:

> It appears, then, that the rule for attaining the third grade of clearness of apprehension is as follows: Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object. (Peirce, 1878)

Pragmatism says that we judge truth by its consequences. On this view, a hypothesized construct need not exist in an objective and independent sense in order for it to play a useful role in a model. A pragmatic philosophy dovetails neatly with the modeling revolution described by Rodgers (2010). Models are simplified views of reality and as such don't reflect reality exactly but rather serve to illuminate it for us.

Both Edwards (2010) and Howell et al. (2007) suggested designing measures using reflective indicators, but in some cases, a reflective version of the construct may not capture the same meaning as a formatively-modeled construct. For example, requesting a person's self-report of their socio-economic status tells us their perceived SES level but may not accurately reflect their objective SES level. Objective SES might be better modeled formatively with indicators like occupational prestige, income, level of education, and place of residence. Another problem with the suggestion that all measures be designed reflectively is that many analysts rely on data that have already been gathered and may not have reflective indicators available for constructs of interest. In that case, the suggested alternative is to eliminate the formative construct and have observed variables influence endogenous reflectively-measured constructs directly. This is not equivalent, since it does not offer a test of the hypothesis that certain of the predictive variables influence outcomes by the same or similar mechanisms, but it does represent an important model specification to consider as an alternative to a formatively-specified model.

Choosing formative vs. reflective approaches. Jarvis, MacKenzie, and Podsakoff (2003) offered a set of questions for judging whether a particular set of indicators should be treated as reflective or formative. These questions fall into four categories: the direction of causality between the construct and the indicators, the interchangeability of indicators, the expected covariance across indicators, and the pattern of antecedents and consequences of indicators. Each of these categories is discussed individually below insofar as they apply to the measurement of social participation using reported of levels of participation in different activities. Chin, Peterson, and Brown (2008) cautioned researchers that although the criteria suggested by Jarvis et al. (2003) are "intuitively reasonable, … it is difficult to meaningfully categorize measurement scales unequivocally as being formative or reflective based on the measurement items alone" (p. 289). Indeed, although some researchers have conceived of social participation measured by reported levels of different kinds of activities as an index (e.g., M. Brown et al., 2004; Dijkers, 2010; Mars et al., 2009), consideration of these criteria does not lead to an unambiguous decision that participation should be modeled formatively.

Direction of causality. Jarvis et al. (2003) suggested researchers should ask themselves, "Would changes in the indicators/items cause changes in the construct or not?" and "Would changes in the construct cause changes in the indicators?" (p. 203). In a formative model, changes in indicators are expected to lead to changes in the level of the formative construct, rather than vice versa, as in a reflective model. On this criterion, a participation construct measured by levels of different types of activities could be conceived of reflectively or formatively. For the reflective conceptualization, social participation could be considered as a kind of latent "participatoriness" or "drive to participate" construct. As this latent psychological construct increased, a person would be expected to increase their activities in various domains, whether in volunteering, sports and recreation, hobbies, or club participation. However, there is also an argument in favor of considering these as formative indicators. As a person increases their participation in one activity (say, volunteering), their overall social participation increases; the change in level of activity drives the change in overall participation, not vice versa. In this conception, participation is just a composite of different kinds of participation.

Interchangeability of indicators. Jarvis et al. (2003) suggested that in a formative model, indicators need not be interchangeable; they need not represent similar content. In a formative model, dropping one indicator may alter the definition of the construct while in a reflective model indicators are theorized to be selected from a domain of interchangeable possibilities (Diamantopoulos & Winklhofer, 2001). On this criterion, social participation seems to be best modeled formatively. Items selected for inclusion in a social participation measure usually will not be interchangeable; dropping one changes the definition of participation. A social participation constructed measured with gardening, golfing, and political campaigning would differ from one measured with volunteering at a dog shelter, attending church or other religious services, and maintaining an online journal. The activity items do not seem to be sampled from a universe of interchangeable possibilities.

Covariance of indicators. In a reflective model, indicators are expected to show moderate to high correlation with one another; higher intercorrelation equates to higher reliability as measured by Cronbach's alpha. In formative models there is no expectation of such internal consistency (Diamantolopoulos & Winklhofer, 2001). In fact, in a formative model, multicollinearity may be problematic, contributing to instability in estimating indicator loadings. We would not necessarily expect correlations across different types of activities. Different types of activity may be correlated with each other to the extent they require an able body and a desire for engagement, but most people will choose particular ways of engaging meaningfully and socially according to their preferences and opportunities; this seems likely to limit the correlation across activity types. Someone who helps in caring for grandchildren may feel no need to participate in the community by volunteering while someone who enjoys participating in a social club may hold no religious views that encourage them towards religious attendance. In fact, a high level of participation in one activity may compromise someone's ability to participate in other activities because of a lack of available time. On this criterion, social participation measured by reported levels of participation in different activities appears to be appropriately modeled formatively, not reflectively.

Antecedents and consequences of indicators. Jarvis et al. (2003) proposed that formative indicators "are not required to have the same antecedents and consequences" (p. 203). If the indicators do not have similar consequences, however, one wonders what the point of introducing the formative construct is. In the proposed model of social participation as related to life satisfaction and social connectedness, the purpose of modeling a common construct of social participation is to clarify the relationship between such social participation and life satisfaction, not to explore how different types of activity differentially affect life satisfaction and social connectedness. The different types of activity may, however, have different antecedents: participation in a social club depends on the availability of such a club and perhaps on whether one's friends have joined the club; attendance at a particular denomination of church may depend on one's religious upbringing; caring for one's grandchildren depends upon having had children who now have children of their own, and so forth. That they may have different antecedents argues for a formative model.

Structural equation modeling versus the general linear model. While SEM offers some advantages over more traditional statistical modeling approaches such as ANOVA, regression, and so forth, it also has some disadvantages. It requires specialized software and is usually taught only after a student has completed courses in introductory statistics, experimental design and ANOVA, regression, and multivariate methods. SEM is a large-sample technique; with small samples, estimation problems may be encountered and achieved power to detect an effect may be low (Kline, 2005). More traditional multivariate techniques may therefore be of interest to the researcher who wants to develop indexes given a set of index indicators and multiple outcome measures. This section discusses the general linear model (GLM), which forms the foundation of many traditional multivariate and univariate statistical methods, and describes its potential usefulness in index construction. Canonical correlation analysis (CCA), the most general of the models supported by the GLM, could be used as an exploratory analysis to identify subsets of variables that might usefully compose an index based on their associations with multiple outcome variables.

Many statistical techniques including bivariate correlation, multiple regression, ANOVA, survival analysis, discriminant analysis, and canonical correlation can be expressed via the GLM, which models additive linear relationships between two or more variables (Tabachnick & Fidell, 2007). CCA represents the most general form of the GLM for a set of continuous variables that has been divided into two groups (Tabachnick & Fidell, 2007). CCA produces pairs of composites (weighted linear combinations or *variates*) of each set of variables so that the composites in each pair are maximally related (Thompson, 1984). Depending on the number of variables available, CCA may produce more than one pair of composites, but usually only the statistically significant pairs are interpreted (Tabachnick & Fidell, 2007). CCA answers some of the same questions that the index construction methodology proposed here seeks to answer. Some of those questions that CCA can answer were formulated by Thompson (1984) as follows:

> (1) To what extent can one set of two or more variables be predicted or 'explained' by another set of two or more variables?
>
> (2) What contribution does a single variable make to the explanatory power of the set of variables to which the variable belongs?
>
> (3) To what extent does a single variable contribute to predicting or 'explaining' the composite of the variables in the variable set to which the variable does *not* belong?
>
> (4) What different dynamics are involved in the ability of one variable set to 'explain' in different ways different portions of the other variable set? (p. 10)

Figure 9 shows a graphical depiction of CCA as it might be used in the current project. Participation indicators are entered as one set of variables, considered as independent variables, and outcome measures are entered as the second set of variables, considered as dependent variables. CCA does not require designating one set of variables as independent and one set as dependent, but that can be done if it enhances interpretation. Perfect reliability is assumed for both sets of variables, with the outcome variables of social connectedness and life satisfaction assumed to be entered as scale scores computed as averages of item scores. When CCA is applied to this data, two pairs of canonical variates will be estimated, since the maximum number of pairs is equal to the minimum number of variables in either of the two sets (Tabachnick & Fidell, 2007). For each pair, the weights of the outcome variables onto the outcome variate represent a distinct pattern of outcomes that might be associated with a subset of the participation variables. The weights of the participation items onto the variate in a particular pair suggest which of those items should be included in a particular index. Participation items which contribute little to the variate could be dropped from an index. The weights of the outcome variable suggest the pattern of relationships that holds between a particular index composed of participation variables and the outcomes.

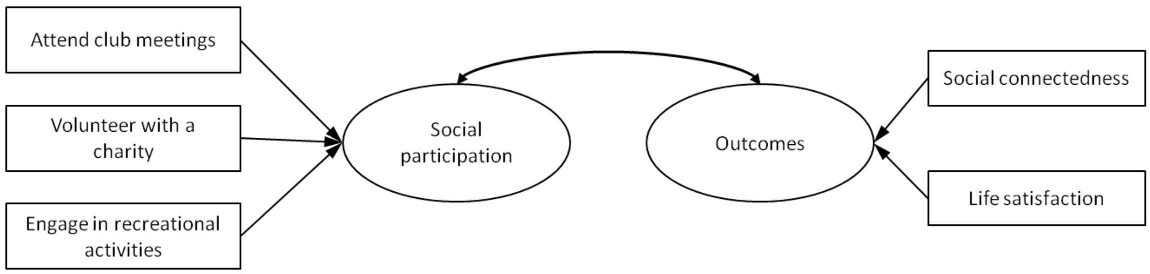


Figure 9. Canonical correlation version of social participation model

While canonical correlation may help a researcher identify indexes in an exploratory fashion, SEM provides the researcher with the ability to specify the structure of the indexes a priori, based on theory and other substantive considerations such as results of prior research, and then conduct a confirmatory analysis. SEM's ability to specify very complex models and statistically test whether these fit the data or not is one of its strongest advantages. In practice, SEM's usefulness as a confirmatory method is compromised by three factors. First, the chi-square test that can definitively reject a model is sensitive to sample size and thus may reject models that usefully describe reality while simplifying it. Instead of the chi-square test, many researchers use approximate fit indexes to evaluate their models but this is considered questionable practice by some SEM theorists (Barrett, 2007; McIntosh, 2007). Second, SEM is often used in a hybrid confirmatory/exploratory mode, where a researcher starts with a hypothesized model but respecifies it based on empirical results such as correlation residuals or modification indexes. Third, SEM can find the best-fitting model among many fitting models but cannot identify the “true” or “best” model. For each model, many equivalent models with equally good fit statistics exist and other, better-fitting models exist too. The use of approximate fit indexes rather than the chi-square test for model rejection, the exploratory respecification of models, and SEM’s lack of ability to identify one best or true model make SEM in practice look similar to practices employed in the use of traditional models such as CCA, which, with the right perspective, can bring a modeling approach to a research problem. Just as exploratory factor analysis (EFA) plays a prominent role in scale development along with CFA implemented within SEM programs, CCA could be an important analytical method in index construction practices, especially if joined with the relatively more confirmatory approaches relying on SEM developed and described here.

Definitions

The majority of conceptual and technical terms are defined in the literature review in this chapter and in chapter two, which covers the methods used in this project. This section defines only a few key terms that might otherwise remain ambiguous.

Conceptual terms. By *social participation,* I mean activity that is meaningful to a person, is voluntarily chosen, and has the potential to bring them into social connection with other people. I construe social participation broadly so as to make possible an elaborated model of social participation that contains multiple dimensions. The specific activities that will be considered under the broad heading of social participation are listed in chapter two.

By *life satisfaction*, I mean a person's overall global assessment of their life quality, judged according to personal criteria (Diener, Suh, Lucas, & Smith, 1985). This particular subjective well-being construct was chosen over measures of moods and emotions (that is, affect), since social activities are hypothesized to have long-term effects on subjective well-being. Long-term effects might be best detected through a cognitive construct such as life satisfaction rather than through emotional constructs such as positive or negative mood.

By *social connectedness*, I am referring to a person's sense of being in relationship with others, of having people they can turn to when they need help, of being embedded in one or more social networks, and of being around people they can relate to. In this study, social connectedness is measured subjectively using positive items from a loneliness scale. There is some question over whether loneliness and its opposite (here, social connectedness) represent two ends of a bipolar scale or are distinct constructs (Russell, 1996) as well as whether social support and loneliness are aspects of one higher-order social attachment factor (Newcomb & Bentler, 1986). This project remains agnostic about such questions, using only positive items from the UCLA Loneliness Scale Version 3 (Russell, 1996) for convenience, to avoid questions of multi-dimensionality, and because incorporating a positive construct (social connectedness) rather than a negative one (loneliness) makes the model both easier to describe and more intuitively appealing.

Technical terms. By *reflective measurement*, I mean the construction and validation of instruments such as questionnaires, tests, and assessments based on classical test theory and the common factor model. In reflective measurement, the latent quantity of interest is modeled as causing the observed indicator variables; these indicators are called *reflective indicators*. By *formative measurement*, I mean the construction and validation of measurement instruments based on causal indicator models, where measured quantities are modeled as composed by observed variables rather than reflecting them. When indicators are specified as causing or composing a variable, they are called *causal indicators* or *formative indicators*. I use the term *formative construct* to mean an unobserved variable that is modeled formatively. Generally, I reserve this term for variables modeled with a non-zero disturbance term. I use the terms *composite variable* or *composite* to refer to variables modeled formatively with zero error specified.

I use Stevens' (1946) definition of *measurement*, "the assignment of numerals to objects and events according to rules" (p. 677). I use this term recognizing that psychometric measurement in the Stevens' (1946) tradition does not often provide interval measures as are used in the physical sciences or might be provided by item response theory techniques such as embodied in the Rasch model described by Bond and Fox, 2007. Scale and index models of constructs may be better thought of as data *summaries* rather than *measures*. It is typical, however, within psychosocial research to call scales or indexes "measures" and to call the use of them "measurement" even when they have not been shown to have interval-level measurement properties. Interval-level measurement such as provided by Rasch models is appropriate and necessary for high-stakes ranking purposes such as test-based accountability in education. For explanatory and predictive purposes such as a better understanding of how social participation is related to well-being, data summaries provided by scale or index models may be both appropriate and practical. Application of item response theory to the measurement of social participation is outside the scope of this project, but could represent an interesting future research project.

Delimitations

The focus of this study was on exploring models for social participation in the context of its relationship with life satisfaction and social connectedness, not on definitively establishing the direction or structure of causal relations between social participation, social connectedness, and life satisfaction. For a variety of reasons, any inference of causality here would be suspect. The study is based on cross-sectional, non-experimental data. Respondents selected their own level of participation; they were not randomized to one level or another and then followed to see what the outcomes were. This means estimation of an effect of participation on well-being may reflect participant differences rather than a causal effect of participation on well-being. Also, the relationships among participation and subjective well-being may very well be reciprocal or may operate solely or mainly in the reverse direction. People who feel generally more satisfied with their lives may be more likely to get out and engage with their community. With respect to social connectedness, those who have ties in their extended family and community will be more likely to participate socially and meaningfully.

A second limitation is that the approach used a secondary analysis with public access data. Data were not gathered specifically for this study, so it was limited in what could be included in the participation measurement model. This made the study realistic in reflecting what secondary data analysts might confront but limited the ability of the study to consider the broadest range of possibilities for modeling and measuring participation. Future research projects may benefit from designing a participation measurement instrument from scratch or starting with an existing formal instrument; this might allow for a broader or more appropriate set of activities to be considered, thus better defining the construct of social participation.

This study did not explore formative measurement models for the outcome constructs in the structural model, that is, for life satisfaction or for perceived social connectedness. Life satisfaction, for example, could conceivably be modeled as composed of different aspects of satisfaction—satisfaction with home life, satisfaction with work life, satisfaction with one's social network—and these aspects could be combined formatively into an overall measure of satisfaction. The measures used in this study seem more reflective in nature, appearing mostly interchangeable and tapping globally into an overall level of satisfaction, but even these could be considered formatively. The Diener Satisfaction with Life Scale (SWLS) used in modeling life satisfaction in this study includes items such as "The conditions of my life are excellent" and "So far I have gotten the important things I want in life" (Diener, Emmons, Larsen, & Griffin, 1985). A respondent might feel that they have achieved the important things they wanted in life yet the conditions of their life may be poor due to health or financial limitations. Thus the two items may not be tapping into one underlying latent variable but might better be considered as composing life satisfaction rather than reflecting it.

Similarly, social connectedness could be considered formatively. One way of doing so would be to ask a person about their social connections and quantify the closeness of the relationship. Do they have a live-in spouse or other partner? Are they in contact with their children or other relatives? How often? In this study, social connectedness is measured by using positive items from the UCLA Loneliness Scale, Version 3 (Russell, 1996), which was developed from a reflective perspective. The scale includes items such as “How often do you feel that you are ‘in tune’ with the people around you?” and “How often do you feel isolated from others?” On the criteria given by Jarvis et al. (2003), these indicators appear to be appropriately modeled reflectively. They can be thought of as caused by an underlying perceived social connectedness factor, appear to be interchangeable, would be expected to have moderate to high covariance, and likely have the same antecedents and consequences.

It was reasonable to treat both life satisfaction and social connectedness as reflective measures, both because they were developed using classical test theory and because the items appear to generally meet the criteria for reflective measurement. These measures have shown high internal consistency in previous studies, which is expected from a scale measure but not from an index. Cronbach’s alpha for the SWLS for the 2004 pilot psychosocial questionnaire was .90 (Clarke, Fisher, House, Smith, & Weir, 2008), suggesting a high level of internal consistency; alpha for the loneliness scale was not reported for the HRS data but ranged from .89 to .94 in Russell’s (1996) reliability analysis. Also, the focus of this study was on modeling social participation; life satisfaction and perceived social connectedness are important as elements of the nomological network being considered but are not of primary interest themselves.

Furthermore, incorporating formative constructs as endogenous variables is an inherently problematic exercise (Cadogan & Lee, in press); attempting to incorporate social connectedness and life satisfaction formatively in this study would have made it very difficult to proceed. Formative indicators that define one construct may have different antecedents, making it difficult or impossible to specify parsimonious yet accurate models with endogenous formative constructs. For example, consider socioeconomic status, modeled formatively with income, wealth, educational level, and location of residence. Each of the formative indicators is itself predicted by different sets of variables. Academic talent and interest may predict educational level while family wealth may predict an individual's wealth. Endogenous variables may be best modeled reflectively, allowing a clear predictive path to the endogenous construct of interest. Indeed, this may point to a serious problem for the formative modeling approach as a whole. If social participation, modeled formatively, cannot be treated as an endogenous outcome in structural equation models, for example in rehabilitation intervention trials, does this mean that the use of social participation indexes is useless in such settings? This is yet another limitation of the study; it does not address how to incorporate formative variables endogenously or what general implications this potential problem has for the formative modeling approach or for the construction of measures of social participation.

Chapter Two: Method

This chapter presents the data set, sample, and variables used in constructing and comparing the reflective (scale) model of social participation with the formative (index) model, outlines screening procedures used for checking whether the data met assumptions of the analytic methods used, and describes the data analytic procedures employed in estimating, evaluating, and comparing models. It closes with a discussion of how each research question was intended to be answered by the proposed procedures.

Two measurement models for social participation were developed, evaluated individually, and then compared with each other in terms of fit to empirical data, parsimony, and criterion-related validity. The two hypothesized models that were tested and refined in the scale and index construction process were not derived from existing models of the different facets of participation, because, as discussed in the literature review, there is not an agreed-upon definition of participation, its dimensions, or what activities compose it. Existing participation measurement instruments model participation dimensions in different ways, and each set of dimensions specified depends on the activities chosen to measure participation. For example, the FAI defined domestic, leisure/work, and outdoors domains (Schuling et al., 1993) while the MSPP categorized activities as consumptive participation (e.g., having a meal at a restaurant), formal social participation (e.g., volunteering), and informal social participation (e.g., spending time with family) (Mars et al., 2009). Ad hoc development of participation measures as well has not typically been based upon a common understanding of social participation; unique models have been developed for different research studies. For example, Baker et al. (2005) divided up participatory activities a priori into paid work, formal volunteering, caregiving, informal helping, and do-it-yourself activities while Harlow and Cantor (1996) used cluster analysis to identify domains of participation empirically. The present study defined its own a priori factor model and scale model, based on the participatory activities included in the data set and on informed speculation about what activities would be highly intercorrelated (for the scale model) or would share predictive associations with the outcome constructs (for the index model). Typically, instrument development, whether based on reflective or formative modeling, would first define domains and then develop items to measure those domains. Since this project used secondary data analysis, the approach to identifying domains was more inductive and exploratory.

The first model, the scale model, comprised one or more reflective measures representing participation, developed using CTT-based procedures (exploratory factor analysis, confirmatory factor analysis, and consideration of scale reliability as measured by Cronbach's alpha). The hypothesized factor model is shown in Figure 10. Here, activities out in the community load on one factor, intellectual activities such as reading books load on another, domestic activities such as baking load on a third, and sports and exercise load on a fourth. Someone who does one type of volunteering is more likely to do another, while someone who enjoys one particular home activity (e.g., gardening) is likely to also participate in others (e.g., baking). The reflective approach seemed likely to identify clusters of activity preferences, based on maximizing intercorrelations across items in factors.

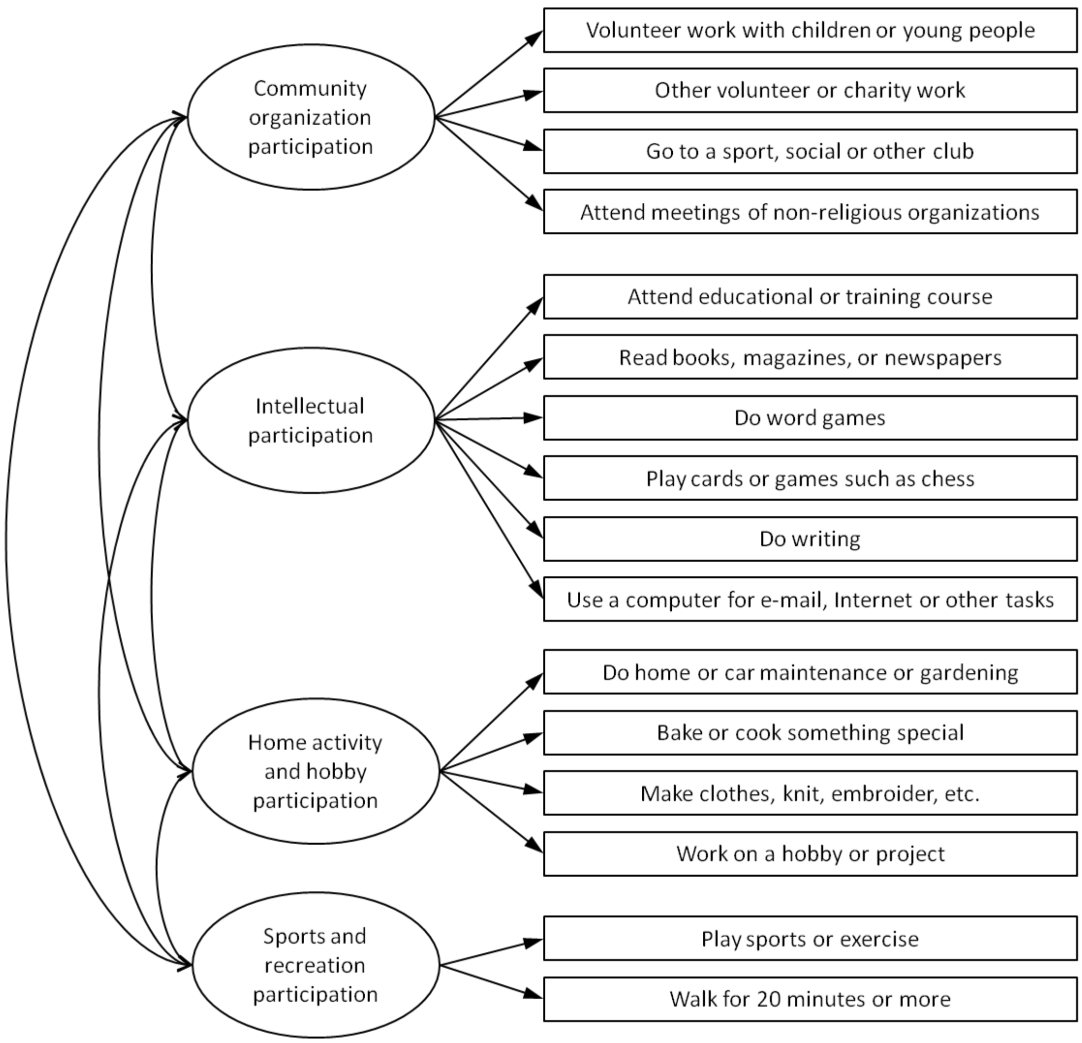


Figure 10. Scale (reflective) model of social participation

The second model, the index model, incorporated formative measures of social participation, constructed and evaluated by estimating composite variables defined by levels of participation in different activity types. The structural model used to identify the formative measurement model included life satisfaction and social connectedness latent variables modeled reflectively as outcomes in addition to participation composites as predictors. An example of such a model is shown in Figure 11; community participation and domestic participation are seen as two kinds of social participation and are modeled formatively as composites. In this approach, multiple measures (in this case, multiple indexes) may be identified just as in the scale approach. However, the two approaches were seen as likely to identify different groupings of items, since the scale approach seeks high intercorrelations within a factor while the formative approach seeks similar patterns of influences on outcomes, by maximizing external variance explained. In the example shown, community participation such as volunteer work, joining clubs, or attending educational courses may have similar effects on the outcomes of life satisfaction and social connectedness while domestic participation such as hobbies, playing cards with friends, or doing home maintenance may themselves share a pattern of influence that is distinct from community participation but similar among themselves. For example, community participation may be more strongly predictive of social connectedness than domestic participation.

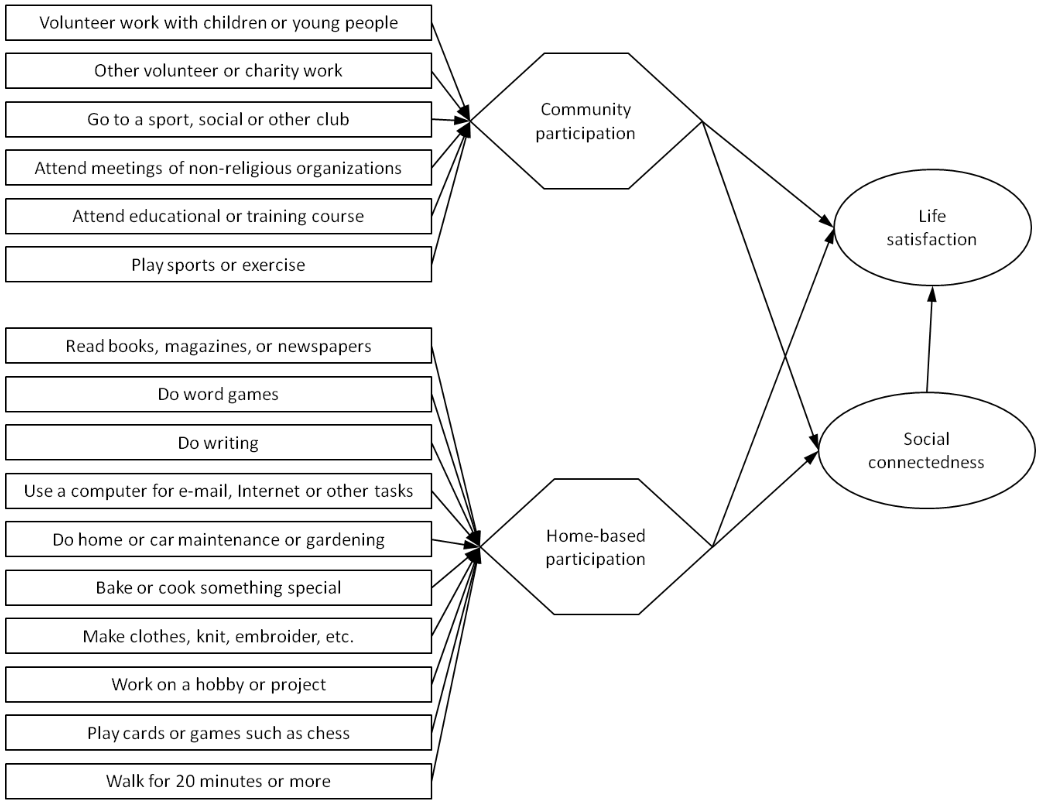


Figure 11. Index (formative) model of participation

After both models of participation were developed and refined, they were compared to each other based on model fit (overall as well as a consideration of areas of poor fit), parsimony (number of constructs required to model participation in the context of the broader structural model including life satisfaction and social connectedness), and predictive validity (ability to explain variance in the outcomes of life satisfaction and social connectedness). This was accomplished by comparing approximate fit indexes, ranking models on the basis of information criteria such as the Bayesian information criterion (BIC; Raftery, 1993), and considering variance explained in the outcomes.

For clarity of the conceptual approach, Figure 10 and Figure 11 do not show measurement errors or endogenous variable disturbances, but of course their specification and estimation is critical to effective structural equation modeling. All reflective indicators will have unique errors specified and estimated; these represent measurement error and other unique unmeasured determinants of the indicators. Treatment of the error terms in the formative model (whether measurement error for causal indicators or disturbance for the formative constructs) is not straightforward; in many cases there is not enough information to estimate the error variance of formative variables. The approach chosen here was to set the error term for each formative construct at zero and consider various fixed measurement error estimates for observed indicators, following the example of Grace and Bollen (2006, 2008). The reasoning behind and implications of these choices are discussed in the literature review in chapter one and later in this chapter.

Sample

The Health and Retirement Study (HRS) is an ongoing panel study completed every two years representing all people over age 50 in the United States, with data going back to 1992 (Leacock, 2006). Respondents are interviewed about a broad array of topics including health and cognitive status, retirement plans, demographic characteristics, and income and wealth. Since 2004, random subsamples of respondents have been asked to complete a psychosocial leave-behind questionnaire (PLQ) addressing factors such as personality traits, life satisfaction and social participation (Clark, Fisher, House, Smith, and Weir, 2008). The sample used for this analysis was limited to non-institutionalized respondents age 65 and over who participated in the 2008 wave of the HRS and completed the PLQ. Of the 4,346 such participants, 1,812 (41.7%) were male and 2,534

(58.3%) were female. The mean age for men was 74.6 years old (*SD* = 6.6) and for women, 74.8 years old (*SD* = 7.2).

In order to ensure that results from analyses of the HRS data generalize to the population, sampling design weights must be used. This analysis used the weights that were constructed for use with the PLQ. These weights are the product of the HRS respondent-level weights and a non-response adjustment factor that predicts the probability of completing the psychosocial leave-behind questionnaire (Health and Retirement Study, 2010).

A subset of the HRS data has been made available in a user-friendly version, known as the RAND HRS, by the RAND Center for the Study of Aging (St. Clair et al., 2010). The RAND HRS includes data from the core interviews and offers constructed values such as a depression scale and imputed values for wealth and income (St. Clair et al., 2010). Data from the RAND HRS were merged with data from the raw 2008 HRS data file, which includes PLQ items (RAND Center for the Study of Aging, 2011).

Measures

Life satisfaction. The HRS 2008 PLQ included Diener's Satisfaction with Life Scale (SWLS), a five-item instrument that has undergone comprehensive reliability and construct validity studies (Diener et al., 1985). The items were "In most ways my life is close to ideal," The conditions of my life are excellent," "I am satisfied with my life," "So far, I have gotten the important things I want in life," and "If I could live my life again, I would change almost nothing" (Clarke et al., 2008). Items were rated on a seven-point scale from 1 (*Strongly disagree*) to 7 (*Strongly agree*). Cronbach's alpha for the sample used in this analysis was .86.

Perceived social connectedness. Perceived social connectedness was measured by items from the PLQ. Question 20 from that questionnaire used 11 items from the UCLA Loneliness Scale, Version 3 (Russell, 1996) such as "How often do you feel that you are 'in tune' with the people around you?" and "How often do you feel isolated from others?" with a three-point response scale (*Often*, *Some of the time*, *Hardly ever or never*). Reversing the scale so that higher is better can be considered a measure of perceived social connectedness. The full scale has shown reliability as measured by Cronbach's alpha ranging from .89 to .94 in past studies (Russell, 1996). Validity was established via correlating the full scale with other measures of loneliness for convergent validity and with measures of health, well-being, and adequacy of personal relationships for construct validity (Russell, 1996). Confirmatory factor analyses have suggested that a model with a global bipolar loneliness factor and two method factors for positive and negative wording fit data collected in earlier studies (Russell, 1996). This study focused on the positive end of the scale (social connectedness rather than loneliness). In order to create a clean unidimensional factor, only positively-worded items were used. Cronbach's alpha for the scale constructed from the seven positive items for the sample used in this analysis was .87.

Social participation. The HRS 2008 administration collected information relating to participation through the PLQ, which asked respondents to indicate how much time they spent in particular activities, using a scale of 1 (*Daily*), 2 (*Several times a* *week*), 3 (*Once a week*), 4 (*Several times a month*), 5 (*At least once a month*), 6 (*Not in the last month*). PLQ participation items were recoded on a scale from 0 (*Not in the last month*) to 5 (*Daily*). The activities the questionnaire asked about were:

- Care for a sick or disabled adult
- Do volunteer work with children or young people
- Do any other volunteer or charity work
- Attend an educational or training course
- Go to a sport, social, or other club
- Attend meetings of non-religious organizations, such as political, community, or other interest groups
- Pray privately in places other than a church or synagogue
- Read books, magazines, or newspapers
- Do word games such as crossword puzzles or Scrabble
- Play cards or games such as chess
- Do writing (such as letters, stories, or journal entries)
- Use a computer for e-mail, Internet or other tasks
- Do home or car maintenance or gardening
- Bake or cook something special
- Make clothes, knit, embroider, etc.
- Work on a hobby or project
- Play sports or exercise
- Walk for 20 minutes or more

A broad set of activities such as this had the potential to demonstrate the differences between scale and index approaches more dramatically than a more constrained set, so most of the activities in this list were used in developing the model of social participation. Two, however, were not used. “Caring for a sick or disabled adult” was not included in the items since it is most likely to represent caring for one’s spouse or other live-in relative and is likely more obligatory than voluntary. The focus of this study is on participation that is specifically chosen by the respondent because it is enjoyable and meaningful to them, and in which they have substantial autonomy. “Praying privately” was not included in the analysis, since it is intrinsically not social and is unique among the activities in incorporating religiosity. While “reading books, magazines, or newspapers” is likewise not inherently social, it was included in the analysis for three reasons: (1) it is likely to be correlated with other intellectual activities such as doing word games or using a computer to browse the Internet, so is expected to load on at least one factor in the scale model and (2) it may lead to increased social connectedness and life satisfaction because it provides a means of connection, discussion, and enjoyment so may play an important role in a social participation composite. In other words, there is a case to be made for both its likelihood to intercorrelate with other participation items (thus representing an important element of one of the scale participation measures) and for its ability to predict the outcomes of life satisfaction and social connectedness (thus representing an important element of one of the index participation measures). Keeping it in the list of activities considered may help illuminate the differences between the two approaches to modeling participation.

Control variables. Life satisfaction and social connectedness have been shown to be related to a number of variables not of specific interest in this study, except to the extent that they may confound the relationship between social participation, social connectedness, and life satisfaction. Control variables used were gender, income and wealth, age, single vs. partnered (defined as married or living with someone), and physical and mental health. Income and wealth were measured at the household level; wealth included the value of any second home. To control for health status, the respondent's self-report of health on a scale from 1 (*Excellent*) to 5 (*Poor*) was used. The mental health measure used a score from the Center for Epidemiologic Studies Depression (CESD) scale (Radloff, 1977), which incorporates negative indicators such as restless sleep or feeling alone with positive indicators feeling happy or enjoying life, all or most of the time (St. Clair et al., 2010). In early reliability studies, the scale showed alpha reliability of .84 or greater (Radloff, 1977). It has been shown to be reliable and valid in studies using samples of community-dwelling older adults (Berkman et al., 1986).

Software

SAS Version 9.2 and PASW Version 18 were used for data merge, variable construction, and data screening. Mplus Version 6.11 (Muthén & Muthén, 2010) was used for fitting structural equation models. Mplus also offers some data screening capabilities, such as the ability to identify multivariate outliers and computation of sample statistics. These facilities were used as convenient.

Data Screening

Prior to analysis, the analysis data set was screened to ensure it met assumptions of structural equation modeling. In case important violations of assumptions were found, data was transformed or analytic procedures adjusted to respond to the violation.

Descriptive statistics. For each observed variable to be used in the analysis, descriptive statistics were inspected and reported. For continuous and ordinal variables statistics reported included mean, standard deviation, minimum, maximum, skewness, kurtosis, and percentage missing. For nominal variables, frequencies were inspected.

Missing data. Missing data patterns can be characterized as MCAR (missing completely at random), MAR (missing at random), or MNAR (missing not at random) (Tabachnick & Fidell, 2007). If data are missing without any systematic pattern, then the missingness is MCAR). If values are missing systematically, but the missingness can be predicted (and hence corrected for) using observed values of other variables, then the missingness is MAR. However, if values are missing systematically and that systematic missingness cannot be predicted based on available observed values, the missingness is considered MNAR. Because missing values are not observed, it is not possible to definitively confirm whether the pattern of missing values should be classified as MCAR, MAR, or MNAR; however, a researcher can investigate whether missingness on a particular variable appears to be predictable by observed values of other variables; if so, the assumption of MAR may be appropriate. The maximum likelihood estimation algorithm used by Mplus 6.11, which will be used for the analyses in this project, handles missing data on outcome variables properly so long as the missing data is MAR or

MCAR, but cases with missing covariates are deleted from the analysis unless special steps are taken to specify distributional assumptions for them (“Missing Data Modeling,” n.d.). In the formative version of the social participation model, the participation indicators function like covariates, since they are specified as exogenous predictors in the model. Missing data on the participation indicators is therefore not handled by Mplus’ maximum likelihood algorithms in the formative version of the model, unless the indicators are modeled as latent variables with fixed measurement error. In that case, the items are no longer exogenous and missingness on them can be handled without case deletion.

The variables used in this data study were, for the most part, expected to show a MAR missingness pattern. For example, someone with poorer physical or mental health may be less likely to answer all questions on the PLQ, but this would be corrected for by the inclusion of health covariates in the analysis, at least when the variables with missing data are specified as endogenous. The core constructs of social participation, social connectedness, and life satisfaction are not particularly sensitive topics, so refusals to answer questions on such topics based on values of the variables likely wouldn’t happen with great frequency. Tabachnick and Fidell (2007) suggested that researchers investigate missingness patterns for any variables with more than 5% missing data. If particular variables showed more than 5% missing data, then an investigation of the patterns of missingness as related to other variables in the analysis data set was undertaken.

Wealth and income, two of the covariates to be used in the analysis, are sensitive topics, and there may be many respondents who were unwilling to provide information about their financial status. However, the cleaned and enhanced HRS data set provided by RAND to be used in these analyses included imputed wealth and income data (St. Clair et al., 2010). Although the use of such singly-imputed values can introduce downwardly-biased standard errors, this was not expected to seriously impact results since wealth and income are not core to the analysis but are used as control variables only.

Multivariate normality. The default maximum likelihood estimator used by Mplus assumes multivariate normality, in which all variables are distributed univariate normal, joint distributions of all pairs of variables are bivariate normal, and bivariate scatterplots of each pair of variables are linear and homoscedastic (Kline, 2005). Since there were so many variables to be used in this study, checking all bivariate scatterplots was not feasible. While there are tests of multivariate normality, these are not available in basic statistical packages. Therefore only univariate normality was checked. Statistical tests of the univariate normality assumption tend to reject the normality hypothesis with large sample sizes so Tabachnick and Fidell (2007) suggested visually inspecting the shape of distributions (such as with frequency histograms) and checking the absolute value of skewness and kurtosis. Kline (2005) suggested that absolute values of skewness greater than three and absolute values of kurtosis greater than 10 are problematic, based on the results of SEM simulation studies. These cutoffs were used to indicate substantial departures from univariate normality that required investigation and possibly, transformation or selection of an estimation algorithm robust to non-normality.

Outliers. An outlier is a case with an extreme value on one variable (a univariate outlier) or with an unusual pattern of values across multiple variables (a multivariate outlier). One reason outliers are of concern in SEM is because they may compromise multivariate normality (Kline, 2005). Mplus has the capability to identify outliers using four methods: Mahalanobis distance, log-likelihood contribution, a measure of log-likelihood distance influence, and Cook's D, which estimates influence ((Muthén & Muthén, 2010). Because the log-likelihood distance influence and Cook's D measures are computationally intensive, requiring the recalculation of each model as many times as there are observations (Muthén & Muthén, 2010), the log-likelihood and Mahalonobis distance statistics were used to identify multivariate outliers. The estimation algorithm chosen was Mplus' MLR option, which produces chi-square statistics and standard errors that are robust to non-normality, such as might be caused by univariate or multivariate outliers. In addition, preliminary scale, index, and full structural model analyses were run both with and without multivariate outliers identified by large log-likelihood values removed. The multivariate outliers did not appear to substantially change results, so all results are reported from analyses with outliers included.

Linearity. Unless the researcher transforms variables, SEM estimates linear relationships so the presence of nonlinearity may compromise the parameter estimates. Linearity was checked not across all pairs of individual indicators, but instead by constructing subscale scores for the constructs of life satisfaction and perceived social connectedness then plotting bivariate scatterplots of those scores against covariates and individual participation item indicators.

Multicollinearity. Variables which are perfect or near-perfect linear combinations of other variables can cause problems with model estimation, because they may cause variance-covariance matrices to be non-invertible (Kline, 2005). The correlation matrix was inspected for extremely high correlations (those greater than .90). This can identify pairwise multicollinearity, but not multicollinearity existing across three or more variables (Kline, 2005). During model estimation, there were no errors referencing non-invertible or singular matrices which might have suggested the presence of multicollinearity.

SEM Procedures

This section discusses procedures common to all models: ensuring model identification, modeling measurement error, evaluating model fit, estimating models, comparing nested and non-nested models, and refining model specifications.

Ensuring model identification. A structural equation model must be *identified* in order for it to be evaluated; this means it must be "theoretically possible to derive a unique estimate of each parameter" (Kline, 2005, p. 105). In order to be identified, all structural equation models must meet the following two necessary conditions:

- The *t* rule. "The number of free parameters in a model, say t, must be less than or equal to the number of nonredundant elements in the covariance matrix of the observed variables in the model." (Bollen & Davis, 1994/2009, p. 502)
- The scaling rule. "Each latent variable in a structural equation model must be assigned a scale for the model to be identified." (Bollen & Davis, 1994/2009, p. 502)

These conditions are necessary but not sufficient to guarantee identification (Bollen & Davis, 1994/2009). Another way of stating the *t* rule is to state that there must be at least as many knowns as unknowns; in this case, the knowns are the variances and covariances of the observed variables and the unknowns are the free model parameters to be estimated (Kline, 2005). The number of knowns is computed as $v\ (v{+}1)\ /2$ where v is the number of observed variables in the model; this represents the number of unique variances and covariances in the covariance matrix (Kline, 2005). Models that are underidentified (i.e., that violate the *t* rule) have more unknowns than knowns and this cannot be rectified by adding cases to the data set, since the knowns are observed variances and covariances (Kline, 2005). It can, however, be addressed by fixing free parameters to particular values or by reducing the number of paths one is trying to estimate. The scale of a latent construct modeled reflectively is usually established by setting one of its factor loadings to one (Bollen & Davis, 1994/2009). This is the approach that was taken with constructs modeled reflectively. Measurement errors and disturbances (unmodeled causes of endogenous variables) also need their scales established; this was accomplished via unit loading identification (ULI) constraints that specify path coefficients of one from the error or disturbance to the variable to which they point (Kline, 2005).

CFA models with one factor are identified if they have three indicators and are modeled unidimensionally, meaning that there are no correlated error terms specified in the model (Kline, 2005). For multi-factor models, unidimensional measurement requires no correlated errors and no indicators loading on more than one factor. Unidimensional multi-factor models are identified if they have at least two indicators per factor, but more indicators per factor are desirable in order to avoid estimation problems (Kline, 2005).

Structural regression (SR) models (latent variable models with directional paths between the latent variables) must meet the *t* rule and the scaling rule but as mentioned, this does not guarantee identification. Bollen (1989) offered a two step rule for evaluating the identification of an SR model. First, respecify the SR model as a CFA model (replacing directional paths with unanalyzed associations among factors) and evaluate this against CFA identification requirements as described above. Second, consider whether the directional portion of the model is *recursive*; if it is, and if CFA identification requirements are met, then the overall model is identified. The directional portion of the model, which can be considered abstractly as a path model that specifies directional associations between non-latent (i.e., observed) variables, is called recursive if it has no uncorrelated disturbances and all causal effects are unidirectional (Kline, 2005).

Models with formative constructs must meet Bollen and Davis' (1994/2009) *2+ emitted paths rule* which states that "Every latent variable with an unrestricted variance (or error variance) must emit at least two directed paths to variables when these latter variables have unrestricted error variances" (p. 503). This is a necessary but not sufficient condition for identification. If the formative construct emits paths to two correlated outcome variables (as is the case in this study), then the model is unidentified (Bollen & Davis, 1994/2009). However, modeling formative constructs as composites, that is, with zero variance, allows the model to be estimated, assuming a scale is set for each composite (Grace & Bollen, 2006). Scales can be set for composites by specifying unit loadings from one of the causal indicators to the composite (Grace & Bollen, 2006).

Modeling measurement error. SEM explicitly models measurement error, that variance in each observed indicator that is not explained by underlying latent factors (Kline, 2005). As shown in Figure 12, an SEM model with reflectively-measured constructs typically models measurement error via individual and unique error terms, labeled e_1, e_2, and e_3. The error terms measure more than just score unreliability; they incorporate all unmeasured influences on the indicator of interest (Kline, 2005). In a unidimensional measurement model, these errors are modeled as uncorrelated; introducing a correlation between two error terms suggests that an additional factor of some sort exists. In developing reflective-indicator models of constructs, this study started with uncorrelated errors and introduced correlated errors only to the extent that they were substantively justified (for example by similar wording or concepts in items).

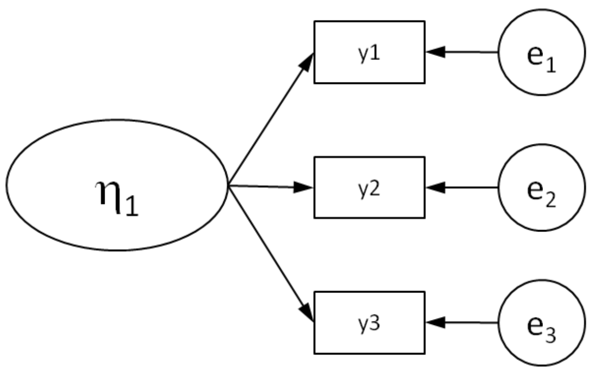


Figure 12. Modeling unique variance in indicator variables

Formatively-modeled constructs do not generally incorporate measurement errors for the causal indicators into the model. The error that may be modeled at the construct level (z_1 in Figure 13) reflects omitted causes of the formative construct (Bollen & Lennox, 1991). Causal indicators, however, just like reflective indicators, are unlikely to be perfectly reliable. This difficulty is addressed in the index construction section later in this chapter.

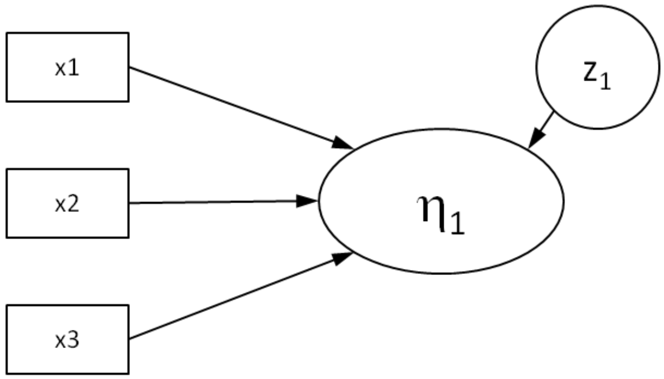


Figure 13. Formatively-measured construct without measurement error

Incorporating covariates. Subjective well-being, of which life satisfaction is one component, has been shown to be related to a variety of other variables including socioeconomic status, physical and mental health, and demographic characteristics (Baker et al., 2005). Similarly, social connectedness is expected to depend not solely on activities but also on other covariates. In order to condition the results on differences across respondents, covariates as described in the measures section were introduced as predictors of the latent life satisfaction and social connectedness variables. They were incorporated as single-indicator variables uncorrected for measurement error.

Multiple group analysis by gender. Gender has been shown to be an important moderator of the relationship between participation and well-being (Harlow & Cantor, 1996). Multiple-group analysis was used to check whether patterns of associations among social participation, social connectedness, and life satisfaction differed by gender. In multiple-group analysis, structural equation models are fit to both groups with and without constraints imposing equality on certain model parameters, to see if there is a significant deterioration of fit when the cross-group equality constraints are imposed (Kline, 2005). When multiple-group analysis is applied to a measurement model, the question of interest is whether measurement model can be treated as invariant across groups (T. Brown, 2006). When applied to the structural model, the question is whether the structural relationships can be treated as invariant across groups. Multiple-group analysis was deployed during measure construction.

Measurement models can be characterized as invariant at four levels: configural, metric, measurement error invariance, and scalar invariance (Cheung & Rensvold, 2002). The first two were considered important for the purposes of this project. First each measurement model was checked for configural invariance. In the scale model, configural invariance would occur if each gender shows the same number of factors with the same items loading on each factor. In the index model, configural invariance requires that a model comprising the same indexes composed of the same items fits adequately for both genders. Those measurement models that show configural invariance were checked for metric invariance, or invariance of factor loadings (scale model) or composite weights (index models). Since the measurement models did not show metric invariance across genders, the structural models were fit individually by gender.

Evaluating model fit. With a sample size of thousands of cases, the chi-square goodness of fit test is likely to reject most models, even if the model is “minimally false” (Bentler & Bonett, 1980, p. 591). Therefore, chi-square statistics were reported but significant values were not be used to reject models. A model was considered to show good fit if CFI was greater than or equal to .95 and RMSEA was less than .05 (Hu & Bentler, 1999) and acceptable fit if CFI was greater than or equal to .90 and RMSEA was less than .10 (Kline, 2005). SRMR was also reported, with values less than .08

considered good (Hu & Bentler, 1999) and less than .10 considered adequate (Kline, 2005).

Comparing models. Two structural equation models are considered nested if one is a subset of the other; that is, if the two models have the same configuration except that one model has more constraints (such as paths or error variances fixed to zero) than the other (Kline, 2005). Nested models were compared by means of chi-square difference tests, where the difference between model chi-squares was computed and tested with degrees of freedom calculated as the difference between the two models' degrees of freedom (Kline, 2005).

When models are not nested, they can be compared by means of predictive fit indexes such as the Akaike information criterion (AIC) and the Bayes information criterion (BIC) (Kline, 2005). Models with lower AIC and BIC values are preferred (Kline, 2005). These require use of the same data sets, with the same variables and same observations; they are meaningless if used to compare models that include different sets of variables. In order to use the AIC and BIC to compare models, it was necessary to ensure that the models to be compared included the same activity items. Since the scale and index approaches use different criteria for including items in measures, use of the same activity items was not automatic. It was achieved by including any activity items individually if they appeared in one set of measures but not the other.

Both the AIC and BIC trade off fit (how well the model reproduces the observed data patterns) and parsimony (how many parameters are used in the model), with the BIC penalizing complexity more (Forster, 2000). A model with more parameters is likely to fit the data better but at the same time is likely to capitalize on chance variation in the data set, and so may be further from the true population model (Zucchini, 2000). Philosophically, the problem that the AIC and BIC attack is the same one addressed by cross-validation, where one subset is used to fit a model and another subset is used to check the model's predictive accuracy: that adding more parameters increases the chance of "overfitting" the model to random noise in the data set. The AIC is asymptotically equivalent to a kind of cross-validation known as leave-one-out cross-validation (Forster, 2000). Unfortunately, both cross-validation and the use of information criteria to compare models suffer from the same problem as use of the model chi-square for model fitting. That is, as sample size increases, they "provide little or no additional information over a direct comparison of models using only the calibration stage." (Busemeyer & Wang, 2000, p. 178). Nevertheless, the BIC, with its greater attention to parsimony, was used to compare non-nested models in this study, given the lack of a better alternative.

Refining model specifications. Model respecification was guided by theoretical considerations first, based on alternative specifications that were considered competing models to the one being tested. Modification indexes were inspected to guide development of models in order to achieve adequate fit when hypothesized models did not provide it.

Estimating models. Because there was evidence of non-normality, the Mplus MLR algorithm which provides robust standard errors (Muthén & Muthén, 2010) was used.

Cross-validation. Since both the scale construction and index construction procedures involved respecification and refinement guided by modification indexes as well as by substantive considerations, the models developed needed to be checked in holdout samples (T. Brown, 2006). In order to validate the developed models and explore their predictive power, the data set was divided in half, with equal numbers of cases randomly assigned to each half. The first half was used for measure construction. The second half was used for checking the fit of the measurement models in a separate sample and for evaluating and comparing the performance of the scale and index models in the overall structural model.

Scale Construction

The scale construction process with activity items entered as reflective indicators followed conventional CTT-based procedures: an exploratory factor analysis to identify the empirical factor structure of the data set, followed by a confirmatory factor analysis to check the fit of the measurement model and refine it. Because EFA did not unequivocally identify how many factors should be used or what their structure should be, a number of CFA models were fit. These used the measurement construction half of the data set. After the factor models were developed, they were fitted in the validation subsample to check fit.

Hypothesized factor structure. Factor analysis should be informed by substantive considerations and the results of prior research as well as by statistical guidelines (T. Brown, 2006). Inspection of the items, considering which activities are likely to be correlated in the population, suggested four areas of activity that may show moderate to high intercorrelation: community organization participation, intellectual participation, home activity and hobby participation, and sports and recreation participation. This hypothesized factor structure is shown in Figure 10 and reflects the assumption that correlations will be based on personal preferences and temperament. For example, someone who enjoys reading may also enjoy crossword puzzles, since both involve language and intellectual engagement while someone who engages in volunteer work may also choose to participate as a member of other community organizations such as social clubs. The number of factors hypothesized was consistent with the number of participation dimensions that participation instrument developers have previously theorized. For example the MSPP (Mars et al., 2009) defined four sub-indexes, the FAI (Schuling et al, 1993) proposed three, and the CIQ (Reistetter et al, 2005) proposed three. Note, however, that Harlow and Cantor's (1996) empirical study of clusters among 33 participation items found eight domains with five items not fitting into any domain. Also, reliabilities for two of the FAI's three subscales were low (Schuling et al., 1993), calling into question whether it makes sense to posit a common latent factor underlying each one.

Identifying the empirical factor structure. EFA seeks to identify the smallest number of underlying common factors that explain the correlations among items (T. Brown, 2006). EFA also assigns items to factors, providing guidance as to which factor an item should load upon. There are a number of estimation methods that may be used to estimate a factor solution including principal factors, maximum likelihood, and generalized (weighted) least squares (Tabachnick & Fidell, 2007). Maximum likelihood has the benefit of allowing for the statistical evaluation of fit of the factor model, but can lead to improper solutions and assumes multivariate normality for its statistical tests to be valid while principal factors (PF) is not as prone to improper solutions and does not have any distributional assumptions (T. Brown, 2006). PF was used to estimate the EFAs because of its practicality.

Of concern in this study is the dimensionality of the participation items; a variety of criteria for deciding how many factors exist among a set of items have been proposed. Many of these procedures are based on use of *eigenvalues,* which summarize variance in the variance/covariance matrix of a data set analyzed with EFA (T. Brown, 2006). These procedures include the Kaiser-Guttman rule (sometimes known as the Kaiser criterion) that counts factors as practically significant if they have eigenvalues greater than one or the scree test, in which the researcher inspects a plot showing eigenvalues plotted in decreasing order of magnitude and selects those factors whose eigenvalues appear before the slope of the plot flattens noticeably (T. Brown, 2006). In this study, I used parallel analysis (Horn, 1965) along with substantive considerations. Parallel analysis compares the eigenvalues from the factors in the data set to eigenvalues from a randomly-generated data set that has the same number of cases and variables as the real data set and retains only those factors from the original data set with eigenvalues greater than the averaged eigenvalues from the generated data set (Tabachnick & Fidell, 2007). I also inspected scree plots and considered the results of applying the Kaiser-Guttman rule. The number of factors wasn’t obvious given the results of these tests, so a number of alternate solutions specifying a fixed number of factors were estimated. Oblique rotations were used to produce more interpretable factor solutions then solutions were inspected to see which seemed most interpretable and which aligned with the hypothesized four-factor model. Solutions were also estimated by gender since the factor structures appeared to differ for men and women.

Refining the factor structure. CFA analysis was conducted in Mplus (Muthén & Muthén, 2010), comparing the hypothesized model presented above to models with additional factors as well as to models informed by the results of the EFA. In CFA, all aspects of the model including the number of factors, which items load on which factors, the pattern of correlation across indicator errors, and so forth must be pre-specified; they are not generated based on the data (T. Brown, 2006). However, this process takes on an exploratory perspective when many different models are compared, as was done in this project. Models were developed both for the data set as a whole and for males and females separately. The four-factor structure shown in Figure 10 was estimated and compared to five, six, and seven-factor models by means of chi-square difference tests. Additional factors were created by splitting up community organization participation, intellectual participation, or both, in the following manner:

- Community organization participation was divided into two factors, one with the two volunteering items and one with "going to clubs" and "attending meetings of non-religious organizations."
- Intellectual participation was divided into two factors: games ("do word games" and "play cards or games such as chess") and writing and reading ("read books, magazines, or newspapers", "use a computer for e-mail, Internet, or other tasks"). The item "Attend educational or training course" was then entered as a single-indicator factor, since it was the only one of the six items that generally takes place outside the household or immediate social group.

The five-factor model was achieved by splitting community organization participation into two factors. The six-factor model was achieved by retaining the community organization participation factor as one factor but splitting intellectual participation into three factors. The seven-factor model split both the community organization participation and intellectual participation factors into multiple factors. The item “Attend educational or training course” was thought to possibly load on the first factor, hypothesized as “community organization participation,” if the factor actually represents out-of-home activities. Models with this item loading with volunteering and club participation items were considered. While these hypothesized factor structures resulted in models with many factors, some of which have only two or even just one indicator, past empirical research had suggested that such solutions might be required to achieve adequate fit, defined in this study as RMSEA less than or equal to .10 and CFI greater than or equal to .90. While researchers typically want to have at least three indicators per item, such a rule of thumb was not used here, since the priority was to find a model that showed good empirical fit and also to explore to what extent the scale model suited this kind of data.

Results of the EFA were used to further adjust the model as needed to achieve adequate fit. Ultimately, this scale development process was more exploratory than confirmatory, given the starting point of a pre-existing list of activities not designed around a dimensional model of participation. Typically in scale construction the researcher would start with a theoretical model, write items to capture each factor, and then test the model as hypothesized. In this case, the goal was to generate a good-fitting and intuitively reasonable representation of the underlying factors of the activities, so as to compare this against a formatively-modeled set of participation indexes.

Once an adequately-fitting model was achieved, it was estimated using CFA in the validation half of the data set. Reliability coefficients for each scale were computed using the validation half and reported.

Fitting the structural model. Once the factor structure of the reflective model of participation was developed, the factors identified were incorporated into the model that included the outcomes of social connectedness and life satisfaction, using the validation half of the data set. An example of the structural model that would have been fit if the hypothesized model were confirmed is shown in Figure 14; the actual model that were fit used the participation factors identified by the scale construction processes described above.

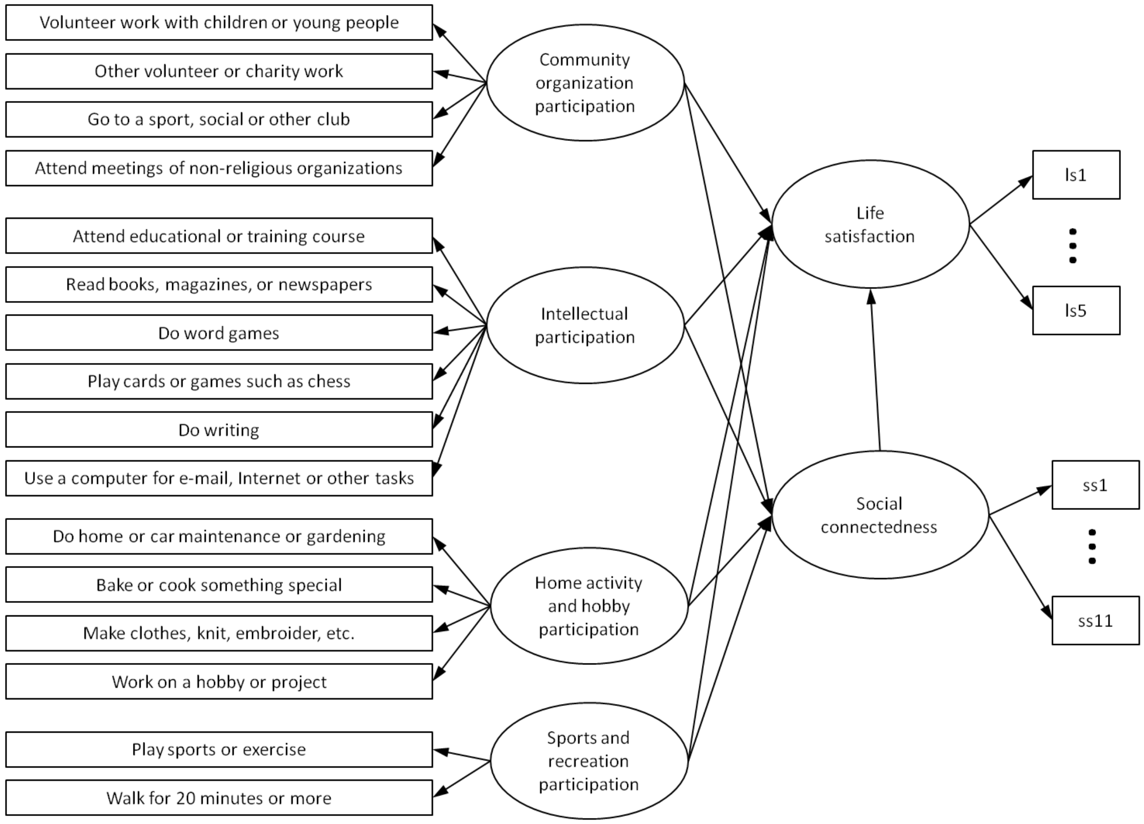


Figure 14. Structural model with hypothesized reflective participation factors

Index Construction

The index construction approach modeled composite variables that showed significant relationships with outcome constructs of interest, in this case perceived social connectedness and life satisfaction. Items combined into one composite should show roughly the same pattern of associations with the outcome constructs of interest. A hypothesized composite model is shown in Figure 11. Like the hypothesized scale model, this was not generated based on existing theory about participation, since there is not agreement on what constitutes participation, what activities to include, or how different domains of participation should be characterized. Instead, it was developed based on consideration of how the different activities included in the data set were likely to share in different patterns of association with the outcome variables of social connectedness and life satisfaction. Activities that engage a person out in their community were expected to show moderate to strong relationships with both life satisfaction and support, while home-based participation such as hobbies or reading may not show as strong a relationship with social connectedness, working mainly directly to improve life satisfaction. Still, even home-based participation was expected to increase a person's sense of social connectedness, since activities like gardening, Internet usage, and reading books can serve as a point of connection to other people.

As with the scale development process, the index construction process generated measures of participation using the measure construction half of the data set. After the index model was constructed, it was fit in the validation half of the data set so its performance could be compared with the performance of the scale model.

Ensuring model identification. As Bollen and Davis' (1994/2009) 2+ emitted paths rule states, a formative construct is unidentified unless it is embedded into a model in which it includes two outgoing paths. Fitting the composite model required fitting the entire structural model including life satisfaction and social connectedness constructs; items from the validated SWLS and UCLA loneliness scales were used to model the outcome constructs of life satisfaction and social connectedness, respectively. Even after embedding the formative constructs into the overall structural model, the model would have remained underidentified unless the constructs were (1) assigned a scale and (2) treated as composites with zero error. The constructs were assigned a scale by setting the loading from one causal indicator per composite to the composite at one. Error at the construct level was not modeled (in other words, it was fixed at zero); thus the variables as modeled are termed *composites* rather than *formative constructs*. The distinction between the two and its implications were discussed in the literature review in chapter one.

Modeling measurement error. The composite model, unlike the reflective model, does not automatically model measurement error for indicators (Bollen & Lennox, 1991). In a reflective model, unique error terms for each indicator are included and can be estimated given the model meets certain identification requirements. In a formative model, error may be modeled at the formative construct level (e.g., for the participation construct); this represents all omitted causes of the latent construct, not measurement error (Bollen & Davis, 1994/2009). In a composite model, such as estimated here, error is not modeled at either the indicator or the construct level. The composite is treated as a perfect and nonrandom linear combination of observed indicator values.

Formative measurement has been criticized because it doesn't generally account for measurement error in the causal indicators (Edwards, 2010). One way to handle measurement error would be to introduce multiple items per activity so as to be able to estimate latent variables for each different activity type. For practical reasons, it is unlikely in index development that multiple indicators with latent constructs would be incorporated in such a way. The data set used in this study does not have multiple indicators per activity type and formal participation instruments such as the FAI (Schuling et al., 1993), the MSPP (Mars et al, 2009), and the POPS (Brown et al., 2004) do not typically ask multiple questions about participation in specific activities. Grace and Bollen (2006) suggested specifying a fixed amount of measurement error for single-indicator measures such as those used in composite variable definitions, and used a value of 10% measurement error for their ecological models that incorporated composites defined by multiple single-indicator measures. During index construction, composite models were estimated with no error for the participation indicators, with 10% error, with 20% error, and with 30% error. The results were compared, to see how sensitive they were to the amount of error specified. A fixed percentage of measurement error was estimated in the model by specifying a residual term for each activity, fixing its path to the activity indicator at one, and then constraining the variance of the residual term to be a specific percentage of the sample variance of the activity indicator (Kline, 2005).

Fitting the model. Initially, a disaggregated model was fit, as shown in Figure 15. This model estimates unique influences on social connectedness and life satisfaction for each activity indicator in the analysis. It was inspected to see if any indicators did not have significant effects on either one of the outcomes. Indicators without significant effects could not be dropped entirely from the model if they were included in one of the scales in the reflectively-based model (that would invalidate model comparison based on the BIC), but a lack of significant predictive relationship suggests that they may not meaningfully participate in any composites representing an index. For comparison purposes, they were included in the model with paths to the outcome constructs constrained to zero.

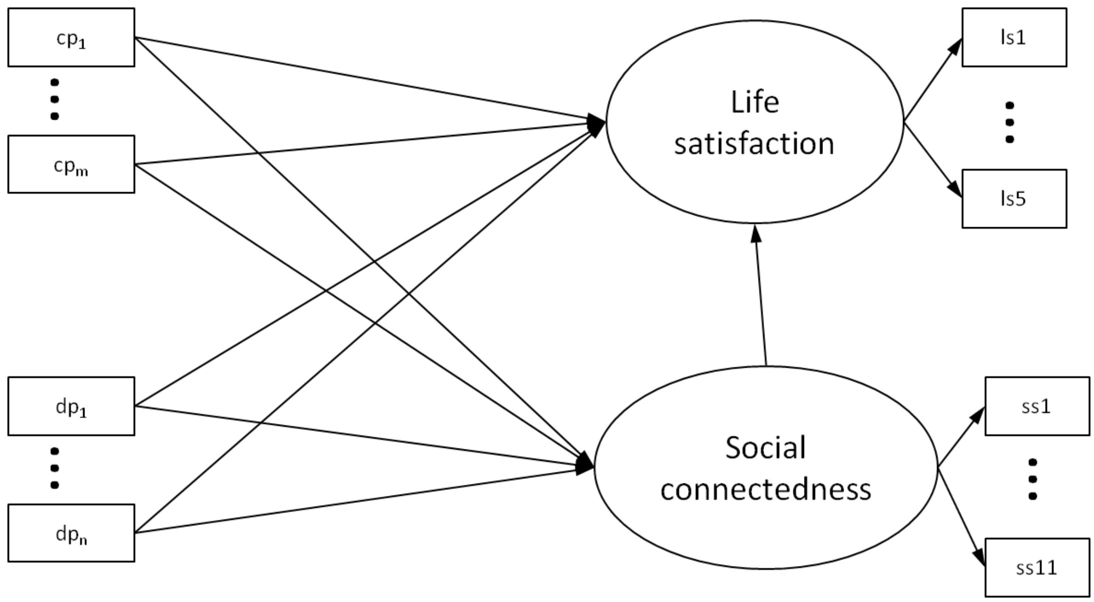


Figure 15. Disaggregated participation model

After the disaggregated model was examined, the model with composites was estimated across genders and for males and females separately (Figure 11). Additionally, a model with just one participation composite involving all activity items was estimated and its fit compared with the hypothesized model.

Refining the model. Once an adequately fitting composite model was developed for males and females, it was refined by trimming non-significant indicators. In this step, a generous alpha level of .10 was used. This was used instead of .05 because correlations across the items would tend to lead to higher *p*-values when all items were in the model, compared to what you would see with a reduced set of items.

Model Comparison

Once adequately-fitting models were constructed using the scale and index construction procedures outlined here, they were incorporated into the full structural model that included the outcomes of social connectedness and life satisfaction as well as covariates of those outcomes such as age, years education, and self-report health status. Since measurement invariance was not established for the scale and index models, the full structural models were fit and compared by gender. Models were compared by means of BIC values since they were not nested. Approximate fit indexes were reported and compared. Variance explained in outcomes, represented by R^2 values for the latent outcome variables, was also compared across models. Results of the model comparison suggested additional models to be estimated and interpreted, so three ad-hoc analyses were completed in order to provide additional information for use in answering the research questions.

Chapter Three: Results

This chapter presents the results of the analyses described in chapter two. First data were screened to ensure they met the assumptions of structural equation modeling. Next the data set was split randomly into a measure construction half and a validation half. Using the measure construction half, scale and index models of social participation were developed. Each was validated using the confirmatory half of the data set. Then the scale and index models were compared by fitting them to the overall structural model, again using the confirmatory half of the data set. There is a brief discussion of how the results answer the research questions and finally, the results of the comparison are used to develop and interpret gender-specific models of the relationship between life satisfaction, social participation, and perceived social connectedness.

Data Screening

Table A.1 in Appendix A shows percentages of participants reporting different levels of participation in the 16 activities to be included in the analysis, ordered according to percentages of respondents reporting daily participation, from largest to smallest. Sixty-eight percent of respondents reported reading daily, while the next most popular daily activities were computer use (24%), word games (21%), or walking at least 20 minutes (20%). Twenty percent of respondents reported participating in sports or exercise several times a week while 16% reported engaging in hobbies and 10% reported baking or cooking with similar frequency. Education and volunteer work were the least popular activities: considering their activities during the past month, 81% of respondents reported they had not attended an educational or training course, 80% reported they had not engaged in volunteer work with youth, and 64% reported they had not engaged in other volunteer work.

Chi-square tests of association were completed to see if there were significant differences in participation levels in different activities by gender. A Bonferroni correction was used to adjust significance levels for the 16 tests completed. Using a significance level of .0031 (= .05/16), only four activities did not show different patterns of activity levels by gender: volunteering (other than with youth), education, non-religious organization participation, and computer usage. These patterns were explored further by inspecting tables of percentages of respondents by gender reporting different levels of participation (Table A.2). Activities showing percentage differences at different levels of participation of at least five percent or greater are noted in the table. More women than men engaged in volunteering with youth: eighty-three percent of men reported no participation in the last month in volunteering with youth compared to 77% of women. Twelve percent more women than men reported daily participation in word games. Five percent more women than men reported writing several times a month. Women generally reported higher frequencies of baking or cooking while men reported engaging in home maintenance, car maintenance, or gardening more frequently than women. Few men reported any participation at all in sewing or knitting: ninety percent reported they had not engaged in it in the last month. Almost 23% of men reported walking daily compared to about 18% of women.

Table A.4 and Table A.5 in Appendix A report correlations for life satisfaction and perceived social connectedness items, respectively. Correlations for life satisfaction items ranged from .45 to .74. Correlations for social connectedness items ranged from .33 to .72, with the lowest correlations (between .33 and .38) holding between the first item, "How often do you feel that you are 'in tune' with the people around you?" and the remaining items. Table A.6 and Table A.7 in Appendix A report correlations by gender for the social participation items. Correlations were generally low, not even reaching .30 in most cases. For males, the highest bivariate correlations computed were for sports/exercise with walk for 20 minutes ($r = .46$, $p < .001$) and for hobbies/projects with home/car maintenance or gardening ($r = .40$, $p < .001$). For females, the highest bivariate correlations were for sports/exercise with walk for 20 minutes ($r = .45$, $p < .001$) and for hobbies/projects with baking/cooking ($r = .49$, $p < .001$).

Univariate normality. Table A.3 shows descriptive statistics for continuous and ordinal variables to be used in the analysis. These variables were screened for univariate normality by inspecting skewness and kurtosis values. Skewness with absolute value greater than three and kurtosis with absolute value greater than ten were considered potentially problematic, based on Kline's (2005) suggestions relying on the results of SEM simulation studies. Two of the participation variables showed extreme skewness: volunteering with youth (skewness = 3.2) and attending educational or training courses (skewness = 3.92). Only attending educational or training courses showed extreme kurtosis (kurtosis = 15.8). None of the life satisfaction or social connectedness indicators showed extreme skew or kurtosis. Among the covariates, only wealth and income showed problematic skewness and kurtosis: wealth had a skewness index of 9.4 with kurtosis of 135.0 while income had a skewness index of 7.9 and kurtosis of 109.5.

Since Mplus offers a version of maximum likelihood estimation (the MLR option) that provides chi-square statistics and standard errors that are robust to non-normality, untransformed participation variables were used in estimating the structural equation models even though two items showed extreme skewness. Muthén (2006) recommended against transforming data if the only purpose is to achieve a more normal distribution and suggested using the non-normality robust estimators instead. Wealth and income were transformed using logarithms, since these two variables are often most accurately modeled as predictors in linear regressions with such transformations; failing to transform them might violate the linearity assumption of SEM (Wooldridge, 2009). Logarithmically transformed wealth and income did not show extreme skewness or kurtosis except for the log of income, which showed kurtosis of 26.1.

One additional issue relating to univariate normality and linearity that was considered was that the response levels for participation did not represent a linearly increasing scale. Daily participation equates to a frequency of about 30 times a month, several times a week implies a frequency of perhaps 12 times a month (four weeks multiplied by three times per week), and once a week corresponds to a frequency of four or five times per month. Squaring the scale points from zero to five would create a scale that roughly represents actual monthly frequencies (Not in the last month = 0, Once a month = $1^2 = 1$, Several times a month = $2^2 = 4$, Once a week = $3^2 = 9$, Several times a week = $4^2 = 16$, and Daily = $5^2 = 25$). However, squaring the scale point values also increased kurtosis and skewness, sometimes to unacceptable amounts. When faced with actual frequency data such as that, a researcher might undertake a square root transformation to make the distributions more nearly normal. It makes some theoretical sense to leave the participation frequency variables as is, since we might expect that the biggest returns to participation will come from moving from no participation to some participation or from very little participation (once a month for example) to a moderate amount (once a week or several times a week). But at the highest levels of participation, the returns in terms of life satisfaction and social connectedness may flatten out; participating several times a week versus daily may look very similar in terms of associations with life satisfaction and social connectedness. Because of normality concerns and also the plausibility of nonlinearity in the relationship between frequency of participation and the outcomes, social participation responses were left untransformed.

Missing data. The life satisfaction and social connectedness indicators showed only small percentages of missing data, less than 5%, so their patterns of missingness were not analyzed. No cases were missing on gender, marital status, age, or race. Years of education, self-report of health, and the depression score were available for all but less than 1% of cases. Wealth and income were available for all cases, since the HRS provides imputed values where reported values are unavailable. Many social participation variables showed greater than five percent missing (see Table A.1). Of the sixteen activities, only reading, walking at least 20 minutes, and baking or cooking showed less than five percent missing. For the other variables, dummy variables indicating missingness were constructed. While missingness of participation variables was significantly related to the outcomes of social connectedness and life satisfaction, this missingness was not statistically significantly related to the outcome scores controlling for covariates such as marital status, self-report of health, and age. This suggests a missing-at-random (MAR) pattern of missingness, which is considered ignorable, meaning it can be dealt with using techniques such as multiple imputation or certain maximum likelihood algorithms, including those used by Mplus.

Mplus drops cases with missing exogenous variables. In the scale model, participation items are endogenous, because they are specified as reflecting latent variables. In the index model, participation items are conceptually exogenous, because they predict the composite participation variables. In the index model, if social participation variables were to be introduced without specifying measurement error as a fixed percentage of the sample variance as planned, cases with missing social participation responses would be dropped. But since social participation items will be modeled latently with a fixed residual variance, this implies that Mplus can impute values for missing items through its maximum likelihood algorithms. Thus, no multiple imputation for missing participation items was necessary in either the scale or the index models of social participation, so long as participation items in the index model were treated as endogenous, as they are when specifying measurement error as a fixed percentage of variance in the indicator.

Additional assumptions of SEM. The data were screened to ensure there were no extreme univariate outliers, that the assumption of linearity appeared to be met, and that no overly high pairwise multicollinearity existed. The data appeared to meet the assumptions of SEM, but given the skewness in some variables it was decided that the Mplus MLR algorithm would be used to ensure that standard errors and chi-square statistics were not distorted by a lack of normality. Following the suggestion of Meyers, Gamst, and Guarino (2006) for identifying multivariate outliers outside of the context of a specific model, a regression with the case ID as the dependent variable and all participation variables and covariates was run. Mahalanobis distance was calculated in this regression to identify potential multivariate outliers. Using a *p*-value of .001 (Tabachnick & Fidell, 2007), 60 out of 1368 available cases (about 4%) for males and 73 out of 1884 available cases (also about 4%) for females were identified as potential multivariate outliers. Note that many cases were dropped from the regression due to missing data (384 for men and 650 for women). This analysis suggested that there could be a sizable number of multivariate outliers. In order to check the sensitivity of results to the presence of such multivariate outliers, scale and index construction as well as structural model analyses were run first with the entire data set, then with multivariate outliers identified by their large log-likelihood contribution excluded. Since there was no evidence that outliers substantially changed results, which is not surprising given the use of an estimation algorithm that gives results that are robust to non-normality, results are reported for the entire data set without excluding outliers.

Scale Construction

Exploratory factor analysis. A preliminary principal components analysis on the exploratory half of the data set identified five factors with eigenvalues greater than one explaining a cumulative 51.6% of the variance in the items. Parallel analysis using principal axis factoring identified seven factors. The scree plot shown in Figure 16 suggested just one important factor. Factor identification was also run with the file split by gender; results were similar: the Kaiser criterion identified five factors in each subpopulation; parallel analysis identified seven factors; scree plots suggested just one important factor. In each case, the Kaiser-Meyer-Olkin measure showed acceptable sampling adequacy and Bartlett's test of sphericity rejected the null hypothesis of an identity correlation matrix.

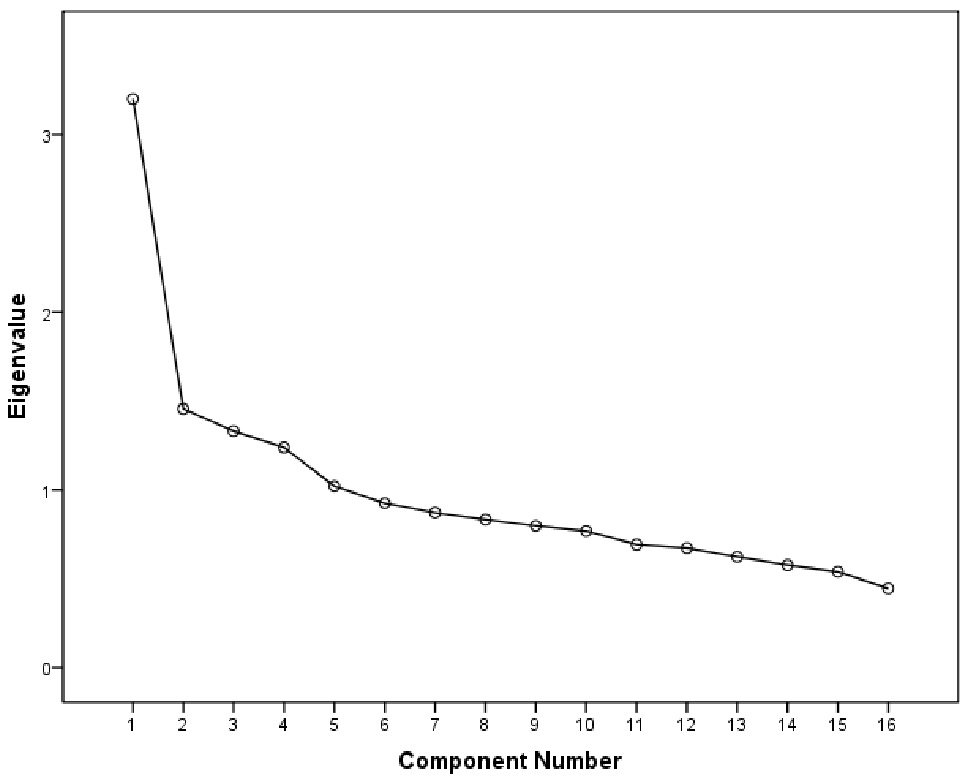


Figure 16. Scree plot from factor analysis of sixteen participation items

Given the equivocal results regarding the number of factors, EFA was run specifying four, five, six, and seven factor solutions to see which appeared most interpretable and whether any conformed to the hypothesized factor structure. While no

EFA criterion consulted specifically identified four factors, the hypothesized model from chapter two had just four (see Figure 10). Note that five, six, and seven factor solutions were also hypothesized as alternatives to be tested if the four-factor solution did not have adequate fit. Solutions were estimated with the entire measurement construction half and then by gender. In each solution, at least two factors were correlated at greater than .30, so a direct oblimin oblique rotation was used. The seven factor solution resulted in cross-loadings for some items and identified some single-item factors; also, it could not be estimated by gender using principal axis factoring. The six factor solution suffered from similar problems. Baking or cooking something special was a single-item factor, while there were two factors for activities out in the community, with attending meetings of non-religious organizations loading on both. Also, a six factor solution for males only could not be estimated. Because of a lack of parsimony and problems of estimation, the six and seven-factor EFA solutions were not reported nor interpreted, but they did imply that such solutions might be required to achieve good fit in the CFA models to be estimated after the EFA.

Table 1 compares the empirical four and five factor solutions to the hypothesized four factor model. The four-factor solutions explained about 45% of variance in the overall measurement data set, in the male subsample, and in the female subsample, while the five-factor solution explained around 52% of variance in each case. The four-factor solution estimated across the entire measurement data set and estimated for males only was mostly consistent with the hypothesized model, except that "Go to a sport, social or other club" loaded with the physical activity items. The four factor solution for females, however, was not as consistent with the hypothesized model. Three items did not have loadings greater than .30 on any factor: volunteer work with young people, doing writing, and using a computer. Domestic hobby tasks divided into two factors, one of which joined physical activity items with baking or cooking something special and with home maintenance or gardening, which seemed an unintuitive solution. Both the four and five factor solutions showed different patterns across male and female respondents, suggesting a lack of configural invariance by gender.

Table 1

EFA factor assignments for four and five-factor solutions

Item	Hypothesized factorial model	Four factors			Five factors		
		Overall	**Male**	**Female**	**Overall**	**Male**	**Female**
Volunteer work with children or young people	Community	Community	Community		Community	Community	
Other volunteer or charity work	Community	Community	Community	Community	Community	Community	Community
Go to a sport, social or other club	Community	Sports and exercise	Sports and exercise	Community		Gaming and socializing	Sports and exercise
Attend meetings of non-religious organizations	Community	Community	Community	Community	Community	Community	Community
Attend educational or training course	Intellectual	Community	Community	Community	Community	Community	Community
Read books, magazines, or newspapers	Intellectual	Intellectual	Intellectual	Intellectual	Intellectual		Intellectual
Do word games	Intellectual	Intellectual	Intellectual	Intellectual	Intellectual	Gaming and socializing	Intellectual
Play cards or games such as chess	Intellectual	Intellectual	Intellectual	Intellectual	Intellectual	Gaming and socializing	Intellectual
Do writing	Intellectual		Intellectual			Intellectual	
Use a computer for email, Internet or other tasks	Intellectual	Intellectual	Intellectual			Intellectual	Intellectual
Do home or car maintenance or gardening	Home and hobbies	Home and hobbies	Home and hobbies	First hobby factor	Outdoor activities	Home and hobbies	Home
Bake or cook something special	Home and hobbies	Home and hobbies	Home and hobbies	First hobby factor			Home
Make clothes, knit, embroider, etc.	Home and hobbies	Home and hobbies		Second hobby factor	Hobbies		Hobbies
Work on a hobby or project	Home and hobbies	Home and hobbies	Home and hobbies	Second hobby factor	Hobbies	Intellectual or Home and Hobbies	Hobbies
Play sports or exercise	Sports and exercise	Sports and exercise	Sports and exercise	First hobby factor	Outdoor activities	Sports and exercise	Sports and exercise
Walk for 20 minutes or more	Sports and exercise	Sports and exercise	Sports and exercise	First hobby factor	Outdoor activities	Sports and exercise	Sports and exercise

Confirmatory factor analysis. Four, five, six, and seven factor hypothesized models as described in chapter two were estimated across the entire measurement data set. While this used confirmatory techniques, the analysis at this stage was still data-driven, so the measurement construction half of the data set was used and the validation half of the data set was saved for confirming the final model. As described in chapter two, modifications of the model (both a priori as hypothesized and ad hoc as informed by EFA results) were pursued until acceptable fit was achieved, specified as CFI greater than or equal to .90 and RMSEA less than .10. RMSEA values were generally below the cutoff for all tested models but CFI values were mostly inadequate. Fit statistics and chi-square difference tests for the CFA models fitted across both genders (n = 1791) are shown in Table 2. All of the chi-square statistics were significant, $p < .0001$. Because the MLR estimation algorithm used for fitting these models produced a scaled chi-square, chi-square difference tests needed to be corrected and are not simple differences between the reported chi-squares (“Chi-square difference testing”, n.d.). The five factor hypothesized model, which split community organization participation into a volunteering factor and a clubs factor, was significantly better than the four factor model. The six factor model which started from the four factor model but split intellectual participation into games, writing and reading, and education was also significantly better than the four factor model, $p < .0001$. The seven factor model significantly improved upon the six factor model, $p < .0001$, but the fit was still inadequate according to the CFI value, $\chi^2(85)=550.78$, $\chi^2/df = 6.5$, RMSEA=.055, CFI=.83, SRMR=.044. In the interests of achieving a parsimonious model, an alternative six factor model with the single-item factor education merged into the clubs factor, representing social activities in the community, was compared to the seven factor model. This model was significantly worse than the seven factor model.

Modification indexes were inspected to see if any respecifications might achieve adequate fit. Starting from the seven-factor model, a residual correlation was added for the two items mentioning sports ("Go to a sport, social or other club" and "Play sports or exercise"). This significantly improved the fit relative to the seven-factor model, but the fit was still inadequate according to its CFI value, $\chi^2(84)=492.85$, $\chi^2/df = 5.9$, RMSEA=.052, CFI=.85, SRMR=.044. BIC values supported choosing the final seven factor model with the one correlated residual over any other tested models. No other suggested modifications appeared to make substantive sense. Given the inadequate fit of the overall model as well as the results of the EFA which suggested lack of invariance across genders, factor models were developed separately for males and females.

Table 2

Scale construction CFA results – All respondents – Measurement construction subsample

No.	***Model***	***χ^2(df)***	***RMSEA***	***CFI***	***SRMR***	***BIC***	***Compare to***	***Corrected $\Delta\chi^2$***	***Δdf***	***p***
A1	Four factors	773.82(98)***	0.062	0.75	0.054	101,091				
A2	Five factors	713.11(94)***	0.061	0.78	0.052	101,032	A1	54.99	4	<.001
A3	Six factors	613.12(90)***	0.057	0.81	0.045	100,931	A1	149.35	8	<.001
A4	Seven factors	550.78(85)***	0.055	0.83	0.043	100,866	A3	52.24	5	<.001
A5	Merge education into clubs factor	568.02(89)***	0.055	0.83	0.044	101,074	A4	20.23	4	<.001
A6	Starting from A4, add correlated residual for items referencing sports	492.85(84)***	0.052	0.85	0.041	100,805	A4	135.56	1	<.001

Note: *** $p < .001$.

CFA with male subsample. Results for males (n = 787) are shown in Table 3. As with the models fit to the entire sample, addition of factors improved the fit significantly. However, merging the single-item education factor did not significantly worsen the fit relative to the seven-factor model, corrected $\Delta\chi^2(4) = 1.00$, $p = .91$. Thus for men, "Attend educational or training course" loaded with "Go to a sport, social or other club" and "Attend meetings of non-religious organizations." From this six-factor model, adding the correlated residual for the two items representing sports significantly improved the fit so that it was acceptable. The resultant six-factor model is shown in Figure 17. This model showed adequate but not good fit, $\chi^2(88) = 166.07$, $\chi^2/\text{df} = 1.9$, $p < .0001$, RMSEA=.034, CFI=.93, SRMR=.035. The BIC value for this model was lower than any of the other tested model, providing additional justification for using this model rather than a more parsimonious one, which might be more conceptually pleasing.

Table 3

Scale construction CFA results – Male– Measure construction half

No.	***Model***	***$\chi^2(df)$***	***RMSEA***	***CFI***	***SRMR***	***BIC***	***Compare to***	***Corrected $\Delta\chi^2$***	***Δdf***	***p***
M1	Four factor	280.19(98)***	0.049	0.85	0.047	41,750				
M2	Five factor	246.75(94)***	0.045	0.87	0.044	41,731	M1	28.74	4	<.001
M3	Six factor	231.99(90)***	0.045	0.88	0.041	41,738	M1	43.90	8	<.001
M4	Seven factor	192.00(85)***	0.040	0.91	0.037	41,715	M3	31.89	5	<.001
M5	Merge education into clubs factor (alternate six factor)	189.06(89)***	0.038	0.92	0.037	41,690	M4	1.00	4	0.91
M6	Starting from M5 (six factors), add correlated residual for items referencing sports	166.07(88)***	0.034	0.93	0.035	41,667	M5	19.32	1	<.001

Note: *** $p < .001$.

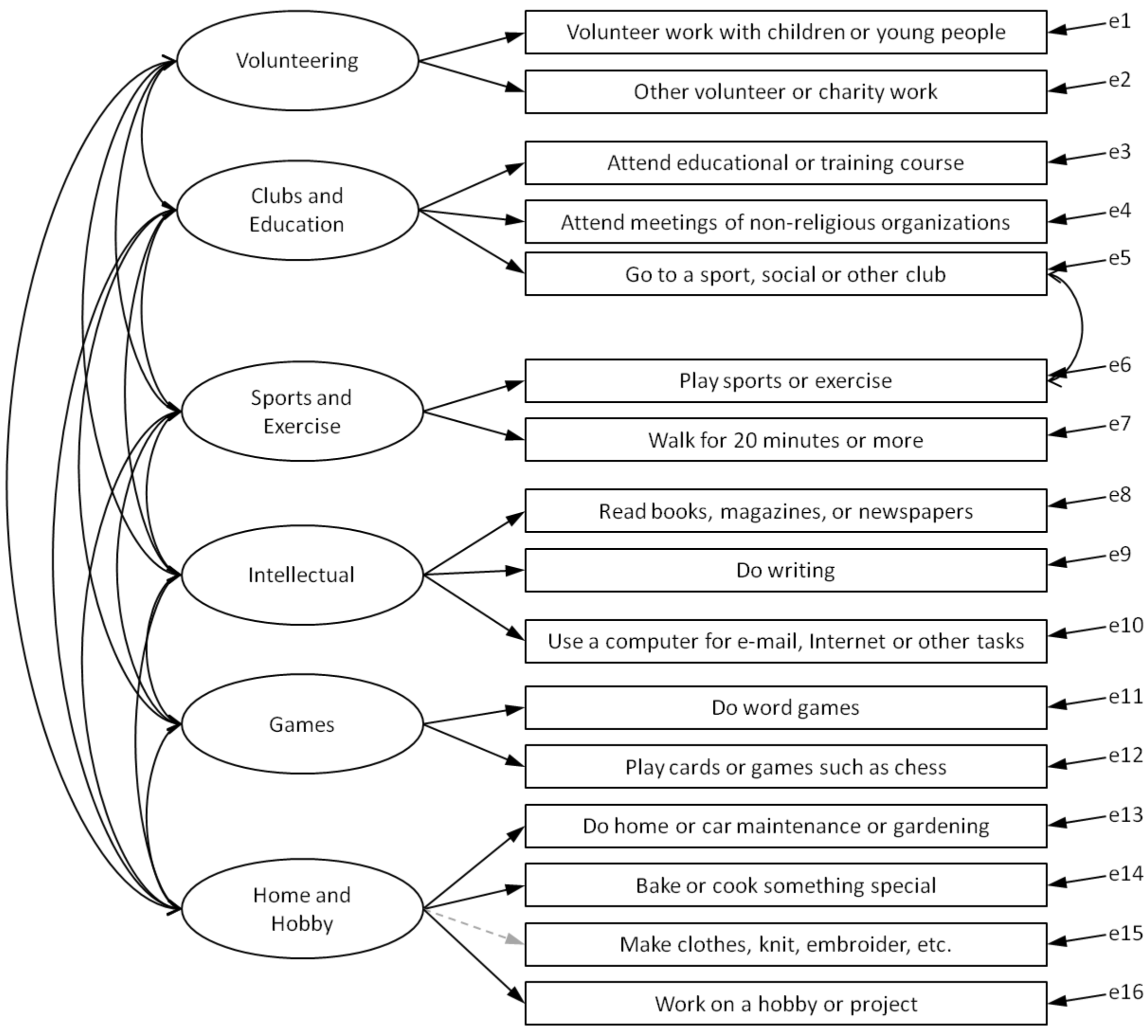


Figure 17. Scale model for social participation – Male – Before dropping items

Estimated model parameters for the six-factor model for males are reported in Table B.1 in Appendix B. Standardized factor loadings for significant loadings ranged from a low of .33 for "Volunteer work with children or young people" on the volunteering factor to a high of .87 for "Other volunteer or charity work," also on the volunteering factor. "Bake or cook something special" also had a low standardized factor loading of .33. "Make clothes, knit, embroider, etc" did not load significantly on Home and Hobbies, p = .51. All factors covaried significantly with each other, p < .001, except for games with volunteering. Standardized residual variances were rather high; for example, “Do word games” had a standardized residual variance of .76 representing 76% variance unexplained and “Read books, magazines, or newspapers had a standardized residual variance of .87. The lowest standardized residual variance was .24, for “Other volunteer or charity work.”

Alpha reliabilities (Table 4) were quite low, ranging from .35 for the games subscale to .64 for sports and exercise. These low reliabilities are related to the small numbers of items per factor but also reflect a fundamental mismatch between the common factor model and the participation data. For example, the originally hypothesized five-item intellectual factor (consisting of reading, word games, cards/chess/other games, writing, and computer usage) which was partially supported in the four-factor EFA solution showed reliability of just .54 for male respondents in the measurement construction subsample, which is still far from adequate.

Table 4

Alpha reliabilities for participation subscales – Male – Measure construction subsample

Subscale	***Cronbach’s α***
Volunteering	.412
Clubs	.451
Sports and exercise	.641
Intellectual	.455
Games	.352
Home and hobbies	.482
Home and hobbies (with sewing/knitting deleted)	.536

Given the non-significant loading for the sewing/knitting item and the low standardized factor loading for the volunteering with youth and baking/cooking items, a trimmed model without these items was estimated. This made volunteering into a single-item factor. The fit for this refined model in the measurement construction half was good,

$\chi^2(50) = 93.60$, $p = .0002$, RMSEA=.033, CFI=.96, SRMR=.032. This trimmed model was fit in the confirmatory half of the data set for validation purposes ($n = 871$). It had adequate fit, $\chi^2(50) = 123.99$, $p < .0001$, RMSEA=.041, CFI=.92, SRMR=.038. The final scale model for males is shown in Figure 18. Parameter estimates for the final model fit with the validation subsample are shown in Table 5, Table 6, and Table 7. All items loaded significantly on their factors. All factors covaried significantly with each other except Volunteering with Games, Home/Hobbies with Games, and Sports with Games. As in the measurement construction half, standardized residual variances for the indicators were rather high, ranging from a low of .33 for "play sports or exercise" to a high of .89 for "read books, magazines, or newspapers." Alpha reliabilities are shown in Table *8*. As in the measure construction half, they were poor, ranging from a low of .35 for Games to a high of .61 for Sports/Exercise.

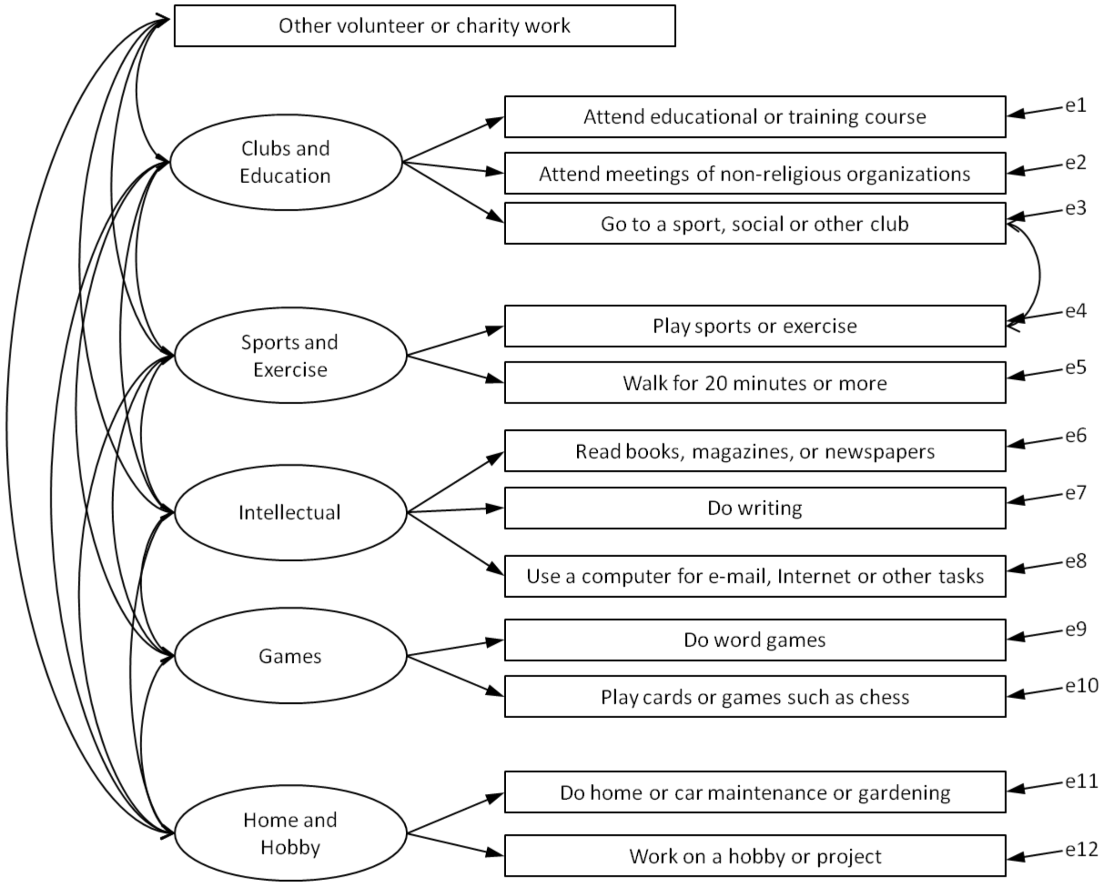


Figure 18. Final six-factor, 13-item scale model for male respondents

Table 5

Factor loadings for final scale model – Male – Validation subsample

	Unstandardized		*Standardized*	
Factor Loadings	***Coefficient***	***SE***	***Coefficient***	***SE***
Social by				
Attend educational or training course	1.00	0.000	0.52***	0.097
Go to a sport, social or other club	1.46***	0.302	0.40***	0.057
Attend meetings of non-religious organizations	1.49***	0.266	0.68***	0.074
Games by				
Do word games	1.00	0.000	0.52***	0.133
Play cards or games such as chess	0.61	0.315	0.37**	0.109
Intellectual by				
Read books, magazines, or newspapers	1.00	0.000	0.33***	0.045
Do word games	2.01***	0.431	0.59***	0.053
Use a computer for e-mail, Internet, or other tasks	2.84***	0.522	0.55***	0.041
Home and Hobbies by				
Do home or car maintenance or gardening	1.00	0.000	0.53***	0.061
Work on a hobby or project	1.49***	0.307	0.75***	0.078
Sports and Exercise by				
Play sports or exercise	1.00	0.000	0.82***	0.084
Walk for 20 minutes or more	0.63***	0.134	0.53***	0.065

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6

Factor covariances and residual correlations for final scale model – Male – Validation subsample

	Unstandardized		*Standardized*	
	Coefficient	***SE***	***Coefficient***	***SE***
Factor Covariances				
Volunteering with Social	0.30***	0.092	0.61***	0.046
Volunteering with Games	0.12	0.094	0.10	0.074
Volunteering with Intellectual	0.22***	0.045	0.39***	0.060
Volunteering with Home and Hobbies	0.26***	0.074	0.22***	0.051
Volunteering with Sports and Exercise	0.42***	0.108	0.20***	0.043
Games with Social	0.14*	0.059	0.36**	0.120
Intellectual with Social	0.10***	0.029	0.59***	0.068
Intellectual with Games	0.17*	0.070	0.39***	0.106
Home and Hobbies with Social	0.08*	0.035	0.21**	0.077
Home and Hobbies with Games	0.16	0.105	0.18	0.092
Home and Hobbies with Intellectual	0.21***	0.049	0.50***	0.065
Sports with Social	0.19**	0.064	0.30***	0.059
Sports with Games	0.09	0.157	0.06	0.093
Sports with Intellectual	0.31***	0.080	0.43***	0.069
Sports with Home and Hobbies	0.59***	0.164	0.38***	0.095
Correlated Residuals				
"Go to a sport, social or other club" with "Play sports or exercise"	0.51***	0.096	0.34***	0.089
Factor Variances				
Volunteering	1.64***	0.131		
Social	0.15	0.079		
Games	0.94	0.488		
Intellectual	0.20**	0.064		
Home and Hobbies	0.87***	0.207		
Sports and Exercise	2.71***	0.562		

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Residual variances for final scale model – Male – Validation subsample

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Attend an educational or training course	0.41***	0.064	0.73***	0.100
Go to a sport, social or other club	1.71***	0.132	0.84***	0.045
Attend meetings of non-religious organizations	0.39***	0.051	0.54***	0.100
Read books, magazines, or newspapers	1.65***	0.136	0.89***	0.030
Do word games	2.54***	0.492	0.73***	0.139
Play cards or games such as chess	2.19***	0.234	0.86***	0.081
Do writing	1.53***	0.164	0.66***	0.063
Use a computer	3.76***	0.246	0.70***	0.045
Do home or car maintenance or gardning	2.19***	0.200	0.72***	0.066
Work on a hobby or project	1.45***	0.394	0.43***	0.118
Play sports or exercise	1.35*	0.555	0.33*	0.137
Walk for 20 minutes or more	2.77***	0.267	0.72***	0.069

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 8

Alpha reliabilities for participation subscales – Male – Validation Subsample

Subscale	***Cronbach's α***
Clubs	.453
Sports and exercise	.612
Intellectual	.443
Games	.353
Home and hobbies	.527

CFA with female subsample. In the female subsample (n = 1004), an eight-factor model was developed. Starting from the four-factor hypothesized model, additional factors (both those hypothesized a priori and those suggested by EFA results) and correlated residuals were added until adequate fit was reached, defined as CFI greater than or equal to .90 and RMSEA less than .10. Fit statistics and chi-square difference tests are shown in Table 9. Addition of factors improved the model up through seven factors. Merging the single-item education factor with the clubs items significantly worsened the fit, corrected $\Delta\chi^2(4) = 26.98$, $p < .0001$. The seven factor model that specified education as a single-item factor did not show adequate fit according to its CFI value, $\chi^2(85) = 429.31$, χ^2 /df = 5.05, $p < .0001$, RMSEA=.064, CFI=.79, SRMR=.053. Based on results of the EFA, the home and hobbies factor was split into two factors. This significantly improved the fit. Once a correlated residual was added for the items referencing sports, a model was achieved that had adequate fit, $\chi^2(77) = 243.10$, $p < .0001$, χ^2 /df = 3.2, RMSEA=.046, CFI=.90, SRMR=.037. BIC values which trade off model fit versus model parsimony supported the choice of the eight-factor model with the correlated residual for sports items, since it had the lowest BIC of any of the models tested. The resultant eight-factor model for females is shown in Figure 19. Relative to the male model, this model had an additional single-item factor for education, measured by "Attend educational or training course" and had separate factors for home and hobby participation. The male model had just one factor for home and hobby participation.

Table 9

Scale construction CFA results – Female – Measurement construction subsample

No.	***Model***	$\chi^2(df)$	***RMSEA***	***CFI***	***SRMR***	***Compare to***	***BIC***	$\Delta\chi^2$	Δdf	**p**
F1	Four factor	567.68(98)***	0.069	0.72	0.062		57,415			
F2	Five factor	538.19(94)***	0.069	0.73	0.060	F1	57,396	27.83	4	<.001
F3	Six factor	460.42(90)***	0.064	0.78	0.054	F1	57,328	98.97	8	<.001
F4	Seven factor	429.31(85)***	0.064	0.79	0.052	F3	57,300	28.60	5	<.001
F5	Merge education into clubs factor	460.87(89)***	0.054	0.80	0.053	F4	57,328	26.98	4	<.001
F6	Starting from F4, split Home and Hobbies into two factors	273.90(78)***	0.050	0.88	0.038	F5	57,167	168.94	7	<.001
F7	Starting from F6, add correlated residual for items referencing sports	243.10(77)***	0.046	0.90	0.037	F6	57,140	277.10	1	<.001

Note: *** $p < .001$.

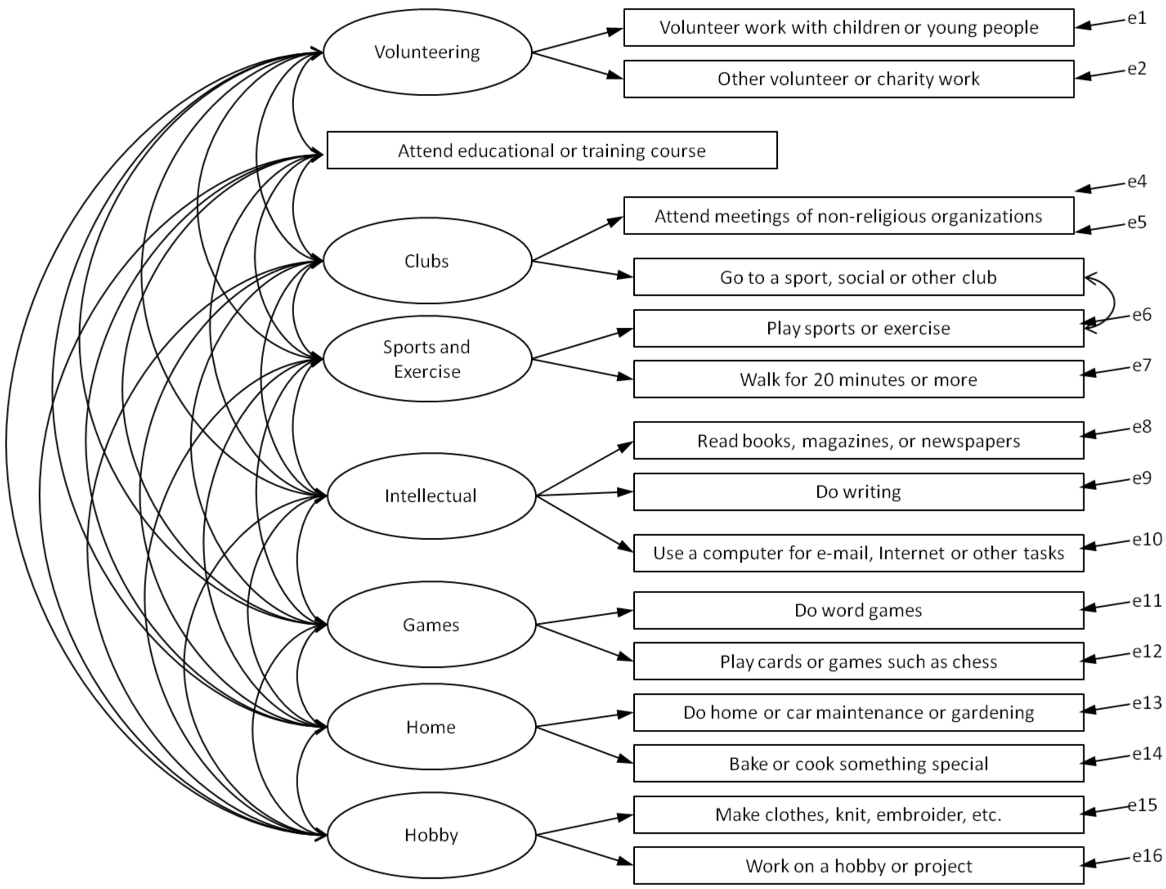


Figure 19. Scale model for social participation – female respondents

Estimated model parameters for the eight-factor model for females fit in the measurement construction subsample are reported in Table B.2 in Appendix B. Standardized factor loadings ranged from a low of .32 for “Volunteer work with children or young people” on the volunteering factor to a high of .97 for “Work on a hobby or project” on the hobbies factor. All indicators loaded significantly on their factors, p < .001. All factors were significantly correlated with each other at the $p < .01$ level. “Volunteer work with children or young people” had the highest standardized residual variance, .90. Residual variance for “Work on a hobby or project” did not differ significantly from zero, $p = .699$. Many items had standardized residual variance greater than .70, suggesting the factor model did not explain a large amount of variance in the indicators. Alpha reliabilities (Table 10) were inadequate, ranging from a low of .368 for volunteering to a high of .642 for hobbies. As with the male model, this is related both to the small number of items per factor and the generally poor fit of the common factor model to the participation data. A five-item intellectual factor consisting of reading, word games, cards/chess/other games, writing, and computer usage showed an alpha of .52, considerably higher than the final three-item intellectual factor's alpha of .42, but still far from adequate.

Table 10

Alpha reliabilities for participation subscales – Female – Measure construction subsample

Subscale	***Cronbach's*** α
Volunteering	.368
Clubs	.413
Sports and exercise	.578
Intellectual	.432
Games	.419
Home	.550
Hobbies	.642

Validating the scale model. Given the low standardized factor loading for the volunteering with youth item, a trimmed model without it was estimated. This made volunteering into a single-item factor. The fit was adequate, $\chi^2(65) = 209.223$, χ^2 /df = 3.2, $p < .0001$, RMSEA=.047, CFI=.91, SRMR=.037. This trimmed model was fit in the confirmatory half of the data set for validation purposes ($n = 1281$). It had good fit, $\chi^2(65) = 159.017$, $p < .0001$, RMSEA=.034, CFI=.95, SRMR=.031. The final scale model for females is shown in Figure 20. Parameter estimates for the final model for female respondents fit in the validation subsample are shown in Table 11, Table 12, Table 13, and Table 14. Alpha reliabilities are shown in Table 15.

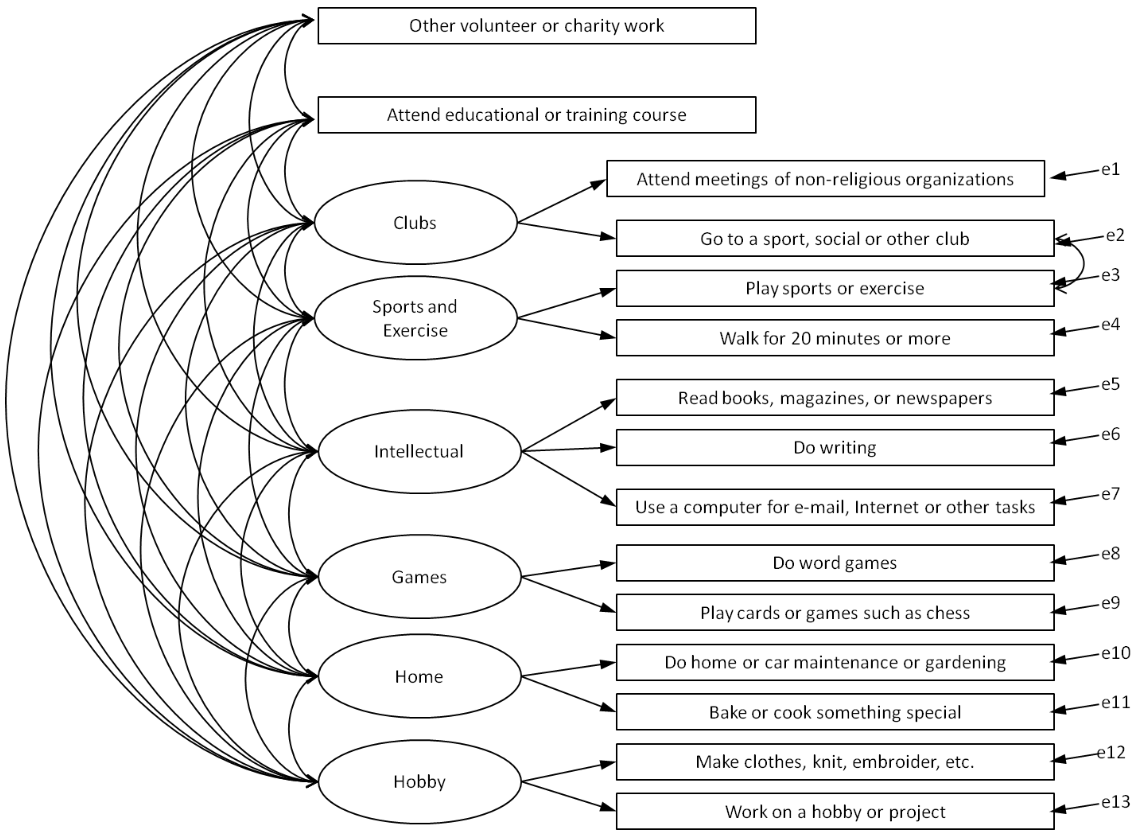


Figure 20. Final eight-factor, 15-item scale model for females

Table 11

Factor loadings for final scale model – Female – Validation subsample

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Clubs by				
Go to a sport, social or other club	1.00	0.00	0.61***	0.06
Attend meetings of non-religious organizations	0.69***	0.11	0.63***	0.05
Games by				
Do word games	1.00	0.00	0.49***	0.05
Play cards or games such as chess	0.97***	0.14	0.59***	0.05
Intellectual by				
Read books, magazines, or newspapers	1.00	0.00	0.37***	0.04
Do writing	1.91***	0.30	0.56***	0.03
Use a computer for e-mail, Internet or other tasks	2.13***	0.33	0.46***	0.04
Home by				
Do home or car maintenance or gardening	1.00	0.00	0.65***	0.04
Bake or cook something special	0.70***	0.08	0.53***	0.04
Hobbies by				
Make clothes, knit, embroider, etc.	1.00	0.00	0.54***	0.04
Work on a hobby or project	2.14***	0.24	0.93***	0.04
Sports by				
Play sports or exercise	1.00	0.00	0.74***	0.04
Walk for 20 minutes or more	0.86***	0.08	0.65***	0.04

Note. * p < .05. ** p < .01. *** p < .001.

Table 12

Factor covariances and residual correlations for final scale model – Female – Validation subsample

	Unstandardized		*Standardized*	
	Coefficient	***SE***	***Coefficient***	***SE***
Factor Covariances				
Education with Clubs	0.34***	0.06	0.48***	0.07
Education with Intellectual	0.13***	0.03	0.31***	0.05
Education with Home	0.17***	0.05	0.15***	0.04
Education with Hobbies	0.12***	0.04	0.18***	0.04
Education with Sports and Exercise	0.31***	0.06	0.25***	0.04
Volunteering with Clubs	0.45***	0.06	0.43***	0.06
Volunteering with Intellectual	0.21***	0.04	0.36***	0.05
Volunteering with Home	0.24**	0.07	0.15***	0.04
Volunteering with Hobbies	0.24***	0.05	0.23***	0.04
Volunteering with Sports and Exercise	0.48***	0.09	0.26***	0.04
Games with Clubs	0.35***	0.07	0.40***	0.07
Intellectual with Clubs	0.19***	0.04	0.47***	0.08
Intellectual with Games	0.33***	0.08	0.66***	0.08
Home with Clubs	0.45***	0.09	0.42***	0.06
Home with Games	0.42***	0.10	0.31***	0.07
Home with Intellectual	0.33***	0.07	0.53***	0.07
Hobbies with Clubs	0.29***	0.06	0.42***	0.06
Hobbies with Games	0.32***	0.07	0.37***	0.06
Hobbies with Verbal	0.24***	0.04	0.61***	0.05
Hobbies with Home	0.55***	0.09	0.53***	0.06
Sports and Exercise with Clubs	0.56***	0.11	0.45***	0.06
Sports and Exercise with Games	0.37***	0.09	0.24***	0.06
Sports and Exercise with Verbal	0.35***	0.06	0.50***	0.06
Sports and Exercise with Home	1.09***	0.12	0.58***	0.05
Sports and Exercisewith Hobbies	0.47***	0.08	0.39***	0.05
Volunteering with Education	0.29***	0.06	0.28***	0.05
Correlated Residual				
"Go to a sport, social or other club" with "Play sports or exercise"	0.27***	0.08	0.19***	0.05

Note. * p < .05. ** p < .01. *** p < .001.

Table 13

Factor variances for final scale model – Female – Validation subsample

Factor	***Variance***	***S.E.***	***p***
Education	1.56	0.09	<.001
Volunteering	0.71	0.09	<.001
Clubs	0.71	0.14	<.001
Games	1.09	0.21	<.001
Intellectual	0.23	0.06	<.001
Home	1.64	0.21	<.001
Hobbies	0.67	0.12	<.001
Sports and Recreation	2.17	0.23	<.001

Table 14

Residual variances for final scale model – Female – Validation subsample

	Unstandardized		***Standardized***	
Participation item	***Coefficient***	***SE***	***Coefficient***	***SE***
Go to a sport, social or other club	1.17***	0.13	0.62***	0.07
Attend meetings of non-religious organizations	0.50***	0.06	0.60***	0.07
Read books, magazines, or newspapers	1.41***	0.11	0.86***	0.03
Do word games	3.54***	0.22	0.76***	0.05
Play cards or games such as chess	1.90***	0.17	0.65***	0.06
Do writing	1.80***	0.11	0.68***	0.04
Use a computer	3.88***	0.18	0.79***	0.04
Do home or car maintenance or gardning	2.21***	0.20	0.57***	0.05
Bake or cook something special	2.02***	0.13	0.72***	0.04
Make clothes, knit, embroider, etc.	1.62***	0.11	0.71***	0.04
Work on a hobby or project	0.46	0.29	0.13***	0.08
Play sports or exercise	1.75***	0.22	0.45***	0.06
Walk for 20 minutes or more	2.20***	0.18	0.58***	0.05

Note. * p < .05. ** p < .01. *** p < .001.

Table 15

Alpha reliabilities for participation subscales – Female – Validation subsample

Subscale	***Cronbach's α***
Clubs	.517
Sports and exercise	.633
Intellectual	.422
Games	.442
Home	.490
Hobbies	.643

Alternate female models. Given the extreme complexity of the eight-factor model for females developed using EFA and (exploratory-style) CFA, two alternate, more parsimonious models for females were developed. The first model had four factors, as shown in Figure 21. This model had poor fit according to its CFI value but adequate fit by RMSEA and SRMR, $\chi^2(71) = 406.080$, $p < .001$, RMSEA = .069, CFI = .77, SRMR = .059. Kenny (2011) asserted that CFI values should not be interpreted if baseline correlations across items are so low that the null model (model with no correlations specified) itself has RMSEA less than about .16. The null model for females with all social participation items specified as uncorrelated had an RMSEA of .12, suggesting that low obtained CFI values reflected low underlying correlations among the participation items. This, in itself, casts doubt on whether any scale model is appropriate for the participation data, given such low correlations. The fact that chi-square difference tests as well as BIC values selected the more complex factor models adds weight to the argument that four- or five-factor solutions are not justified. However, in the interests of comparing both more parsimonious and less parsimonious scale models with the index models, this alternative model was also introduced into comparisons.

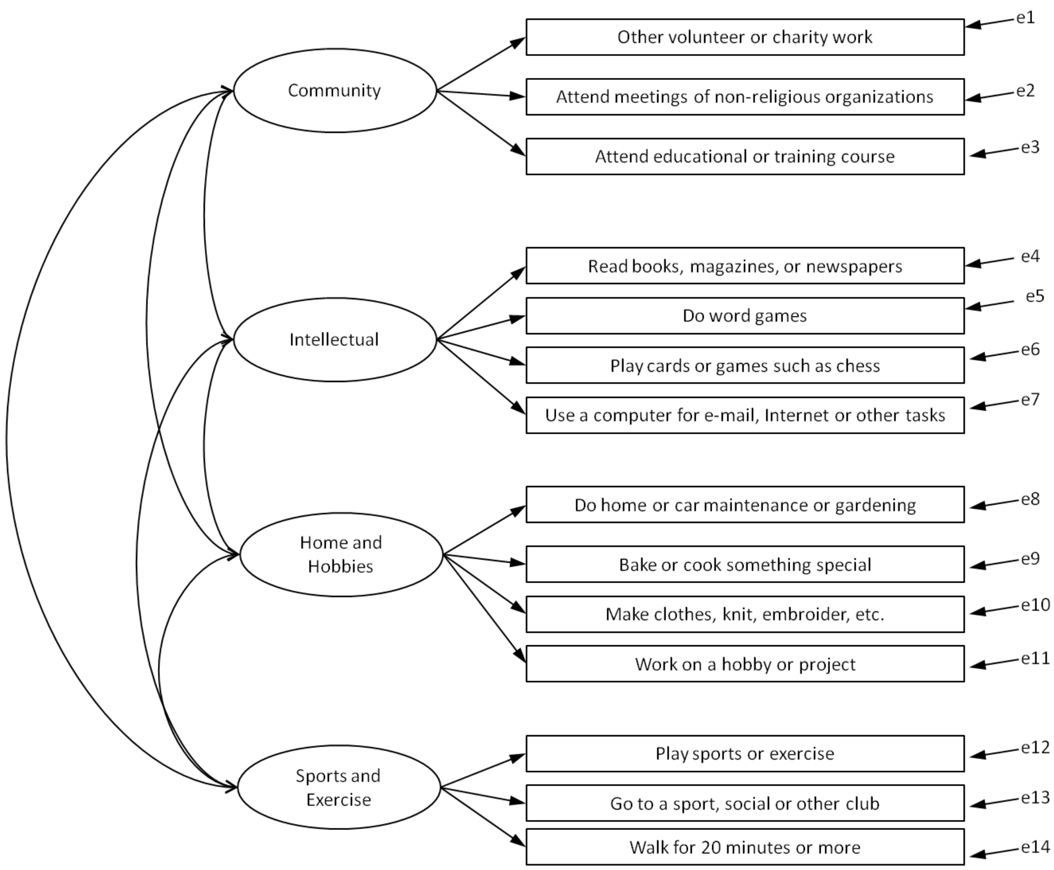


Figure 21. Four-factor alternate scale model for females

For this four-factor model, two items were dropped: volunteer work with children, since this was shown to be a poor item in the analyses already run, and do writing, since this item had not loaded on any factors in the four or five-factor EFA solutions. This left three or four items on each factor and did not include any correlated residuals. In the validation half of the data set, this model still had poor fit according to CFI value, but adequate otherwise, $\chi^2(71) = 368.00$, $p < .001$, RMSEA = .057, CFI = .83, SRMR = .053. Parameter estimates for the alternate scale model for females estimated in the validation half of the data set are reported in Table B.3 in Appendix B. Standardized factor loadings were generally low, less than the .60 level recommended by Kline (2005). Only "play sports or exercise" and "walk for 20 minutes" on the Sports/Exercise factor and "work on a hobby or project" on the Home/Hobbies factor had standardized factor loadings greater than .60. Standardized residual variances were consequently high; for example, 83% of the variance of "bake or cook something special" remained unexplained as did 81% of the variance of "use a computer for email, Internet, or other tasks." Alpha reliabilities for the four subscales in this model using the confirmatory half of the data set were .53 for Community, .60 for Sports/exercise, .47 for Intellectual, and .61 for Home/Hobbies.

The second alternate model was a one-factor model of participation. Scree plots for EFA conducted across both genders and by gender each suggested the presence of just one important factor. No other criterion, however, identified just one factor in the data set. For females, the first factor explained just 20.1% of the variance in the participation items, so it seems unlikely that a researcher would choose such a solution. However, it does offer an extremely parsimonious representation and might result in a useful scale model of participation. Furthermore, it allows one to explore participation in general rather than along specific dimensions. The one-factor model using all items except the volunteering with youth item (dropped for poor performance in earlier analyses) had poor fit, $\chi^2(90) = 715.91$, $\chi^2/df = 8.0$, RMSEA = .083, CFI = .61, SRMR = .069. All items significantly loaded on the single factor at $p < .001$, but standardized loadings were low. The highest loading was for the hobbies/projects item ($\beta = .58$, $SE = .04$, $p < .001$). All other standardized loadings were below .5. Cronbach's alpha for this 15-item factor was .70. Reliability analysis was undertaken to see if reliability could be increased with the elimination of items. It could not be, so the single-factor participation scale model for females was left as is with all of the original items except the volunteering with youth item. In the validation half of the data set, this model showed poor fit according to its CFI value but adequate fit otherwise, $\chi^2(90) = 676.16$, $\chi^2/df = 7.5$, RMSEA = .071, CFI = .69, SRMR = .065. Reliability in the validation half was .74.

Summary of scale construction. EFA and CFA techniques using the exploratory half of the data set were used to develop a six factor, 13-item model for males and an eight factor, 15-item model for females. Additionally, two alternate simpler models were developed for females: a four-factor model and a single-factor model. Criteria for determining number of factors were equivocal and the hypothesized four-factor model fit poorly (as did the alternate female four-factor model). Chi-square difference tests supported the addition of factors and one correlated residual to each model. The models showed adequate to good fit when estimated with the hold-out validation subsample but each of the main models had many small factors including one single-item factor for the model for males and two single-item factors for the model for females. There was not configural invariance across gender: the basic factor structures appeared to differ for men and women. The main scale models did not meet the criteria for unidimensional measurement as defined by Anderson and Gerbing (1988) which requires that each indicator load on a single factor and that error terms are independent; each had a correlated residual for the items that referred to sports. The alternate female model did meet the criteria for unidimensional measurement. Relative to the six factor model for males, the eight-factor model for females split the home and hobbies factor into two factors and specified education as a single-item factor rather than loading it with sports/social club and non-religious organization participation. The alternate four-factor model for females dropped the volunteer with children and do writing items, loaded sports/social clubs on the Sports/Exercise factor, and loaded education/training on the Community factor. The subscales for both men and women showed low reliability, ranging from a low of .35 for the Games factor for males to a high of .64 for the Hobbies factor for females in the eight-factor female model.

Index Construction

Measurement modeling of outcomes. The index model used the outcomes of life satisfaction and social connectedness for identification purposes. Before constructing indexes, measurement modeling of these outcome variables was undertaken to ensure that the fit of the outcome measurement model did not detract from the overall fit of the index model or from the structural models used in comparing the scale and index models. When fit to the exploratory half of the data set (n = 1790), the two-factor model of life satisfaction and perceived social connectedness with five life satisfaction items and seven social connectedness items had adequate fit, $\chi^2(53) = 468.96$, $p < .0001$, RMSEA = .066, CFI = .93, SRMR=.046. In order to achieve good fit (defined as CFI greater than or equal to .95), correlated residuals were added as suggested by modification indexes if they made substantive sense. Two pairs of correlated residuals were added within each factor. The final measurement model for the outcomes is shown in Figure 22. This model had good fit, $\chi^2(49)=119.53$, $p < .0001$, RMSEA = .028, CFI = .99, SRMR = .036. Parameter estimates generated using the measure construction subsample are reported in Table 16.

The model showed good fit in the validation subsample as well, $\chi^2(49)=132.99$, $p < .0001$, RMSEA = .028, CFI = .99, SRMR = .030.

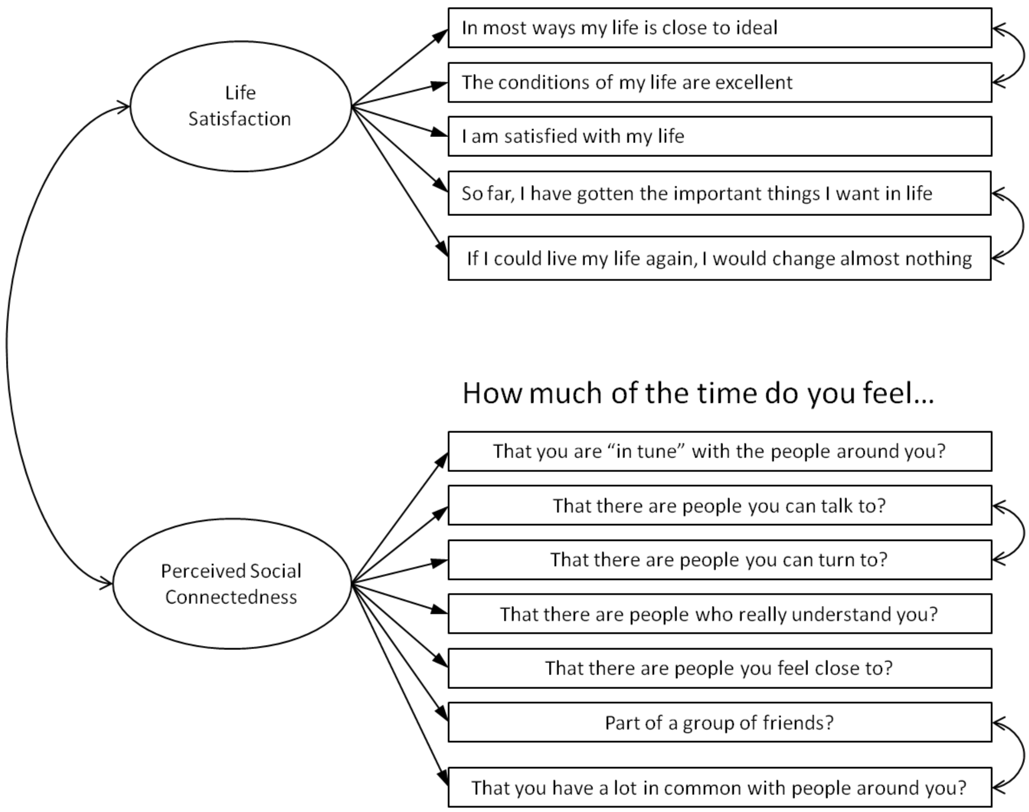


Figure 22. Refined measurement model for outcome variables

Table 16

Parameter estimates for life satisfaction and social connectedness measurement model

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Factor loadings				
Life satisfaction by				
LS1. Life close to ideal	1.00***	0.000	0.76***	0.021
LS2. Conditions of life excellent	1.13***	0.033	0.83***	0.015
LS3. Satisfied with my life	1.09***	0.043	0.90***	0.012
LS4. Have gotten the important things I wanted	0.81***	0.046	0.70***	0.023
LS5. Would change almost nothing	0.77***	0.047	0.52***	0.026
Social connectedness by				
SC1. In tune with people	1.00***	0.000	0.52***	0.024
SC2. People you can talk to	1.28***	0.070	0.73***	0.019
SC3. People you can turn to	1.36***	0.072	0.78***	0.016
SC4. People who understand you	1.41***	0.078	0.78***	0.015
SC5. People you feel close to	1.32***	0.077	0.78***	0.018
SC6. Part of a group of friends	1.34***	0.074	0.64***	0.021
SC7. Have a lot in common with people around you	1.25***	0.065	0.66***	0.019
Factor covariances				
Social connectedness with Life satisfaction	0.16***	0.020	0.34***	0.033
Residual correlations				
SC2 with SC3	0.06***	0.008	0.37***	0.041
SC6 with SC7	0.09***	0.011	0.33***	0.034
LS1 with LS2	0.50***	0.077	0.40***	0.043
LS4 with LS5	0.48***	0.081	0.24***	0.037
Factor variances				
Life satisfaction	1.90***	0.133		
Social connectedness	0.12***	0.012		
Residual variances				
LS1. Life close to ideal	1.44***	0.109	0.43***	0.032
LS2. Conditions of life excellent	1.07***	0.086	0.31***	0.024
LS3. Satisfied with my life	0.53***	0.061	0.19***	0.022
LS4. Have gotten the important things I wanted	1.31***	0.085	0.51***	0.032
LS5. Would change almost nothing	2.99***	0.121	0.73***	0.028
SC1. In tune with people	0.31***	0.014	0.73***	0.025
SC2. People you can talk to	0.17***	0.011	0.46***	0.027
SC3. People you can turn to	0.14***	0.010	0.39***	0.025
SC4. People who understand you	0.15***	0.008	0.39***	0.024
SC5. People you feel close to	0.13***	0.009	0.39***	0.028
SC6. Part of a group of friends	0.30***	0.015	0.59***	0.027
SC7. Have a lot in common with people around you	0.23***	0.011	0.56***	0.026

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

The measurement model for the outcomes was tested for measurement invariance across gender. Model fit statistics and chi-square difference testing results are shown in Table 17. The measurement model showed configural invariance (same pattern of items on factors) and weak metric invariance (factor loadings equal across gender) but not strong metric invariance.

Table 17

Measurement invariance results for life satisfaction and perceived social connectedness

Model	$\chi^2(df)$	χ^2/df	***CFI***	***RMSEA***	***SRMR***	***Corrected*** $\Delta\chi^2(\Delta df)$	p
Configural	170.85(98)***	1.74	0.988	0.029	0.040		
Weak metric	182.19(108)***	1.69	0.988	0.028	0.043	11.23(10)	.340
Strong metric	204.70(118)***	1.73	0.986	0.029	0.047	23.44(10)	.009

Note. * $p < .05$ ** $p < .01$ *** $p < .001$

Accounting for measurement error. In order to check the sensitivity of models to measurement error, the initial, disaggregated model (Figure 23) was fit with three levels of measurement error specified, respectively at 10%, 20% and 30%. This was achieved by fixing the residual variance of each indicator to the product of the measurement error amount (10%, 20%, or 30%) with the observed sample variance for the indicator. Because measurement error tends to attenuate regression coefficients, and participation items enter as regression predictors in the disaggregated model, coefficient estimates may increase as measurement error increases. Results are shown in Table C.1 in Appendix C. *P*-values generally increased with increased measurement error, but patterns of significant and non-significant coefficients stayed largely the same. Reading was a significant predictor in the model without measurement error specified but not in any of the models with measurement error specified. The magnitude of coefficient estimates increased in some cases but not in amounts that seemed practically significant. Overall, different levels of measurement error specified for the disaggregated model didn't seem to affect results in an important way. Given this result, it was decided to fit the index construction models with 10% measurement error following the example of Grace and Bollen (2006).

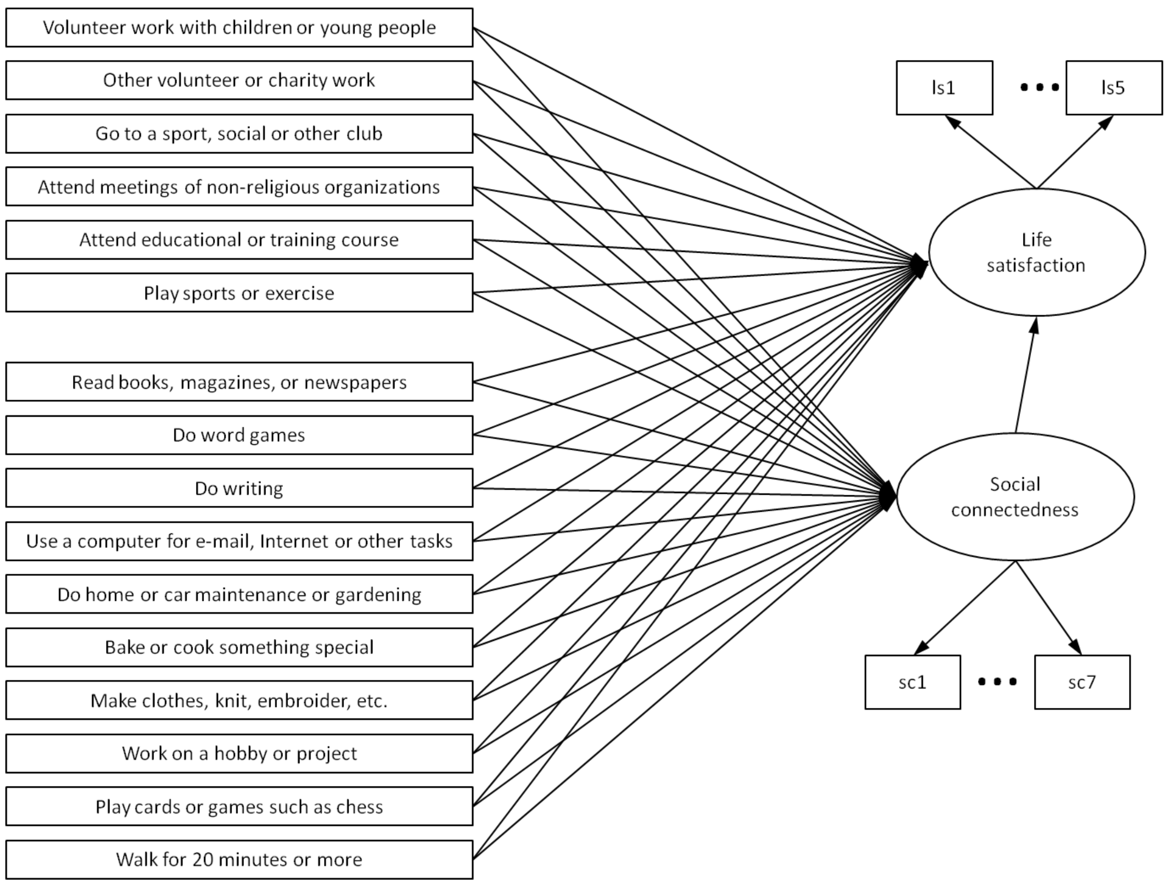


Figure 23. Disaggregated model of social participation

Fitting the disaggregated model. Before specifying any composites, the disaggregated model in which each participation item predicted both perceived social connectedness and life satisfaction was fit (Figure 23), for all respondents and then separately by gender, with 10% measurement error specified in each case. Estimated parameters for the model fit with the entire measurement construction half of the data set (n = 1791) are reported in Table 18. Coefficients significant at the $p < .10$ level are highlighted. This generous significance level was chosen for initial screening since once correlated items have their coefficients constrained to zero, p-values for the remaining coefficients were expected to decrease, due to correlations across participation items. Social connectedness significantly predicted life satisfaction ($\beta = .28$, $SE = .033$, $p < .001$). Volunteering (other than with youth) significantly predicted both life satisfaction ($\beta = .07$, SE = .030, $p = .030$) and social connectedness ($\beta = .7$, $SE = .10$, $p = .002$). Three other activities significantly predicted life satisfaction at the $p = .10$ level: sports/social club participation ($\beta = .05$, $SE = .030$, $p = .088$), computer usage ($\beta = .12$, $SE = .030$, $p < .001$) and hobbies ($\beta = .08$, $SE = .037$, $p = .023$). Two activities other than volunteering also predicted social connectedness: reading ($\beta = .07$, $SE = .035$, $p = .054$) and baking or cooking ($\beta = .13$, $SE = .032$, $p < .001$).

Table 18

Parameter estimates for disaggregated model – All respondents

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.147***	0.147	0.28***	0.033
sp1. Volunteer with youth	-0.005	0.048	0.00	0.031
sp2. Volunteer – other	0.074*	0.034	0.07*	0.030
sp3. Education	-0.059	0.070	-0.03	0.034
sp4. Sports/social club	0.056†	0.033	0.05†	0.030
sp5. Non-religious organizations	0.006	0.056	0.00	0.029
sp6. Read	0.023	0.036	0.02	0.038
sp7. Word games	-0.010	0.021	-0.02	0.031
sp8. Cards/chess/other games	-0.016	0.027	-0.02	0.030
sp9. Writing	-0.035	0.034	-0.04	0.035
sp10. Computer	0.076**	0.020	0.12**	0.030
sp11. Home maintenance/ gardening	0.041	0.026	0.06	0.034
sp12. Bake or cook	0.043	0.027	0.05	0.032
sp13. Sew or knit	-0.017	0.038	-0.01	0.031
sp14. Hobby	0.065*	0.029	0.08*	0.037
sp15. Sports/Exercise	0.022	0.026	0.03	0.036
sp16. Walk 20 minutes	0.038	0.026	0.05	0.034
Social connectedness on				
sp1. Volunteer with youth	-0.001	0.013	0.00	0.033
sp2. Volunteering – other	0.028**	0.009	0.10**	0.033
sp3. Education	-0.006	0.015	-0.01	0.030
sp4. Sports/social club	-0.005	0.010	-0.02	0.037
sp5. Non-religious organizations	0.023†	0.014	0.05†	0.030
sp6. Read	0.016†	0.008	0.07†	0.035
sp7. Word games	0.001	0.006	0.01	0.033
sp8. Cards/chess/other games	0.006	0.007	0.03	0.034
sp9. Writing	0.011	0.008	0.05	0.034
sp10. Computer	0.006	0.006	0.04	0.035
sp11. Home maintenance/ gardening	0.001	0.006	0.00	0.035
sp12. Bake or cook	0.027***	0.007	0.13***	0.032
sp13. Sew or knit	-0.004	0.009	-0.01	0.030
sp14. Hobby	0.000	0.008	0.00	0.039
sp15. Sports/Exercise	0.010	0.007	0.06	0.039
sp16. Walk 20 minutes	0.002	0.007	0.01	0.036

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

The model for males could not be estimated with item number 13 “Make clothes, knit, embroider, etc” included because of small numbers of respondents who reported any participation at all in that activity. It was dropped from the model since it had been dropped from the scale model already, so it did not need to be kept in to maintain comparability of the models developed for male respondents. Table 19 reports regression coefficients with standard errors and *p*-values, both unstandardized and standardized, for the model estimated with males only. Social connectedness significantly predicted life satisfaction ($\beta = .31$, $SE = .047$, $p < .001$). Two participatory activities predicted life satisfaction at the $p < .10$ level for male respondents: computer usage ($\beta = .15$, $SE = .04$, $p = .001$) and home maintenance, car maintenance, or gardening ($\beta = .047$, $SE = .047$, $p = .056$). Only volunteering (other than with youth) significantly predicted social connectedness for males ($\beta = .11$, $SE = .046$, $p = .015$).

Table 19

Parameter estimates for disaggregated model – Male

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.213***	0.207	0.31***	0.047
sp1. Volunteer with youth	0.019	0.064	0.01	0.033
sp2. Volunteer – other	0.040	0.051	0.03	0.044
sp3. Education	-0.094	0.094	-0.04	0.043
sp4. Sports/social club	0.040	0.049	0.04	0.047
sp5. Non-religious organizations	0.011	0.075	0.01	0.040
sp6. Read	0.065	0.051	0.07	0.056
sp7. Word games	-0.050	0.031	-0.07	0.040
sp8. Cards/chess/other games	-0.012	0.039	-0.01	0.041
sp9. Writing	0.001	0.049	0.00	0.047
sp10. Computer	0.095**	0.028	0.15**	0.042
sp11. Home maintenance/ gardening	0.071†	0.037	0.09†	0.047
sp12. Bake or cook	0.035	0.041	0.04	0.046
sp14. Hobby	0.056	0.043	0.07	0.055
sp15. Sports/Exercise	0.011	0.040	0.02	0.055
sp16. Walk 20 minutes	0.033	0.038	0.04	0.052
Social connectedness on				
sp1. Volunteer with youth	0.011	0.022	0.02	0.044
sp2. Volunteer – other	0.034**	0.014	0.11**	0.046
sp3. Education	0.029	0.020	0.05	0.035
sp4. Sports/social club	-0.018	0.014	-0.07	0.054
sp5. Non-religious organizations	0.000	0.021	0.00	0.044
sp6. Read	0.017	0.012	0.08	0.051
sp7. Word games	0.003	0.009	0.01	0.048
sp8. Cards/chess/other games	0.005	0.011	0.02	0.047
sp9. Writing	0.019	0.015	0.07	0.054
sp10. Computer	0.009	0.008	0.06	0.050
sp11. Home maintenance/ gardening	0.019	0.011	0.09	0.057
sp12. Bake or cook	0.007	0.010	0.03	0.045
sp13. Sew or knit				
sp14. Hobby	0.007	0.011	0.04	0.056
sp15. Sports/Exercise	0.014	0.011	0.08	0.059
sp16. Walk 20 minutes	-0.003	0.010	-0.02	0.055

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

Table 20 reports regression coefficients, both unstandardized and standardized, with standard errors and *p*-values for the model estimated with female respondents only. As with men, social connectedness significantly predicted life satisfaction in women (β = .27, *SE* = .046, $p < .001$). Four activities significantly predicted life satisfaction at the $p < .10$ level: volunteering other than with youth (β = .09, *SE* = .043, p = .046), sports/social club participation (β = .06, *SE* = .036, p = .081), computer usage (β = .09, *SE* = .043, p = .032), and baking or cooking (β = .10, *SE* = .045, p = .032). Three activities significantly predicted social connectedness: volunteering other than with youth (β = .09, *SE* = .045, p = .045), non-religious organization participation (β = .09, *SE* = .038, p = .024), and baking or cooking (β = .16, *SE* = .045, p = .001).

Table 20

Parameter estimates for disaggregated model – Female

	Unstandardized		Standardized	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.120***	0.213	0.27***	0.046
sp1. Volunteer with youth	-0.009	0.060	-0.01	0.045
sp2. Volunteer – other	0.096*	0.048	0.09*	0.043
sp3. Education	-0.016	0.099	-0.01	0.050
sp4. Sports/social club	0.076†	0.044	0.06†	0.036
sp5. Non-religious organizations	-0.040	0.084	-0.02	0.042
sp6. Read	-0.023	0.053	-0.02	0.052
sp7. Word games	0.029	0.029	0.04	0.043
sp8. Cards/chess/other games	-0.018	0.037	-0.02	0.043
sp9. Writing	-0.041	0.045	-0.05	0.050
sp10. Computer	0.060*	0.028	0.09*	0.043
sp11. Home maintenance/ gardening	0.007	0.035	0.01	0.047
sp12. Bake or cook	0.082*	0.038	0.10*	0.045
sp13. Sew or knit	0.017	0.044	0.02	0.044
sp14. Hobby	0.045	0.041	0.06	0.053
sp15. Sports/Exercise	0.024	0.035	0.03	0.046
sp16. Walk 20 minutes	0.030	0.034	0.04	0.045
Social connectedness on				
sp1. Volunteer with youth	-0.007	0.015	-0.02	0.047
sp2. Volunteer – other	0.024*	0.012	0.09*	0.045
sp3. Education	-0.027	0.021	-0.06	0.045
sp4. Sports/social club	0.017	0.014	0.06	0.047
sp5. Non-religious organizations	0.041*	0.018	0.09*	0.038
sp6. Read	0.015	0.012	0.06	0.047
sp7. Word games	-0.007	0.007	-0.04	0.045
sp8. Cards/chess/other games	0.009	0.010	0.04	0.047
sp9. Writing	0.002	0.010	0.01	0.046
sp10. Computer	0.002	0.008	0.02	0.049
sp11. Home maintenance/ gardening	-0.009	0.008	-0.05	0.044
sp12. Bake or cook	0.031**	0.009	0.16**	0.045
sp13. Sew or knit	-0.005	0.011	-0.02	0.048
sp14. Hobby	-0.007	0.011	-0.04	0.062
sp15. Sports/Exercise	0.008	0.009	0.04	0.050
sp16. Walk 20 minutes	0.008	0.008	0.04	0.045

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

Comparing the disaggregated and composite models. The disaggregated models for the entire sample, for males only, and for females only were compared to the two-composite hypothesized model (Figure 24), specifying 10% measurement error in the participation items. Adding composites requires both the addition of parameters (from participation indicators to the composite) and deletion of parameters (directly from the participation indicators to the outcomes) so that composite constraints do not result in nested models relative to models without the composites. Therefore, the model with composites was compared with the disaggregated model using BIC values. BIC was chosen because it penalizes complexity more than does the AIC (Kline, 2005). Fit statistics and BIC values are shown in Table 21. The two-composite hypothesized model (Figure 24) had lower BIC values in each case (all respondents, males only, females only) as shown in Table 21 below. The fit of each of the two-composite models was good, with RMSEA < .05, CFI > .95, and SRMR < .05 in each case. Given the good fit of the two-composite models, one-composite models were fit. The one-composite model (Figure 25) joined all participation indicators in a single composite that predicted social connectedness and life satisfaction, with social connectedness mediating the relationship between the participation composite and life satisfaction. The one composite model had lower BIC values in each of the three cases, suggesting that this was a better model than the two-composite model in an overall tradeoff between fit and parsimony. Each of the one-composite models had good fit considering approximate fit indexes but each had significant chi-squares at the $p < .001$ level. In order to check sensitivity of results to the amount of measurement error specified, the one-composite model across the entire sample was estimated with 10%, 20%, and 30% measurement error specified for each participation item. The three models identified the same participation items as having significant predictive power in the model. Regression coefficients and standard errors were generally similar across models, suggesting that it would acceptable to consider only models with 10% measurement error specified, as was done in the remainder of the index construction process.

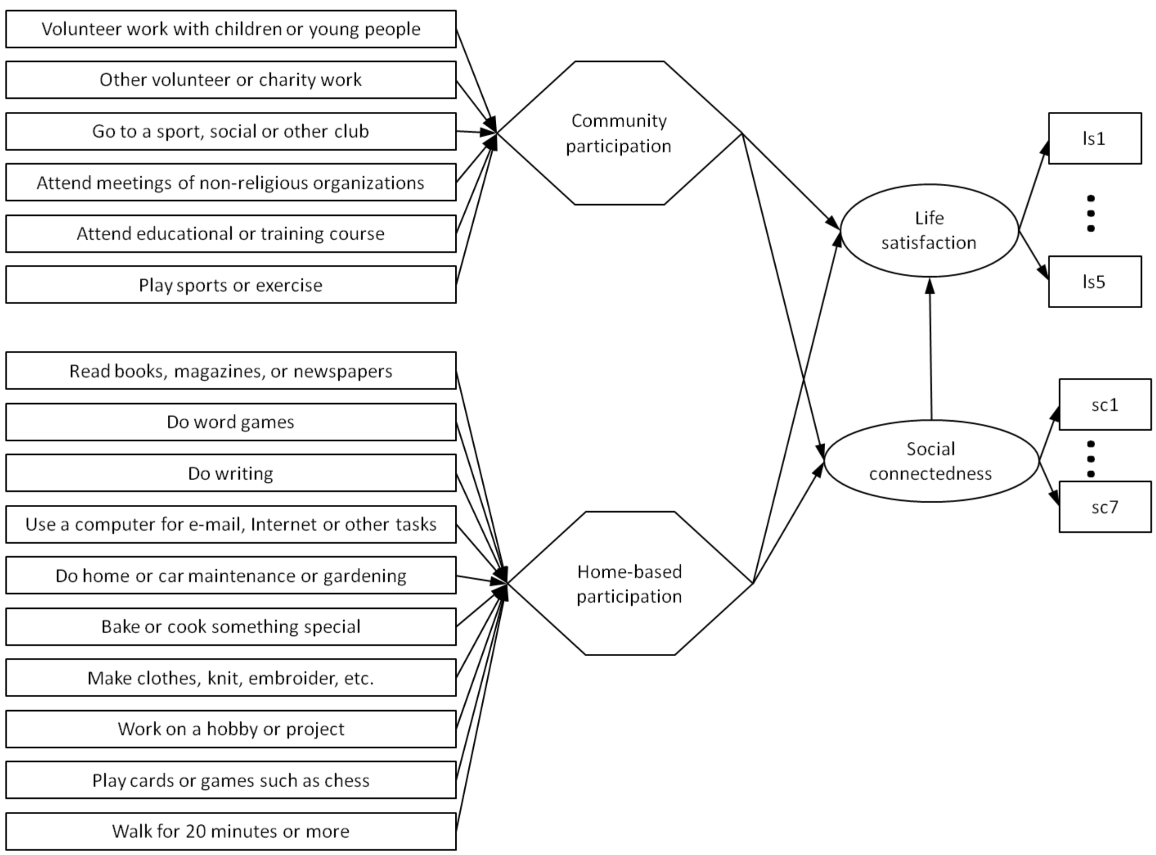


Figure 24. Hypothesized two-composite model of social participation

Table 21

Index construction model comparisons – Disaggregated, two composite, and one composite

Sample	*Model*	*BIC*	$\chi^2(df)$	χ^2/df	*RMSEA*	*CFI*	*SRMR*
All	Disaggregated	149570.46	486.75(209)***	2.33	0.027	0.98	0.030
	Two composites	149491.05	506.50(223)***	2.27	0.027	0.97	0.031
	One composite	149484.46	506.08(224)***	2.26	0.027	0.98	0.031
Male	Disaggregated	63865.98	320.97(226)***	1.42	0.023	0.98	0.033
	Two composites	63786.67	328.55(239)***	1.37	0.022	0.98	0.033
	One composite	63784.37	331.51(240)***	1.38	0.022	0.98	0.034
Female	Disaggregated	84459.74	424.41(209)***	2.03	0.032	0.97	0.034
	Two composites	84378.73	434.50(223)***	1.95	0.031	0.97	0.035
	One composite	84372.03	433.47(224)***	1.94	0.031	0.97	0.035

*** $p < .001$

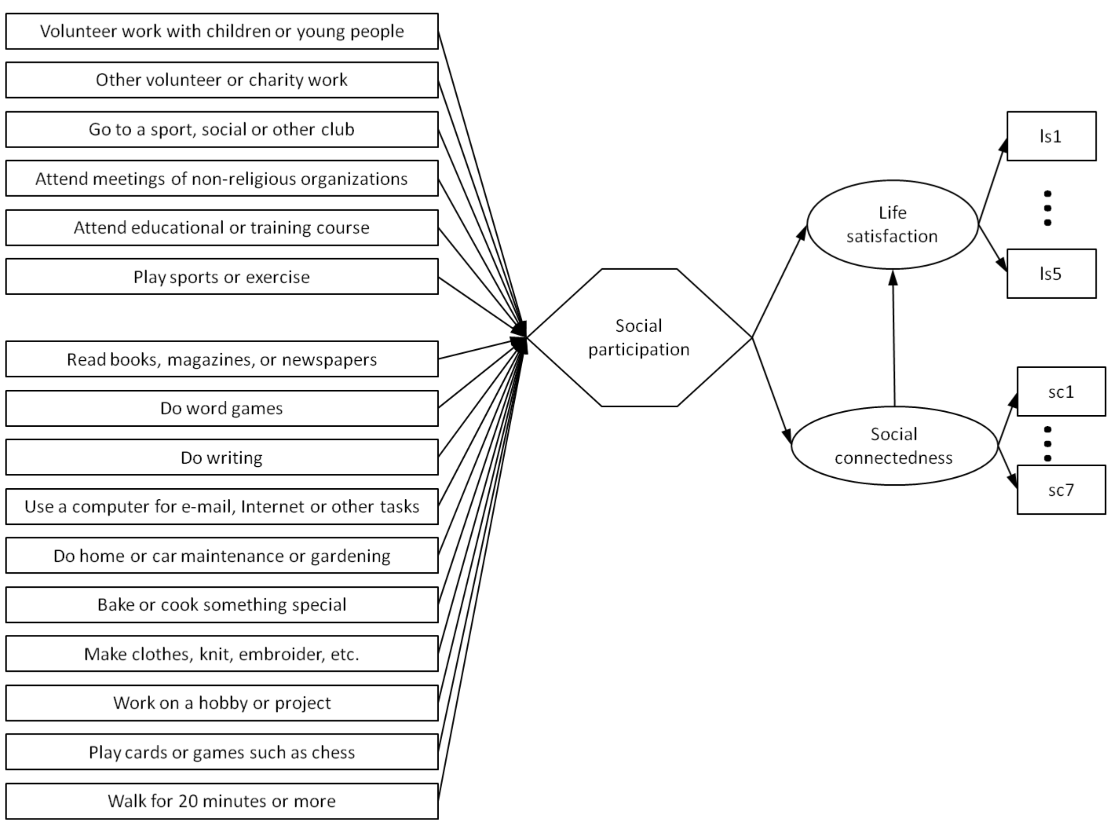


Figure 25. One composite model of social participation

Refining the model. Table 22 shows the estimated regression weights for the one-composite model fit across the entire sample, using 10% residual variance for the social participation items. Social connectedness significantly predicted life satisfaction ($\beta = .27$, $SE = .033$, $p < .001$), as did the participation composite ($\beta = .26$, $SE = .033$, $p < .001$). Participation also significantly predicted social connectedness ($\beta = .26$, SE = .032, $p < .001$). Items that predicted the participation composite at a significance level of $p < .10$ are indicated in the table. This significance level was chosen because items of marginal significance may become significant at the $p < .05$ level when other predictors' coefficients are constrained to zero, given the moderate to large correlations between some participation items. The five items that predicted the composite at the $p < .10$ level were volunteering (other than with youth), reading, computer usage, baking or cooking, and hobbies. Note that standardized and unstandardized coefficients have different sampling distributions so the standardized coefficient may be significant when the unstandardized is not and vice versa (Muthén & Muthén, 2010).

Table 22

Regression weights for one-composite model – All respondents

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.107***	0.147	0.273***	0.033
Participation	0.093***	0.025	0.257***	0.033
Social connectedness on				
Participation	0.023***	0.007	0.261***	0.032
Participation on				
sp1. Volunteer with youth	-0.046	0.365	-0.011	0.086
sp2. Volunteer – other	1.000	0.000	0.316***	0.083
sp3. Education	-0.454	0.418	-0.079	0.077
sp4. Sports/social club	0.221	0.275	0.072	0.086
sp5. Non-religious organizations	0.516	0.466	0.096	0.080
sp6. Read	0.463†	0.305	0.175†	0.098
sp7. Word games	-0.036	0.160	-0.019	0.083
sp8. Cards/chess/other games	0.036	0.215	0.015	0.086
sp9. Writing	0.019	0.241	0.007	0.091
sp10. Computer	0.561***	0.230	0.309***	0.085
sp11. Home maintenance/ gardening	0.250	0.206	0.120	0.091
sp12. Bake or cook	0.806**	0.284	0.352***	0.083
sp13. Sew or knit	-0.173	0.280	-0.050	0.080
sp14. Hobby	0.375	0.251	0.174†	0.097
sp15. Sports/Exercise	0.340	0.226	0.166†	0.105
sp16. Walk 20 minutes	0.249	0.204	0.121	0.093

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

Regression coefficients for the model estimated with male respondents only are shown in Table 23. As in the overall model, social connectedness and participation both significantly predicted life satisfaction and participation predicted social connectedness. The activities that predicted participation at the $p < .10$ level for male respondents were volunteering – other, reading, computer usage, and home maintenance/gardening.

Table 23

Regression weights for one-composite model – Male

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.173***	0.206	0.297***	0.047
Participation	0.083***	0.038	0.301***	0.047
Social connectedness on				
Participation	0.022***	0.011	0.316***	0.044
Participation on				
sp1. Volunteer with youth	0.370	0.666	0.052	0.082
sp2. Volunteer – other	1.000	0.000	0.236*	0.106
sp3. Education	0.073	0.810	0.009	0.100
sp4. Sports/social club	-0.168	0.412	-0.045	0.111
sp5. Non-religious organizations	0.081	0.660	0.012	0.096
sp6. Read	0.783	0.543	0.241*	0.117
sp7. Word games	-0.249	0.306	-0.090	0.097
sp8. Cards/chess/other games	0.037	0.353	0.011	0.103
sp9. Writing	0.442	0.505	0.116	0.113
sp10. Computer	0.793†	0.452	0.338**	0.104
sp11. Home maintenance/ gardening	0.853†	0.512	0.301**	0.109
sp12. Bake or cook	0.366	0.351	0.113	0.097
sp14. Hobby	0.496	0.436	0.176	0.120
sp15. Sports/Exercise	0.395	0.390	0.148	0.133
sp16. Walk 20 minutes	0.132	0.320	0.049	0.119

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

Regression coefficients for the model estimated with female respondents only are shown in Table 24. As in the overall model, social connectedness and participation both significantly predicted life satisfaction and participation predicted social connectedness. The activities that predicted participation at the $p < .10$ level for female respondents were volunteering – other, sport/social club participation, computer use, and baking/cooking.

Table 24

Regression weights for one-composite model – Female

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.071***	0.217	0.254***	0.048
Participation	0.097***	0.035	0.234***	0.046
Social connectedness on				
Participation	0.024***	0.009	0.244***	0.041
Participation on				
sp1. Volunteer with youth	-0.185	0.429	-0.057	0.137
sp2. Charity work	1.000	0.000	0.371**	0.123
sp3. Education	-0.637	0.514	-0.133	0.119
sp4. Sports/social club	0.751†	0.421	0.259*	0.115
sp5. Non-religious organizations	0.624	0.643	0.130	0.121
sp6. Read	0.187	0.380	0.076	0.148
sp7. Word games	0.009	0.217	0.006	0.133
sp8. Cards/chess/other games	0.092	0.275	0.044	0.130
sp9. Writing	-0.176	0.300	-0.080	0.132
sp10. Computer	0.363	0.260	0.229	0.132
sp11. Home maintenance/ gardening	-0.145	0.251	-0.081	0.139
sp12. Bake or cook	1.079*	0.443	0.526***	0.120
sp13. Sew or knit	-0.017	0.308	-0.007	0.129
sp14. Hobby	0.094	0.297	0.050	0.154
sp15. Sports/Exercise	0.289	0.274	0.160	0.150
sp16. Walk 20 minutes	0.318	0.270	0.175	0.130

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

A model was fit to all respondents in the measure construction subsample that constrained regression weights to zero for participation variables that did not show up as significant at the $p < .10$ level in any of the three models (entire sample, males only, females only). This model had good fit, $\chi^2(233) = 517.79$, $p < .0001$, $\chi^2/df = 2.22$, RMSEA = .026, CFI = .974, SRMR = .033. It was not significantly worse than the unconstrained model, corrected $\Delta\chi^2(9) = 11.17$, $p = .26$. Coefficients are shown in Table 25. This model expresses a one-index measure of social participation, consisting of responses on charity work, sports/social club participation, reading, computer usage, home maintenance/gardening, baking/cooking, and hobbies. Standardized coefficients were each significant at the $p < .05$ level except for home maintenance/gardening (β = .175, *SE* = .091, $p = .056$). Note that remaining participation items were not dropped in estimating the model. They had their coefficients set to zero in the composite but they were still modeled latently with residual variance specified and were allowed to covary with each other. A model that dropped the trimmed participation items entirely was also run and it showed similar fit statistics and coefficient estimates, which suggested that either way would produce similar results.

Table 25

Regression weights for refined index model of social participation – All respondents

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Std. Coef.***	***SE***
Life satisfaction on				
Social connectedness	1.122***	0.145	0.276***	0.033
Participation	0.096***	0.024	0.250***	0.033
Social connectedness on				
Participation	0.024***	0.007	0.253***	0.032
Participation on				
sp2. Charity work	1.000	0.000	0.337***	0.077
sp4. Sports/social club	0.470†	0.261	0.163*	0.080
sp6. Read	0.492†	0.280	0.199*	0.096
sp10. Computer	0.572**	0.210	0.337***	0.082
sp11. Home maintenance/ gardening	0.343†	0.201	0.175†	0.091
sp12. Bake or cook	0.769**	0.253	0.359***	0.082
sp14. Hobbies	0.408†	0.221	0.203*	0.088

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

Evaluating invariance across gender. Configural invariance was tested by estimating a multiple group model for males and females with all parameters left free to vary by group. The model showed good fit, $\chi^2(496) = 1057.043$, $p < .0001$, $\chi^2/\text{df}=2.1$, RMSEA = .036, CFI = .95, SRMR = .044. However, patterns of significant coefficients and sign of coefficients differed by gender (Table 26), so it appeared that the index model was not, in fact, configurally invariant (a concept which was developed for factor models, not regression models such as this one). For men, charity work, reading, using a computer, doing home maintenance/car maintenance/gardening, and hobbies were statistically significant, $p < .05$. For women, charity work, going to a sport/social club, and baking/cooking were statistically significant, $p < .05$. All the remaining coefficients were in the expected direction (positive) with the exception of doing home or car maintenance or gardening for women, which was negative, but not statistically significant (β = -.04, *SE* = .15, *p* = .78). This item was dropped for the index model for females. The nonsignificant baking/cooking item was dropped from the male model, since this had been dropped from the scale model so this would allow the final models to be comparable in terms of which items were incorporated into the analysis. The remaining nonsignificant items were kept for both male and female models, since even if they weren't significant, they added to the predictive power of the index. Gelman and Hill (2007) suggested keeping regression coefficients in a model to the extent that they are substantively justified and show the expected sign.

Table 26

Standardized coefficients for one composite, seven item model fit to male and female subsamples

	Male		***Female***	
	Coef.	***SE***	***Coef.***	***SE***
Volunteering	0.30**	0.10	0.38**	0.11
Sports/social clubs	0.01	0.10	0.38***	0.11
Reading	0.26*	0.11	0.08	0.14
Using computer	0.39***	0.09	0.23†	0.13
Home/car maintenance or gardening	0.31**	0.11	-0.04	0.15
Baking/cooking	0.15	0.10	0.56***	0.12
Hobby/project	0.23*	0.12	0.10	0.13

Note. † *p* < .10. * *p* < .05. ** *p* < .01. ** *p* < .001.

Validating the index model. The male and female index models were fit in the validation subsample. The male model had one participation composite comprising six items: volunteering, sports/social clubs, reading, using a computer, doing home or car maintenance or gardening, and working on a hobby or project. The female model had one composite comprising six items: all the items in the male model except baking or cooking something special instead in the place of doing home or car maintenance or gardening. Items not appearing in the index were not included in the analysis. Measurement error of 10% of observed sample variance was specified for each indicator. Participation items were allowed to covary. The male model showed good fit, $\chi^2(114) = 194.23$, $p < .0001$, χ^2/df=1.7, RMSEA = .028, CFI = .98, SRMR = .031, but the coefficient for reading was negative, though nonsignificant (Table 27). Only two standardized coefficients were statistically significant: volunteering ($\beta = .50$, *SE* = .21, $p = .016$) and sports/social clubs ($\beta = .42$, *SE* = .19, $p = .026$). No unstandardized coefficients were statistically significant.

Table 27

Coefficient estimates for participation index – Male – Validation subsample

	Unstandardized			***Standardized***		
	Coef.	***SE***	***p***	***Coef.***	***SE***	***p***
Volunteering	1.00	0.00	999	0.50*	0.21	0.016
Sports/social clubs	0.76	0.53	0.153	0.42*	0.19	0.026
Reading	-0.13	0.38	0.726	-0.07	0.19	0.717
Using computer	0.24	0.29	0.419	0.21	0.21	0.312
Home/car maintenance or gardening	0.35	0.42	0.403	0.24	0.23	0.304
Hobby/project	0.45	0.38	0.241	0.32	0.21	0.131

Note. † *p* < .10. * *p* < .05. ** *p* < .01. ** *p* < .001.

The female model also showed good fit, $\chi^2(114) = 243.87$, $p < .0001$, χ^2/df=2.1, RMSEA = .030, CFI = .97, SRMR = .038. All coefficients on the participation composite were in the expected positive direction (Table 28). Three of the six items were statistically significant: volunteering ($\beta = .38$, *SE* = .12, $p = .002$), using the computer ($\beta = .33$, *SE* = .13, $p = .009$), and doing a hobby or project ($\beta = .34$, *SE* = .12, $p = .006$).

Table 28

Coefficient estimates for participation index – Female – Validation subsample

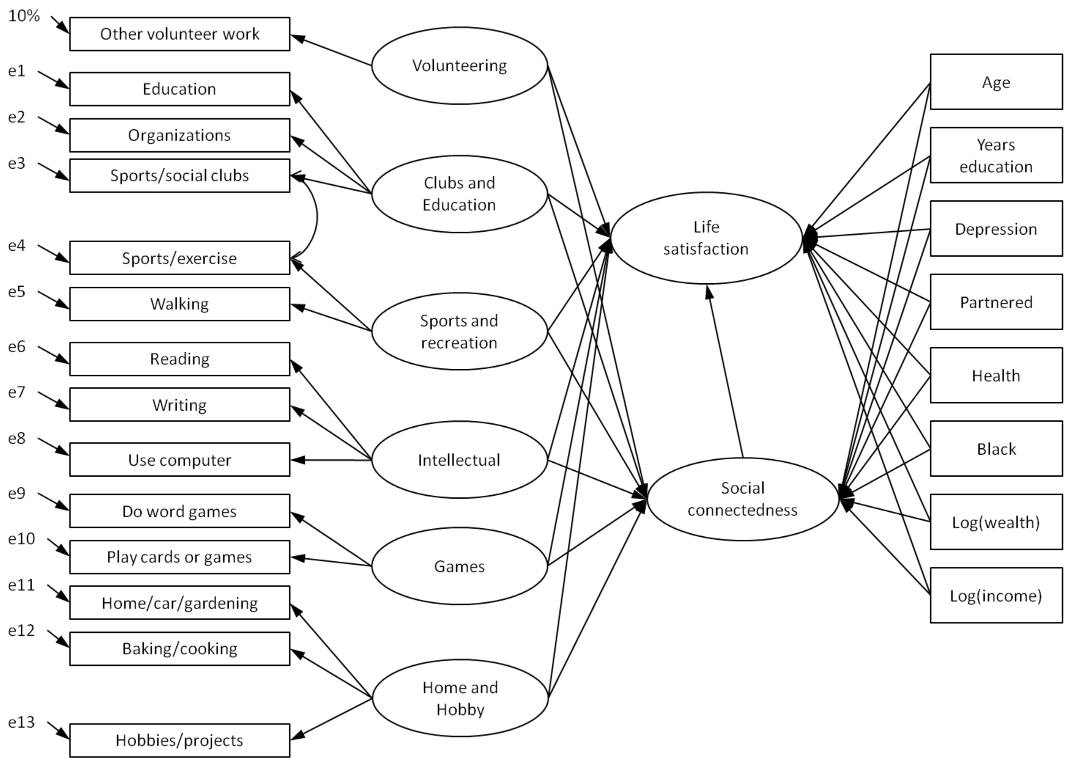
	Unstandardized			***Standardized***		
	Coef.	***SE***	***p***	***Coef.***	***SE***	***p***
Volunteering	1.00	0.00	999	0.38**	0.12	0.002
Sports/social clubs	0.27	0.36	0.452	0.11	0.14	0.426
Reading	0.48	0.39	0.214	0.19	0.13	0.142
Using computer	0.50†	0.27	0.069	0.33**	0.13	0.009
Baking/cooking	0.68†	0.39	0.084	0.34*	0.14	0.015
Hobby/project	0.60	0.32	0.064	0.34*	0.12	0.006

Note. † $p < .10$. * $p < .05$. ** $p < .01$. ** $p < .001$.

Nonsignificant coefficients in the validation sample do not necessarily represent a problem so long as the coefficients are in the expected direction. The fact that the coefficient on reading was negative in the male subsample could indicate a problem with the index definition. The index models were left as is, however, since the validation sample was what would be used in comparing the scale and index models. Adjusting the index model based on these results would give the index model an unfair advantage, since the scale model was not adjusted based on results from validation. Thus the index models validated here were the ones used in comparing results with the scale model despite the potential problems stemming from including the reading item.

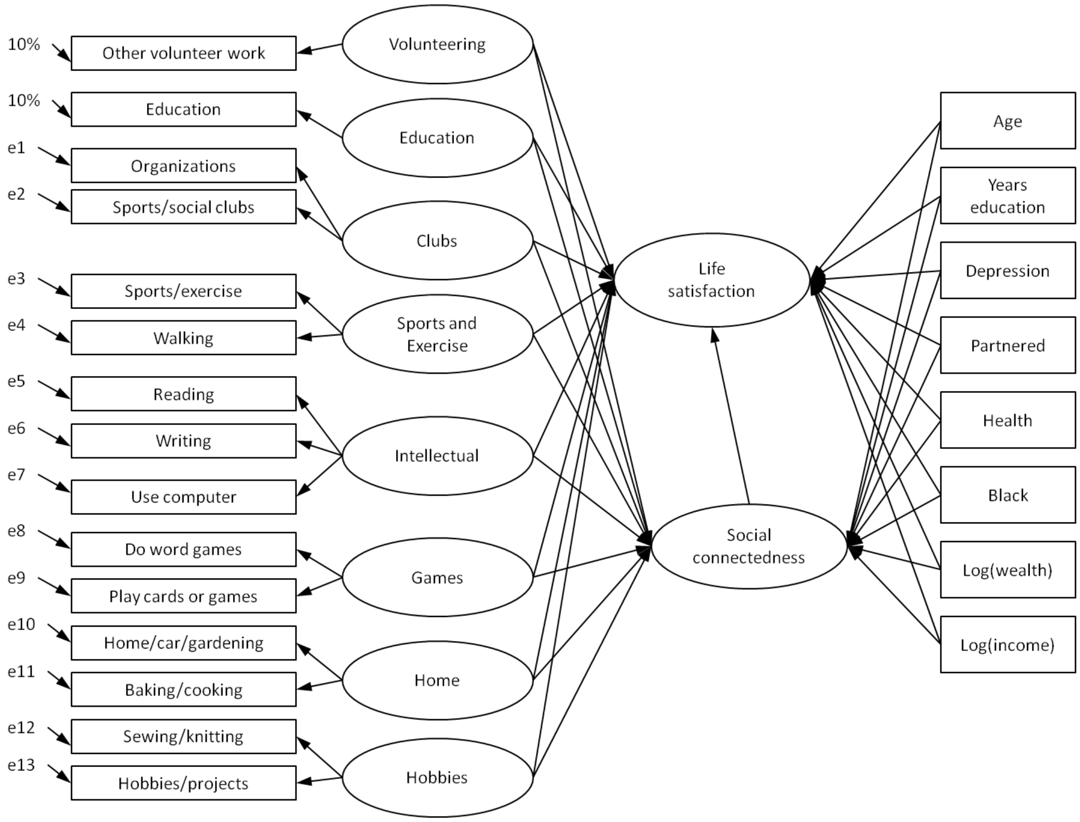
Comparison of Scale and Index Models

Once the scale and index models of participation were developed, they were fitted using the confirmatory half of the data set in the full model including both outcomes and covariates. The scale model was not configurally invariant for male and female respondents and the index model did not show the same patterns of coefficient significance by gender, so the models were compared by gender. Figure 26 and Figure 27 show the structural model using scales for males and females, respectively. Figure 28 shows the full structural model using the alternate four-factor participation scale model for females. Factor covariances are not drawn for clarity, but all factors were allowed to covary (including single-item factors). Figure 29 and Figure 30 show the structural model using indexes for males and females, respectively. Participation indicators in the index models had 10% residual variance specified. Single-item factors in the scale models also had 10% residual variance specified, to make the models as comparable as possible. Because the scale and index models were not nested, they were compared by means of information criteria and approximate fit statistics rather than chi-square difference tests. Comparison using information criteria such as the Bayesian Information Criterion (BIC) requires that models be estimated with the same cases and same variables. To ensure comparability across the scale and index models, the same sets of participation variables were used in the scale and index models for males and females, respectively. For the male model, all participation variables except volunteering with youth, baking/cooking, and sewing/knitting were used. For the female model, all participation variables except volunteering with youth were used. In the index model, participation variables that were not part of the final participation index had their regression weights for the participation composite set to zero.



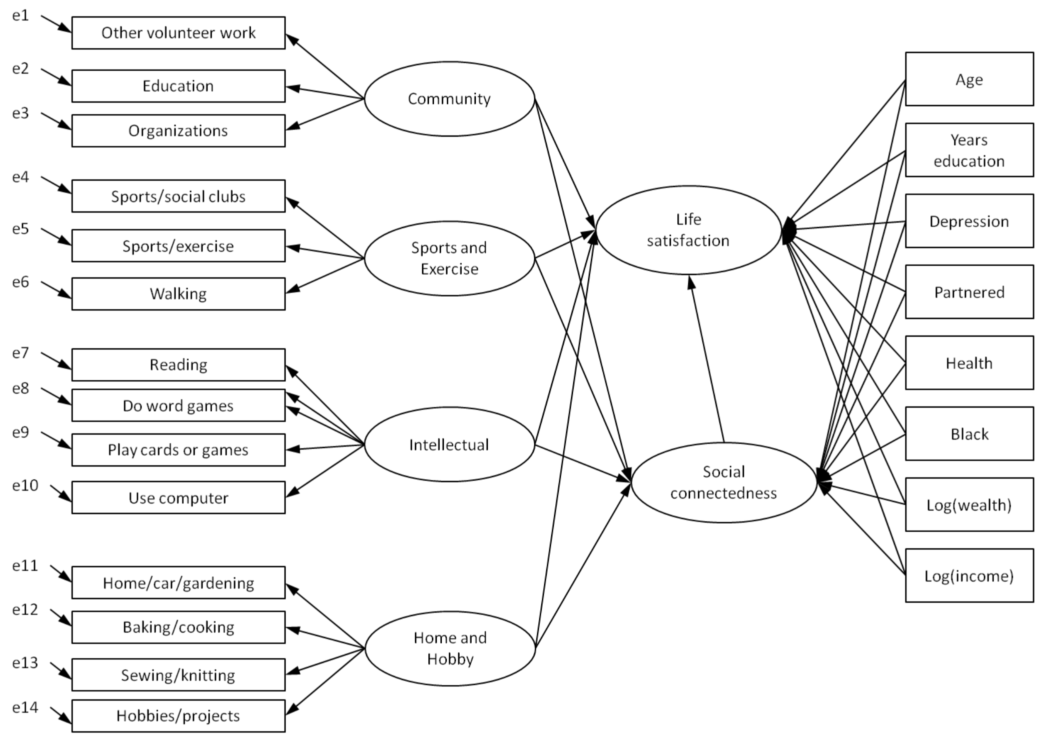
Note. All latent participation factors specified to covary with each other (not shown in figure). Single-item factors have residual error variance fixed at 10% of sample variance for the observed indicator.

Figure 26. Full structural model using participation scales, male version



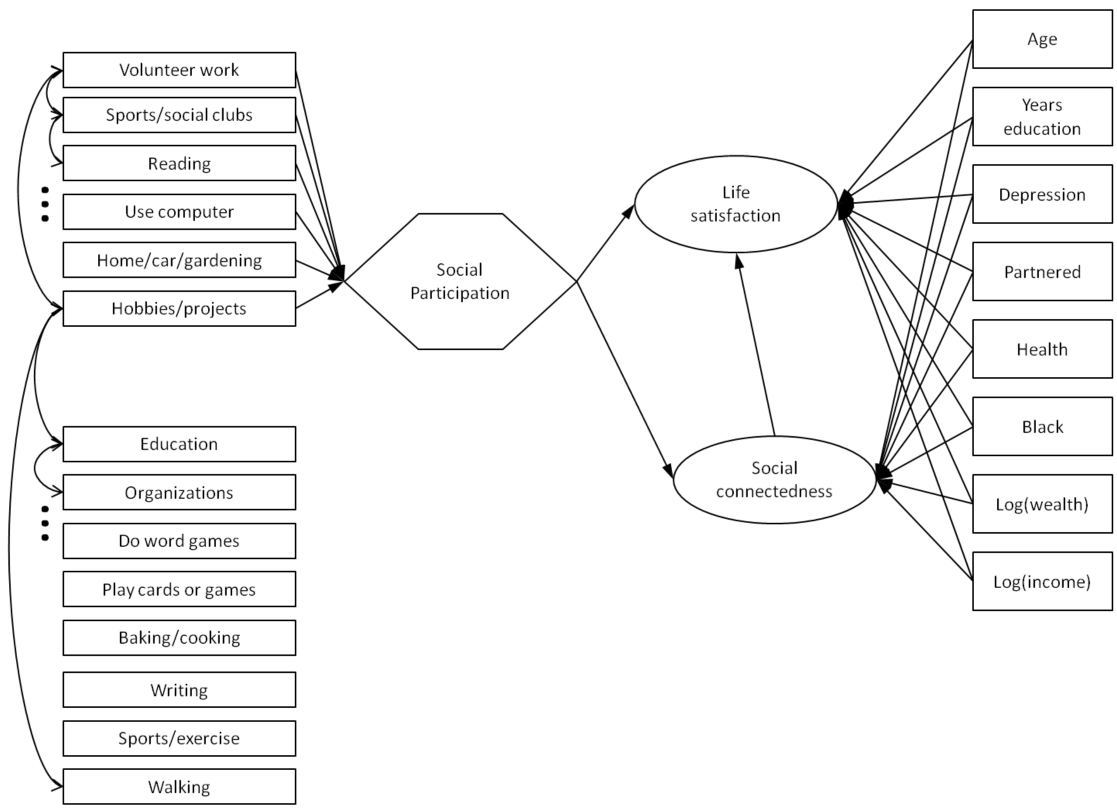
Note. All latent participation factors specified to covary (not shown in figure). Single-item factors have residual error variance fixed at 10% of sample variance for the observed indicator.

Figure 27. Full structural model using participation scales, female version



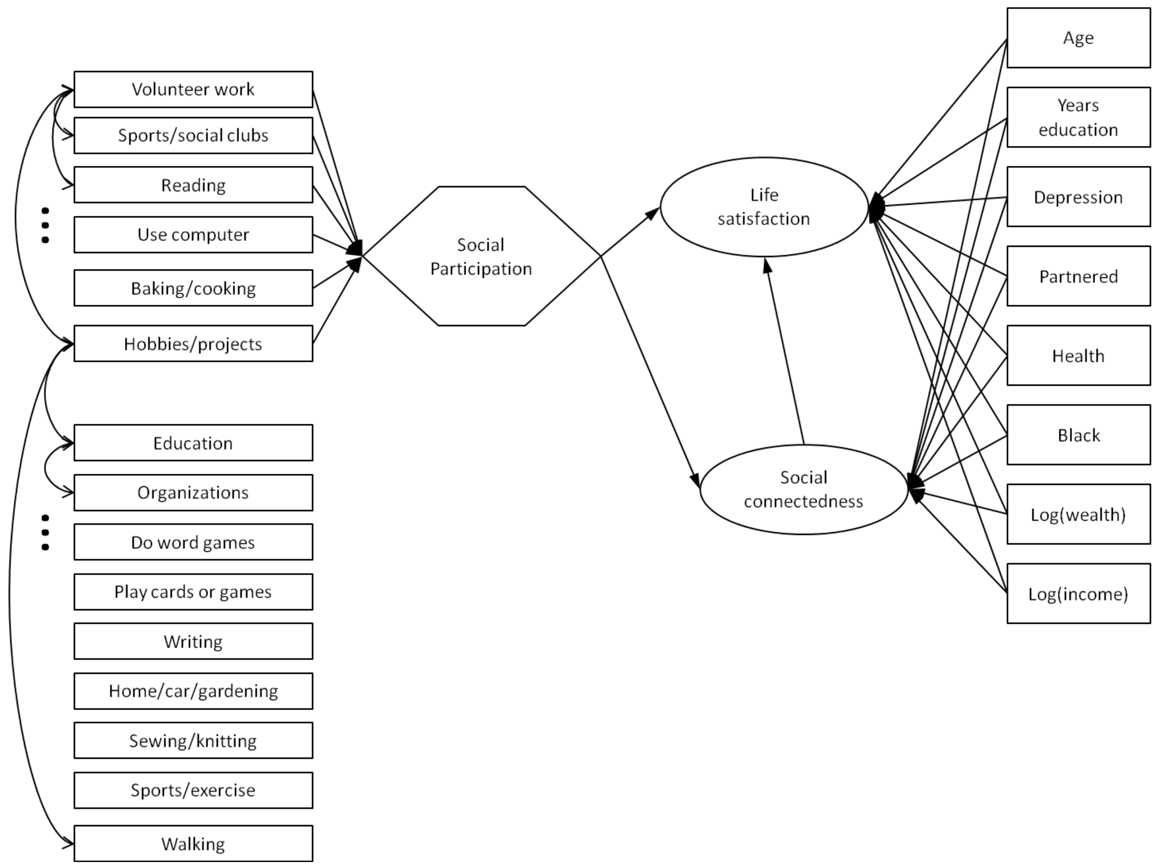
Note. All latent participation factors specified to covary (not shown in figure).

Figure 28. Full structural model using participation scales, female version – alternate four-factor model



Note. All participation items modeled latently with 10% residual error and specified to covary with each other.

Figure 29. Full structural model using participation index, male version



Note. All participation items modeled latently with 10% residual error and specified to covary with each other.

Figure 30. Full structural model using participation index, female version

Fit statistics, BIC values, and R^2 values for the outcomes are shown in Table 29. For males, the index model had a lower (better) BIC while for females, the eight-factor scale model had a lower (better) BIC. The index model for females had a lower BIC value than either the four-factor or single-factor female scale models. Conceptually, the index model seems simpler than the eight-factor scale model for females since it uses just one composite to model participation. The BIC values are measuring how well the model captures covariances across the entire data set, including participation variables not specified as part of the index composite. Perhaps it is therefore not surprising that the BIC values identified the scale model for females as superior, since it accounted for correlations among participation variables using latent factors while the index model accounted for them with pairwise correlations.

Table 29

Model comparison of scale and index models

									R^2 values	
	Model	***n***	***BIC***	χ^2	χ^2/df	***RMSEA***	***CFI***	***SRMR***	***LS***[a]	***SC***[b]
Male	Scale	868	62,247	1185.45(428)***	2.77	0.045	0.86	0.070	0.33	0.08
	Index	868	62,237	920.58(382)***	2.41	0.040	0.90	0.064	0.30	0.07
Female	Scale – eight factors	1279	98,378	1462.72(476)***	3.07	0.040	0.88	0.070	0.31	0.13
	Scale – four factors	1279	98,736	1903.43(506)***	3.76	0.046	0.83	0.079	0.29	0.12
	Scale – single factor	1279	98,832	2025.09(517)***	3.92	0.048	0.81	0.076	0.29	0.11
	Index	1279	98,569	1313.62(422)***	3.11	0.041	0.89	0.069	0.30	0.12

Note. † *p* < .10. * *p* < .05. ** *p* < .01. ** *p* < .001.

Approximate fit statistics for the scale and index models favored the index models but the six and eight-factor scale models explained more variance in the outcomes. The model incorporating scale measures of participation for males did not have adequate fit according to its CFI value, $\chi^2(428) = 1185.45$, $\chi^2/df = 2.77$, RMSEA = .045, CFI = .86, SRMR = .070. The model using the participation index for males had adequate fit, $\chi^2(382) = 920.58$, $\chi^2/df = 2.41$, RMSEA = .040, CFI = .90, SRMR = .064. The scale model for men explained more variance in life satisfaction ($R^2 = .33$) and in social connectedness ($R^2 = .08$) than the index model explained ($R^2 = .30$ for life satisfaction and $R^2 = .07$ for social connectedness). For comparison purposes, a model with only the covariates predicting the outcomes was run. This model explained 31% of the variation in life satisfaction and about five percent of the variation in social connectedness. Thus the index model actually explained less variation in life satisfaction than a model with only covariates explained.

The model incorporating the scale model of participation for females did not have adequate fit, $\chi^2(476) = 1462.72$, $\chi^2/df = 3.07$, RMSEA = .04, CFI = .88, SRMR = .070. The model incorporating the participation index for females did not have adequate fit either, $\chi^2(422) = 1313.62$, $\chi^2/df = 3.11$, RMSEA = .041, CFI = .89, SRMR = .069. As with the male model, the scale model for women explained more variance in life satisfaction ($R^2 = .31$) and in social connectedness ($R^2 = .13$) than the index model explained ($R^2 = .30$ for life satisfaction and $R^2 = .12$ for social connectedness). A model predicting the outcomes with only covariates explained 30% of variation in life satisfaction and about 12% in social connectedness, so it appeared that the scale model added little explanatory power and the index model almost none. The four-factor and single-factor alternate female models were poor overall, with worse fit values, higher BICs, and lower variance explained in life satisfaction compared to the eight-factor scale model and single-composite index models for female. However, the single-factor scale model could be improved, which is discussed in the last section of this chapter.

Scale model for men. Estimated structural coefficients for the model estimated on the male sample are reported in Table 30. Life satisfaction was significantly predicted by social connectedness ($\beta = .27$, $SE = .049$, $p < .001$) but not by any of the social participation scales. Social connectedness was significantly predicted by volunteering ($\beta = .14$, $SE = .047$, $p = .003$). Among the covariates, depression, self-report health status, and log-transformed wealth statistically significantly predicted life satisfaction while depression significantly predicted social connectedness. This model is consistent with the hypothesis that the effect of volunteering participation on life satisfaction in males may be fully mediated by volunteering's effects on social connectedness. Using Mplus' MODEL INDIRECT bootstrapping method for calculating indirect and total effects with standard errors showed a significant indirect effect of volunteering on life satisfaction via social connectedness (β = .037, SE = .014, p = .007) but a nonsignificant total effect of volunteering on life satisfaction (β = .018, SE = .048, p = .71) due to a negative (though nonsignificant) direct effect of volunteering on life satisfaction (β = -.019, SE = .047, p = .682).

Table 30

Structural coefficients – Scale model of participation – Male

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	0.98***	0.186	0.27***	0.049
Volunteering	-0.02	0.047	-0.02	0.047
Social activities	0.10	0.074	0.06	0.046
Games	-0.05	0.066	-0.06	0.067
Intellectual	-0.37	0.308	-0.14	0.112
Home and Hobbies	0.13	0.089	0.10	0.067
Sports/Exercise	0.01	0.049	0.01	0.063
Social connectedness on				
Volunteering	0.04**	0.013	0.14*	0.047
Social activities	0.01	0.020	0.03	0.045
Games	0.02	0.027	0.08	0.081
Intellectual	-0.11	0.090	-0.15	0.122
Home and Hobbies	0.03	0.028	0.09	0.076
Sports/Exercise	0.03	0.015	0.12	0.068
Life satisfaction on				
Age	0.00	0.007	0.01	0.040
Years education	-0.02	0.016	-0.06	0.043
Depression	-0.20***	0.036	-0.26***	0.046
Partnered	0.22	0.115	0.08	0.040
Health	0.22***	0.056	0.20***	0.048
Black	-0.13	0.167	-0.03	0.034
Log(wealth)	0.20*	0.087	0.10*	0.041
Log(income)	0.05	0.066	0.04	0.052
Social connectedness on				
Age	0.00	0.002	-0.01	0.042
Years education	0.00	0.005	0.02	0.050
Depression	-0.03*	0.010	-0.13**	0.048
Partnered	-0.03	0.036	-0.05	0.047
Health	0.02	0.015	0.06	0.049
Black	0.01	0.043	0.01	0.033
Log(wealth)	-0.01	0.029	-0.01	0.051
Log(income)	0.03	0.016	0.08	0.048

Note. * $p < .05$. ** $p < .01$. ** $p < .001$.

Index model for men. Estimated structural coefficients for the model estimated on the male sample using the index measure of participation are reported in Table 31. Participation significantly predicted social connectedness (β = .18, *SE* = .05, $p < .001$) but not life satisfaction (β = .03, *SE* = .07, $p = .65$). Social connectedness predicted life satisfaction (β = .27, *SE* = .05, $p < .001$). Among the covariates, years education, depression score, partnered, self-report health status, and log of wealth statistically significantly predicted life satisfaction, $p < .05$. Only the depression score significantly predicted social connectedness (β = -.13, *SE* = .05, $p = .005$). The participation composite was significantly predicted by volunteering (β = .56, *SE* = .24, $p = .02$) and by sports/social club participation (β = .53, *SE* = .22, $p = .02$). Except for reading, the other predictors were in the expected direction (positive), but were not significant. This model is similar to the scale model in suggesting that the relationship between life satisfaction and participation is fully mediated by the effect of participation on social connectedness. Beyond what the scale model suggests, it suggests that two kinds of participation – both volunteering and sports/social club participation – may influence social connectedness and therefore affect life satisfaction. Using Mplus' MODEL INDIRECT statement to calculate total and indirect effects showed a significant indirect effect of participation on life satisfaction via social connectedness (β =.048, SE = .014, $p = .001$). The total effect of participation on life satisfaction, however, was not significant (β =.058, SE = .061, p = .345), perhaps because it incorporated the greater uncertainty in the direct relationship between participation and life satisfaction (β =.009, SE = .060, $p = .882$).

Table 31

Structural coefficients – Index model of participation – Male

	Unstandardized		*Standardized*	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.00***	0.187	0.27***	0.049
Participation	0.01	0.033	0.01	0.060
Social connectedness on				
Participation	0.03*	0.013	0.18***	0.045
Participation on				
sp2. Volunteer – other	1.00	0.000	0.56*	0.242
sp4. Sports/social club	0.85	0.640	0.53*	0.223
sp6. Read	-0.40	0.637	-0.24	0.326
sp10. Computer	-0.02	0.291	-0.02	0.298
sp11. Home maintenance/ gardening	0.06	0.397	0.04	0.297
sp14. Hobby/project	0.53	0.437	0.42	0.255
Life satisfaction on				
Age	0.00	0.007	0.00	0.040
Years education	-0.03*	0.017	-0.09	0.044
Depression	-0.20***	0.036	-0.26***	0.046
Partnered	0.24*	0.114	0.08*	0.039
Health	0.22***	0.055	0.20***	0.047
Black	-0.11	0.168	-0.02	0.035
Log(wealth)	0.21*	0.089	0.10*	0.042
Log(income)	0.04	0.066	0.03	0.052
Social connectedness on				
Age	0.00	0.002	0.00	0.041
Years education	0.00	0.005	0.02	0.053
Depression	-0.03**	0.010	-0.13**	0.048
Partnered	-0.03	0.036	-0.04	0.047
Health	0.02	0.015	0.08	0.049
Black	0.02	0.044	0.01	0.033
Log(wealth)	-0.01	0.029	-0.02	0.052
Log(income)	0.03	0.016	0.08	0.047

Note. * $p < .05$. ** $p < .01$. ** $p < .001$.

Scale model for women. Estimated structural coefficients for the model estimated on the female subsample using the eight-factor scale model of participation are reported in Table 32. Social connectedness significantly predicted life satisfaction (β = .27, *SE* = .042, *p* < .001). None of the eight participation factors significantly predicted life satisfaction or social connectedness. Among the covariates, age, depression, partnered, health, and log-transformed wealth significantly predicted life satisfaction, *p* < .05, while depression and years education significantly predicted social connectedness. This model does not support the hypothesized mediational model of how participation relates to life satisfaction. Results were similar with the alternate four-factor model of participation. Life satisfaction was significantly predicted by social connectedness (β = .26, *SE* = .04, *p* < .001) but none of the four participation factors significantly predicted life satisfaction or social connectedness. Thus the alternate, more parsimonious scale model did not offer any support for the hypothesized mediational model either.

Table 32

Structural coefficients – Scale model of participation – Female

	Unstandardized		Standardized	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.340***	0.229	0.27***	0.042
Volunteering	0.171	0.091	0.15	0.077
Club participation	0.081	0.130	0.05	0.075
Games	-0.189	0.179	-0.12	0.108
Intellectual	0.256	0.223	0.19	0.165
Home	-0.752	0.656	-0.25	0.217
Hobbies	0.133	0.092	0.12	0.082
Sports/exercise	0.133	0.111	0.08	0.064
Education	0.033	0.066	0.04	0.069
Social connectedness on				
Volunteering	-0.001	0.018	-0.002	0.077
Club participation	-0.040	0.032	-0.11	0.088
Games	0.026	0.036	0.08	0.106
Intellectual	-0.040	0.043	-0.15	0.159
Home	0.150	0.132	0.24	0.210
Hobbies	-0.027	0.019	-0.12	0.085
Sports/exercise	0.020	0.024	0.06	0.069
Education	-0.004	0.015	-0.02	0.077
Life satisfaction on				
Age	0.020**	0.006	0.11**	0.032
Years education	-0.026	0.017	-0.06	0.035
Depression	-0.162***	0.027	-0.23***	0.038
Partnered	0.211**	0.099	0.08**	0.035
Health	0.217***	0.050	0.17***	0.038
Black	-0.179	0.145	-0.04	0.028
Log(wealth)	0.303**	0.097	0.11**	0.034
Log(income)	0.009	0.042	0.01	0.035
Social connectedness on				
Age	0.001	0.001	0.02	0.036
Years education	0.008*	0.004	0.09*	0.038
Depression	-0.037***	0.007	-0.26***	0.044
Partnered	-0.036	0.021	-0.06	0.037
Health	0.016	0.011	0.06	0.038
Black	0.020	0.033	0.02	0.031
Log(wealth)	0.024	0.022	0.04	0.038
Log(income)	-0.007	0.010	-0.03	0.042

Note. * $p < .05$. ** $p < .01$. ** $p < .001$.

Index model for women. Estimated structural coefficients for the model estimated on the female subsample using the index measure of participation are reported in Table 33. In this model, social connectedness significantly predicted life satisfaction ($\beta = .25$, $SE = .038$, $p < .001$) as did participation ($\beta = .13$, $SE = .037$, $p = .001$). Participation also statistically significantly predicted social connectedness ($\beta = .10$, $SE = .040$, $p = .012$). The participation composite was significantly predicted by charity work ($\beta = .54$, $SE = .19$, $p = .007$), baking/cooking ($\beta = .51$, $SE = .22$, $p = .021$), and hobbies ($\beta = .49$, $SE = .22$, $p = .026$). Life satisfaction was significantly predicted by age, depression, partnered, self-report health status and log-transformed wealth. Social connectedness was significantly predicted by years education and depression. This model supports the hypothesized mediational model of the relationship between participation and life satisfaction. In contrast to the male models, which suggested full mediation by social connectedness of the relationship between participation and life satisfaction, these results are consistent with partial mediation. Using Mplus' bootstrapping capabilities to calculate total and indirect effects with standard errors estimated a significant total effect of participation on life satisfaction ($\beta =.15$, SE = .036, $p < .001$) and a significant indirect effect via social connectedness ($\beta =.025$, SE = .011, $p = .021$).

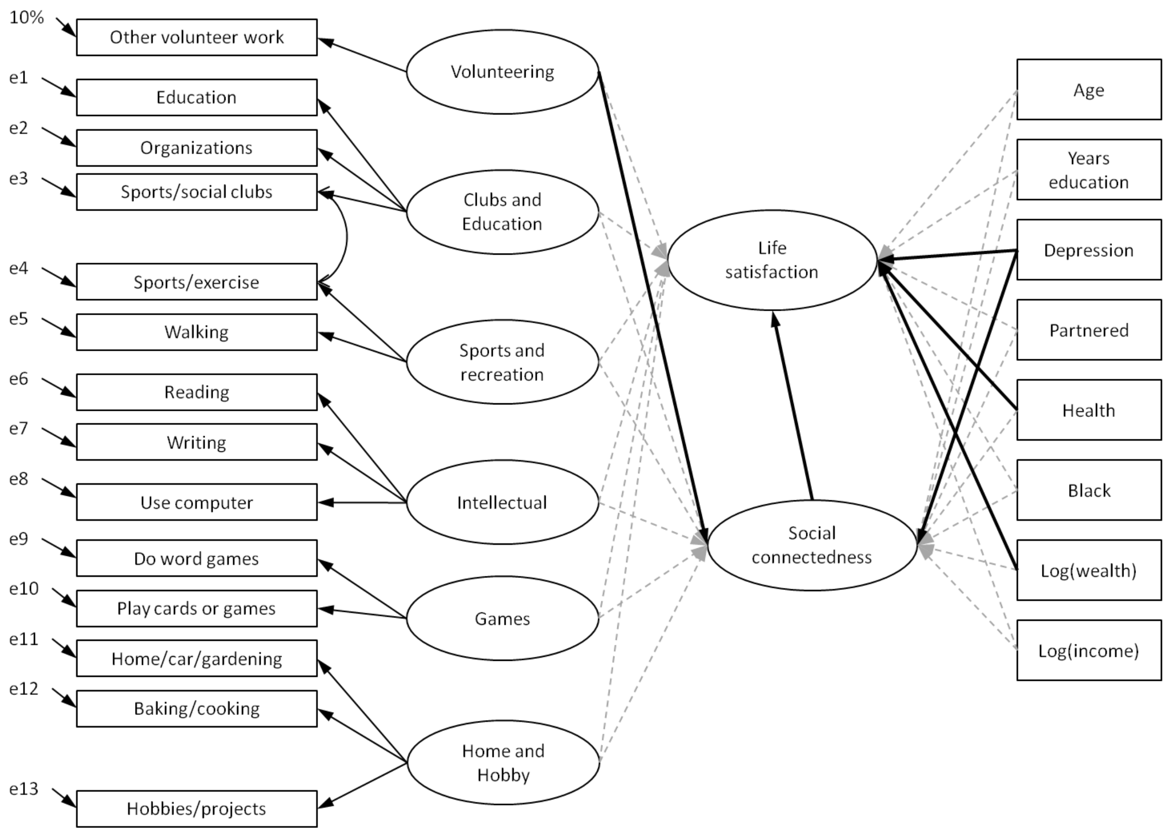
Table 33

Structural coefficients – Index model of participation – Female

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.221***	0.204	0.25***	0.038
Participation	0.079*	0.035	0.13**	0.037
Social connectedness on				
Participation	0.013**	0.007	0.10*	0.040
Participation on				
sp2. Charity work	1.000	0.000	0.54**	0.198
sp4. Sports/social club	-0.490	0.461	-0.29	0.251
sp6. Read	-0.482	0.487	-0.26	0.247
sp10. Computer	0.147	0.288	0.14	0.258
sp12. Bake or cook	0.712	0.484	0.51*	0.223
sp14. Hobby	0.606	0.398	0.49*	0.221
Life satisfaction on				
Age	0.021**	0.006	0.12***	0.033
Years education	-0.027	0.017	-0.06	0.035
Depression	-0.171***	0.027	-0.24***	0.038
Partnered	0.196	0.096	0.07*	0.034
Health	0.214***	0.050	0.16***	0.038
Black	-0.207	0.143	-0.04	0.028
Log(wealth)	0.305**	0.095	0.11**	0.033
Log(income)	0.004	0.042	0.00	0.036
Social connectedness on				
Age	0.001	0.001	0.03	0.036
Years education	0.011**	0.004	0.11**	0.037
Depression	-0.037***	0.007	-0.26***	0.043
Partnered	-0.040	0.021	-0.07	0.037
Health	0.014	0.011	0.05	0.038
Black	0.015	0.032	0.01	0.030
Log(wealth)	0.022	0.022	0.04	0.038
Log(income)	-0.005	0.010	-0.02	0.043

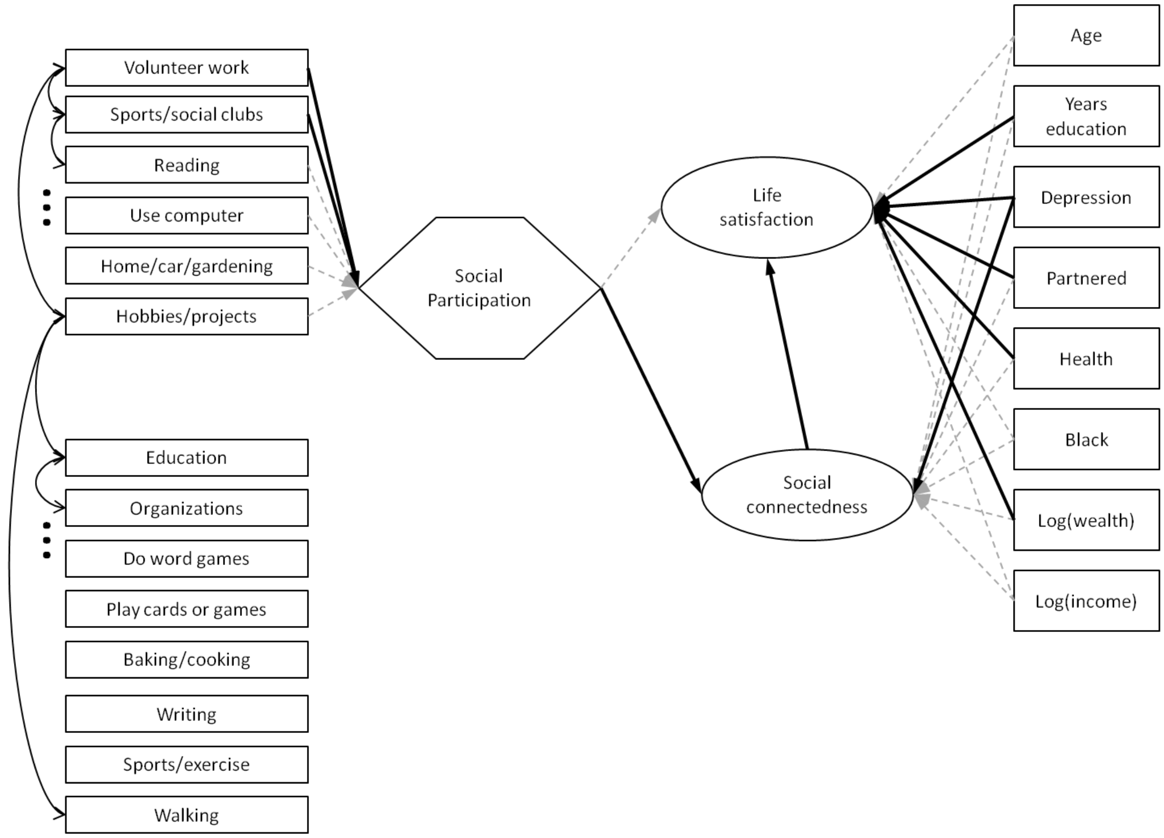
Note. * *p* < .05. ** *p* < .01. ** *p* < .001.

Comparison of four models. For comparison purposes, results for each full structural model (male – scale, male – index, female – scale, and female – index) are illustrated in Figures 15, 16, 17, and 18. Structural paths significant at the $p < .05$ level are shown as black, non-dashed lines while nonsignificant paths are grayed out and dashed. These results make clear that the choice of a scale versus index approach does not lead to the same conclusions about the structural relations across participation, social connectedness, and life satisfaction, for female respondents more so than male respondents in this case. The results also suggest that the structural relations differ by gender. The model for females using the index measure of participation most nearly corresponded to the hypothesized model in which social connectedness partially mediates the relationship between participation and life satisfaction. The index model for females combined multiple types of participation – volunteering, baking/cooking, and hobbies – in its participation index. All of the models used multiple correlated predictors, so the lack of statistically significant prediction from different kinds of participation may reflect correlations across predictors. Dropping some participation subscales from the scale model might result in statistically significant predictive power for the remaining subscales. Likewise, a more parsimonious index definition might result in additional statistically significant participation composite predictors.



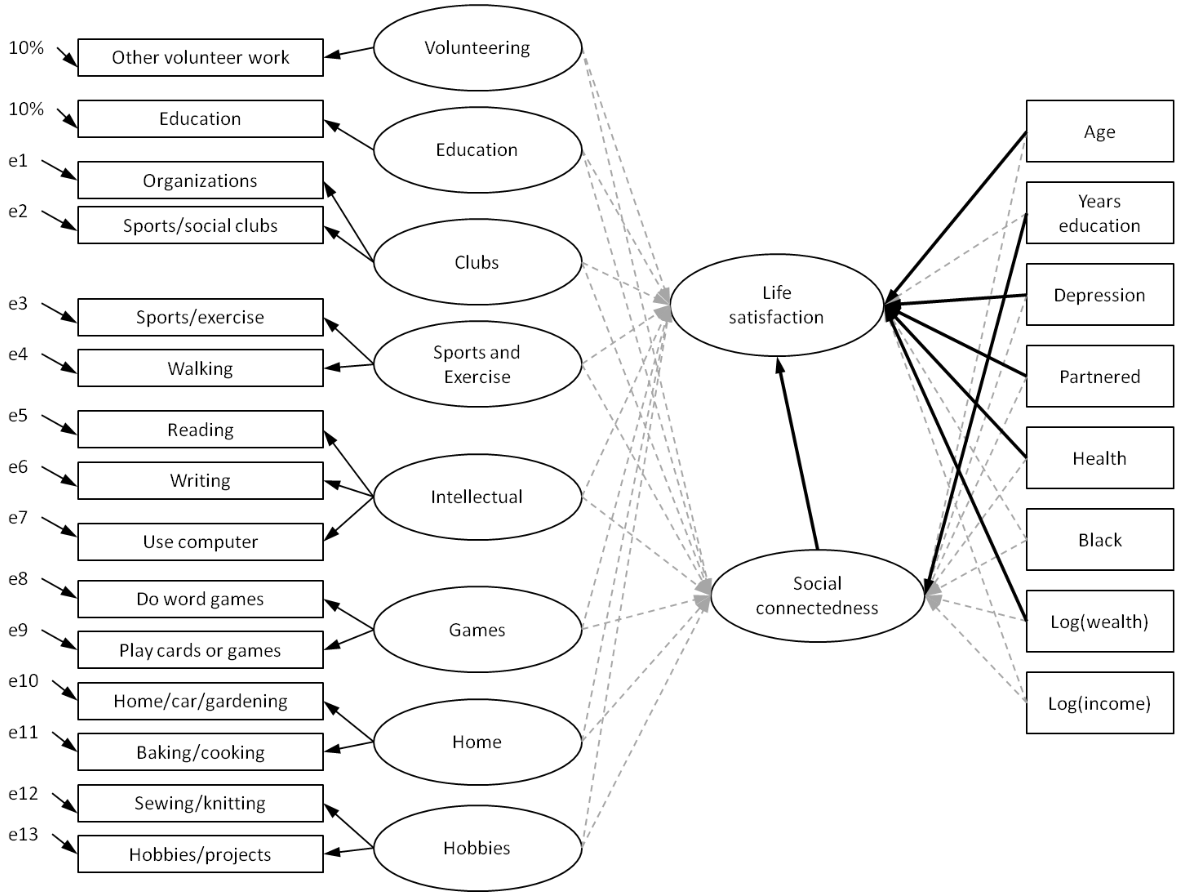
Note. All participation factors specified as covarying (not shown in figure). Darkened structural relations are significant at $p < .05$.

Figure 31. Structural model results – Scale model of participation – Male



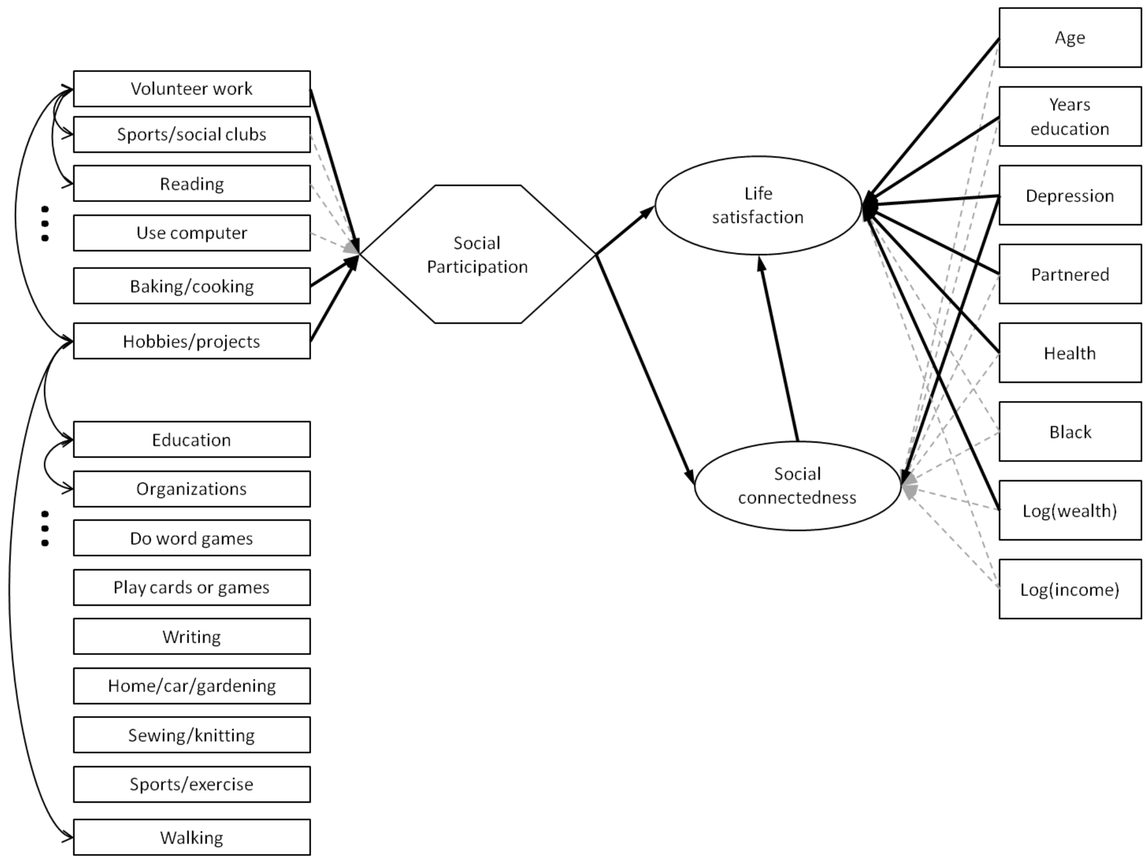
Note. All participation items modeled as single-indicator latent factors with 10% residual variance specified. Darkened structural relations are significant at $p < .05$.

Figure 32. Structural model results – Index model of participation – Male



Note. All participation factors specified as covarying (not shown in figure). Darkened structural relations are significant at $p < .05$.

Figure 33. Structural model results – Scale model of participation – Female



Note. All participation items modeled as single-indicator latent factors with 10% residual variance specified. Darkened structural relations are significant at $p < .05$.

Figure 34. Structural model results – Index model of participation – Female

Interpreting the Index Models

The index models showed more promise than the scale models in answering the question, "how does social participation in late adulthood relate to social connectedness and life satisfaction?" The scale construction process identified many small factors so that when these were embedded into the full structural model there was no representation of the theoretical construct of participation at a general level. When more generic factors were used in the alternate female model, fit to empirical data was compromised. The index model, on the other hand provided a way of considering participation in a general fashion rather than in disaggregated fashion while still maintaining adequate fit. In this section, results from the full structural models fit to the validation subsample using participation indexes by gender are reported and interpreted.

Gelman and Hill (2007) suggested keeping predictors in regression models even if they are not significant, so long as they are in the expected direction but suggest deleting nonsignificant predictors that are in the “wrong” direction. This study hypothesized that different types of participation sum together to produce an overall level of participation that is positively associated with social connectedness and life satisfaction. Therefore, predictors with negative coefficients are not in the expected direction. For males, two items in the participation composite were negative, reading and using a computer. For the final structural model reported and interpreted here, these two items were dropped. For females, there were also two items with negative coefficients in the participation composite estimated in the full structural model: sports/social club participation and reading. For the final model for females reported and interpreted here, those two items were dropped. This resulted in four-item participation composites for men and for women. The covariates income and black were eliminated from the analyses since they hadn’t shown any significant association with the outcomes in prior analyses.

Final full structural model for men. The full structural model incorporating the four-item model for men consisting of volunteering, sports/social club participation, home/car maintenance or gardening, and hobbies/projects had good fit, $\chi^2(176) = 382.927$, RMSEA = .037, CFI = .95, SRMR = .052. Modification indexes suggested that fit could be improved by modeling sports/social club participation predicted by wealth, home maintenance/gardening predicted by age, and hobbies/projects predicted by years of education. This expanded model also had good fit, $\chi^2(173) = 347.90$, RMSEA = .037, CFI = .96, SRMR = .045, and showed significantly better fit than the original model, scaled $\Delta\chi2(3) = 26.67$, $p < .001$. Estimated structural parameters are reported in Table 34; full model results are reported in Table D.1 in Appendix D.

Table 34

Structural coefficients for final model of life satisfaction and social connectedness as related to social participation – Male respondents

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.011***	0.186	0.274***	0.049
Participation	-0.005	0.037	-0.008	0.057
Social connectedness on				
Participation	0.030*	0.012	0.173***	0.045
Participation on				
Volunteering	1.000	0.000	0.637**	0.219
Sports/social clubs	0.679	0.504	0.481*	0.225
Home/car maintenance or gardening	-0.030	0.370	-0.025	0.321
Hobby/project	0.424	0.367	0.384	0.258
Life satisfaction on				
Age	0.000	0.007	0.001	0.040
Years education	-0.027	0.015	-0.072	0.040
Depression	-0.198***	0.036	-0.264***	0.046
Partnered	0.264*	0.110	0.093*	0.038
Self-report health	0.223***	0.054	0.200***	0.046
Log(wealth)	0.243**	0.078	0.116**	0.038
Social connectedness on				
Age	-0.001	0.002	-0.013	0.041
Years education	0.002	0.005	0.023	0.045
Depression	-0.027**	0.010	-0.131**	0.048
Partnered	-0.019	0.034	-0.025	0.044
Self-report health	0.023	0.015	0.076	0.049
Log(wealth)	0.004	0.026	0.008	0.046
Sports/social clubs on				
Log(wealth)	0.327**	0.118	0.141**	0.051
Home/car maintenance or gardening on				
Age	-0.036***	0.010	-0.147**	0.042
Hobby/project on				
Years education	0.074**	0.022	0.137**	0.041

Note. * $p < .05$. ** $p < .01$. ** $p < .001$.

Life satisfaction was significantly predicted by social connectedness ($\beta = .27$, $SE = .05$, $p < .001$) but not by participation ($\beta = -.01$, $SE = .06$, $p = .88$). Social connectedness was predicted by participation ($\beta = .17$, $SE = .05$, $p < .001$) which was significantly defined by volunteering ($\beta = .63$, $SE = .22$, $p = .004$) and by sports/social club participation ($\beta = .48$, $SE = .23$, $p = .03$). Home/car maintenance or gardening was a negative but nonsignificant predictor of participation ($\beta = -.03$, $SE = .32$, $p = .94$) while hobbies/projects showed a positive but nonsignificant contribution to the participation composite ($\beta = .38$, $SE = .25$, $p = .14$). Bootstrapped calculations for indirect and total effects showed a significant indirect effect from participation to life satisfaction via social connectedness ($\beta = .05$, $SE = .01$, $p = .001$) but a nonsignificant total effect of participation on life satisfaction ($\beta = .04$, $SE = .06$, $p = .51$). The direct effect from participation to life satisfaction was negative but nonsignificant ($\beta = -.01$, $SE = .06$, $p = .89$). This model did not support the hypothesized model of the relationship between participation and life satisfaction, which theorized only partial mediation. The results are consistent with full mediation, since any effect of participation on life satisfaction appears to be fully explained by participation's effect on social connectedness.

Sports and social club participation was significantly predicted by log-transformed wealth ($\beta = .14$, $SE = .05$, $p = .006$). Home/car maintenance or gardening was negatively predicted by age ($\beta = -.15$, $SE = .04$, $p = .001$). Hobby/project participation was positively predicted by years education ($\beta = .14$, $SE = .04$, $p = .001$). While these predictors were added based solely on empirical considerations, they do make substantive sense. Wealthier adults are more likely to join country or sports clubs, which may require membership or other fees and some of which cater only to those of high socioeconomic status. Physical projects such as home maintenance and gardening not only become less feasible as one ages and perhaps becomes less physically capable but also are less necessary if someone has moved into a living situation where maintenance tasks are taken care of by someone else. The association between hobby/project participation and level of education deserves further attention, as it is not immediately obvious why the less educated should be less likely to participate in hobbies or projects. Perhaps both these reflect an underlying drive to engage with productive activity such as education at younger ages and hobbies or projects at older ages.

Final full structural model for women. The full structural model incorporating the four-item model for women consisting of volunteering, computing, baking/cooking, and hobbies/projects had adequate fit, $\chi^2(176) = 688.41$, RMSEA = .048, CFI = .91, SRMR = .07. Modification indexes suggested that some of the participation items should be modeled as predicted by demographic covariates such as wealth, age, and level of education. An expanded model which included predictors of participation items was fit. Computer usage was predicted by age, education, and log-transformed wealth; baking or cooking something special was predicted by partner status (married/living with a partner vs. single); and hobby/project participation was predicted by years of education. This model had good fit, $\chi^2(171) = 440.59$, RMSEA = .035, CFI = .95, SRMR = .049 and it was significantly better than the model without participation predictors, scaled $\Delta\chi^2(5) = 315.58$, $p < .001$. Structural parameters are reported in Table 35; full results of the final model are reported in Table D.2 in Appendix D.

Table 35

Structural coefficients for final model of life satisfaction and social connectedness as related to social participation – Female respondents

	Unstandardized		***Standardized***	
	Coefficient	***SE***	***Coefficient***	***SE***
Life satisfaction on				
Social connectedness	1.205***	0.205	0.247***	0.038
Participation	0.075*	0.034	0.129**	0.038
Social connectedness on				
Participation	0.014	0.007	0.120**	0.041
Participation on				
Volunteering	1.000	0.000	0.493**	0.181
Computer	0.197	0.292	0.171	0.234
Baking/cooking	0.660	0.481	0.437*	0.216
Hobby/project	0.646	0.408	0.477*	0.197
Life satisfaction on				
Age	0.022**	0.006	0.121***	0.034
Years education	-0.034*	0.016	-0.072*	0.034
Depression	-0.168***	0.027	-0.238***	0.038
Partnered	0.199*	0.094	0.071*	0.033
Self-report health	0.211***	0.050	0.161***	0.037
Log(wealth)	0.313***	0.088	0.110***	0.031
Social connectedness on				
Age	0.001	0.001	0.036	0.036
Years education	0.009*	0.004	0.090*	0.037
Depression	-0.037***	0.007	-0.255***	0.043
Partnered	-0.045*	0.020	-0.078*	0.035
Self-report health	0.011	0.010	0.042	0.038
Log(wealth)	0.014	0.020	0.024	0.035
Computer on				
Age	-0.073***	0.008	-0.266***	0.029
Education	0.213***	0.021	0.295***	0.030
Log(wealth)	0.714***	0.128	0.167***	0.031
Baking/cooking on				
Partnered	0.675***	0.102	0.210***	0.032
Hobby/project on				
Years education	0.131*	0.019	0.213***	0.031

Note. * $p < .05$. ** $p < .01$. ** $p < .001$.

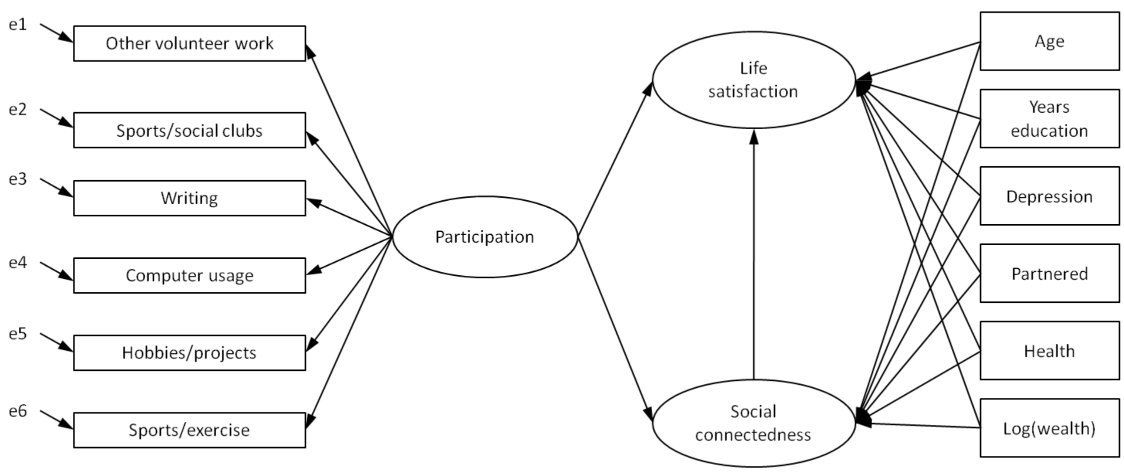
Life satisfaction was significantly predicted by social connectedness ($\beta = .25$, $SE = .04$, $p < .001$) and by participation ($\beta = .13$, $SE = .04$, $p = .001$). Social connectedness was significantly predicted by participation ($\beta = .12$, $SE = .04$, $p = .003$). The indirect effect of participation on life satisfaction via social connectedness was significant ($\beta = .03$, $SE = .01$, $p = .007$) and the total effect was also statistically significant ($\beta = .16$, $SE = .04$, $p < .001$). These findings are consistent with the hypothesized mediational model of the relationship between social participation and life satisfaction which proposed partial mediation via social connectedness. In this model, the participation composite was significantly defined by volunteering ($\beta = .49$, $SE = .18$, $p = .007$), baking/cooking ($\beta = .44$, $SE = .22$, $p = .043$), and by hobby/project participation ($\beta = .48$, $SE = .20$, $p = .016$).

Considering a One-Dimensional Scale Model

The scale model development process was fraught with complexity and ambiguity. Different criteria for identifying the number of factors pointed to varying numbers of factors: parallel analysis suggested the presence of seven, the Kaiser criterion identified four, and scree plots suggested the presence of just one important factor. Items jumped between factors when four and five-factor solutions were estimated. Chi-square difference tests supported the addition of many more factors than seemed reasonable. The fact that the index model worked well with a unidimensional solution suggested that participation might be best treated unidimensionally. However, the one-factor model for females using fifteen items didn't compare well to the index model that used fewer items, perhaps because it was only included as an afterthought. Might a better one-factor model, one with a reduced set of items, produce performance that was comparable to or better than the one-index model? Ad hoc analyses were undertaken to explore whether this was the case. The analysis was limited to the female subsample in order to keep the scope manageable. Principal axis factoring on the fifteen participation items other than the volunteering with youth item was estimated, specifying just one factor. Items with loadings greater than .40 were selected for inclusion in the refined one-factor model; almost all items had loadings greater than .30. This resulted in a six-item model of social participation that included volunteering (other than with children), sports/social club participation, writing, computer usage, hobbies/projects, and sports/exercise. This model had substantial overlap with the six-item index model of social participation for females. It showed adequate fit in the exploratory half of the data set, $\chi^2(9) = 39.93$, RMSEA = .059, CFI = .91, SRMR = .035 and good fit in the confirmatory half, $\chi^2(9) = 24.83$, RMSEA = .037, CFI = .97, SRMR = .027. Reliabilities were still rather low, $\alpha = .59$ in the exploratory half and $\alpha = .64$ in the confirmatory half. These reliabilities could not be increased by deleting any items.

The full structural model using this one-factor model of participation (Figure 35) was fit in the validation half of the data set. As in fitting the final index models, the covariates black and log-transformed income were dropped so the remaining covariates were age, years education, partnered, health, and log-transformed wealth. The model showed adequate fit, $\chi^2(224) = 823.54$, RMSEA = .046, CFI = .90, SRMR = .075. Modification indexes suggested that participation should be predicted from covariates such as age, education, and health status so all covariates in the model were added as predictors of the latent participation variable. This model also had adequate fit, $\chi^2(218) =$

564.36, RMSEA = .035, CFI = .94, SRMR = .042. Life satisfaction was significantly predicted by both social connectedness (β = .25, *SE* = .04, *p* < .001) and by participation (β = .13, *SE* = .20, *p* = .044). The indirect effect of participation on life satisfaction via social connectedness was also significant (β = .04, *SE* = .02, *p* = .048). The total effect of participation on life satisfaction was significant as well (β = .17, *SE* = .07, *p* = .015). These results are quite similar to those from the final index model of participation for female respondents, and are consistent with the hypothesized model of the relationship between life satisfaction and social participation as partially mediated by perceived social connectedness. Parameter estimates for this model are reported in Table D.3.



Note. Life satisfaction and social connectedness indicators and indicator errors not shown.

Figure 35. Full structural model with single-factor scale model for females

Summary

While a four-factor scale model of participation and a two-composite index model were hypothesized, results suggested that treating participation (at least as measured by the items in this study) unidimensionally makes sense. Was the scale or index model better? Conventional scale construction techniques such as inspection of percent variance explained, the Kaiser criterion, and parallel analysis (from EFA) as well as chi-square difference tests and approximate fit indexes (from CFA) led away from a unidimensional model of participation. The index construction techniques offered here, however, led to a unidimensional model; a single-composite model fit as well as did the hypothesized two-composite model and it performed well when pitted against the multi-factor models developed using conventional scale development techniques.

It was the unidimensionality finding of the index approach that led this researcher to reconsider a one-factor scale model solution. The scree plots showed just one important factor in the participation data, whether considered across the data set or by gender. Perhaps this plot should have been paid more heed, since the ad hoc development of a one-factor participation model for females showed good performance in both measurement and structural models. This single-factor model was consistent with the hypothesized model of the relationship between social participation, suggesting that the relationship is partially mediated by social participation's effects on social connectedness. Implications of these findings are discussed in detail in chapter four.

Chapter Four: Discussion

Social participation in older adulthood can improve life satisfaction and a sense of social connectedness. As people age, however, they may face physical and social barriers that prevent them from participating in all the activities they would like. Researchers who study successful aging, whether in the presence of disability or not, are interested in measuring and modeling participation. Typical psychometric methods for scale development based on reflective measurement and the common factor model may not be the most appropriate way of modeling participation when using indicators measuring a respondents' frequency of participation in different types of activities. Because of low intercorrelation across activity types and correlation patterns that do not match researchers' conceptions of the dimensions of participation, models built using conventional approaches based on classical test theory may show poor reliability and a proliferation of subscales that do not have content or predictive validity.

Because of the problems encountered using reflective indicators to develop participation scales, some researchers have proposed that an index model, which could be expressed using a formative measurement approach, may better transform participation frequency data into a useful quantification of a person's level of participation. Index models based on formative measurement select items based on their predicted outcomes and their unique predictive ability net of other predictors in the model while factor models select and group items based on intercorrelations with each other. This research study compared participation scales constructed with conventional psychometrics based on classical test theory with reflective indicators to participation indexes constructed with formative or causal indicators that use patterns of prediction of outcomes rather than indicator intercorrelations to define measures.

Summary

The purpose of this study was to compare a scale model of social participation, based upon reflective measurement, to an index model of social participation, based upon formative measurement, in the context of participation's association with perceived social connectedness and life satisfaction. The results of this comparison generated recommendations about usefully modeling participation.

The research questions were:

(1) What measurement model for social participation has stronger validity: reflective (scale) or formative (index)? There are at least two types of validity at issue here: first, how well does the model fully capture a researcher's conception of social participation (content validity) and second, how well does the model predict outcomes (external or predictive validity)?

(2) What dimensions of social participation should be modeled? Are unidimensional or multidimensional conceptions of participation more conceptually and predictively valid?

(3) What activities should be included in the definition of social participation? This is related to the dimensionality of participation. Unidimensional representations of

participation, for example the single-factor model of participation or single-index model of participation, eliminated certain activities. When is this warranted and when isn't it?

These research questions were driven by a recognition in the research literature, especially in disability and rehabilitation research, that scale models of participation have often failed to accurately represent participation and haven't always provided useful ways to develop single-dimensional or multi-dimensional measures of participation. The formative approach, however, which specifies that levels of different activities sum together to produce an overall level of participation suffers from inadequate methodological development and questions about its philosophical groundings.

Summary of method. Data from the 2008 wave of the Health and Retirement Study representative of community-dwelling adults age 65 and over were used to produce the scale and index models from 16 items that covered a broad range of participatory activities such as volunteering, hobbies, and sports. The scale and index approaches were evaluated according to their theoretical meaningfulness, parsimony, fit to empirical data and predictive validity. A mediational model of the relationships across social participation, perceived social connectedness, and life satisfaction was hypothesized in which participation was specified as directly influencing life satisfaction and indirectly influencing it via social connectedness. This model was used to both identify the index model and to characterize the predictive validity of both scale and index models. Cross-validation was used in order to avoid capitalization on chance variation in the data set; the data set was divided in two and one half was used for measure construction while the other half was used for measurement model validation and structural model comparison.

During measure construction, invariance across genders was explored, since patterns of participation and their associations with outcomes were considered likely to differ by gender.

Scale construction results. Exploratory and confirmatory factor analyses were used to construct the scale model of social participation. EFA criteria for selecting number of factors were equivocal: the Kaiser criterion suggested four factors while parallel analysis identified seven and the scree plot just one. These results held whether EFAs were estimated across the entire exploratory sample or by gender. The four-factor hypothesized model had inadequate fit to the data, whether estimated across both genders or for males or females only. The model was re-specified using changes hypothesized a priori as well as changes suggested by modification indexes, so long as they made substantive sense. It was not possible to develop a model that fit adequately for both genders, so separate scale models were constructed. This implied a lack of configural invariance across genders. The final male model had six factors: volunteering, clubs and education, sports and exercise, intellectual, home and hobby, and games. It included a correlated residual across two items mentioning sports that loaded on two different factors. The scales had low reliability in the exploratory half of the data set, ranging from .35 for games to .64 for sports and exercise. This was (at least partially) related to the low numbers of items for each dimension. The female model had eight factors: volunteering, education, clubs, sports and exercise, intellectual, games, home, and hobby. Like the model for males, it had a correlated residual for the two items mentioning sports. The subscale reliabilities for the scale model for females were also low, ranging from .36 for volunteering to .64 for hobbies. Again, this was in part due to the small number of items per factor. The items had not been designed from a common factor model perspective and so didn't present multiple redundant measures of underlying factors. Because of the extreme complexity of the female model, alternate single-factor and four-factor models were also developed and considered in the comparison with the index models of social participation.

Index construction results. A two-index model specifying domestic and community participation measures was hypothesized. It was estimated using a structural model that specified life satisfaction and social connectedness as outcomes, with the indexes modeled as zero-error composites made up of participation items. Participation items were modeled as latent variables with 10% residual variance in order to account for measurement error. Models with zero, 20%, and 30% error were also estimated; results appeared similar across the different levels of residual variance specified. The two-index model had adequate fit in the exploratory half of the data set. A single-index model (combining all sixteen participation items in one composite predicting social connectedness and life satisfaction) was compared to the two-index model to see if the fit significantly deteriorated; it did not, so the single-index model was retained. The two-index and single-index models were also fit in the male and female subsamples; again, the single-index model was not significantly worse than the two-index model, so it was retained.

Items were trimmed from the model if they did not have statistical significance at the $p < .10$ level for the model fit to the entire sample, the model fit with male respondents only, or the model fit with female respondents. This generous significance level was used to screen items for the composite given that intercorrelations among items would tend to inflate *p*-values when all items were included. The screening process resulted in a seven-item index of social participation. When coefficients on the remaining items were constrained to be zero so that they did not enter into the index, the fit was adequate, for the entire sample, for the female sample, and for the male sample, though not all seven items had statistically significant contributions to the model in each case. The male and female models were refined by dropping items that either had negative coefficients (male model) or had already been dropped by the scale model (female respondents). This led to the elimination of the baking/cooking item from the male index model and the elimination of the home/car maintenance or gardening item from the female index model. Thus both genders had six-item index models to be used in comparison with the scale models, but the two indexes had different items. The index models as developed, like the scale models, were therefore not configurally invariant across genders.

Comparison of models. The scale and index models were fit in the overall structural model in the validation half of the data set with covariates specified as predicting the outcome variables. Because neither the scale nor the index models showed invariance across genders, the models were fit for male and female subsamples separately. The scale models explained more variance in outcomes than the index models while the index models had slightly better fit statistics. BIC values chose the male index model and the female scale model, but this criterion cannot tell the difference between important and unimportant complexity. For example, the index models were disadvantaged by having to explain correlations across participation items that were specified as having zero contribution to the index. These items had to be included in order to make the models comparable for the purposes of using information criteria but in actual use of the index model these would have been entirely dropped. Also, the scale models didn't have any small factors trimmed, even though that might have been a reasonable step to take were there not the requirement to maintain the same items in each model for comparison purposes. BIC values therefore didn't seem very useful in distinguishing among the models. In terms of variance explained in outcomes, the multi-factor scale models seemed superior: the six-factor male and the eight-factor female models did explain more variation in life satisfaction and social connectedness than the single-index models explained.

The index model for female respondents provided evidence in favor of the hypothesized relationships among social participation, social connectedness, and life satisfaction although the male index model did not. In the index model for females, volunteer work, baking/cooking, and hobbies all significantly contributed to social participation, which itself predicted both social connectedness and life satisfaction. Social connectedness also predicted life satisfaction, so the hypothesized mediational model was supported. In the index model for male respondents, social participation directly predicted social connectedness and indirectly predicted life satisfaction but did not have a direct effect on life satisfaction. In this model, only sports/social club activities significantly contributed to social participation. The scale model for male respondents suggested that among the participation subscales, only volunteering predicted social connectedness and none of the participation subscales predicted life satisfaction. In the eight-factor scale model for female respondents, none of the participation items predicted social connectedness or life satisfaction.

Ad hoc analysis of single-factor female model. Because the index model suggested a unidimensional measure of participation, a single-factor model of social participation for female respondents was explored. A trimmed single-factor social participation model was developed using results of a principal axis factoring specifying just one factor. The six items that loaded at greater than .40 were kept in the model: volunteering (other than with children), sports/social club participation, writing, computer usage, hobbies/projects, and sports/exercise. This model showed adequate fit in exploratory and confirmatory subsamples. When fit in the full structural model with covariates, the model showed good fit. The structural relationships among the participation latent factor, perceived social connectedness, and life satisfaction were similar to those estimated in the final structural model using the index model of social participation for females. Reliability of this unidimensional participation scale was low, around .60. It was, however, the best-performing and most parsimonious scale model developed.

Major Findings

Research question one: Scale or index model of social participation? Neither the scale nor the index model was entirely satisfactory. Both lacked predictive validity in the full mediational model with covariates, explaining little or no additional variation that could be explained by the covariates such as age, health status, and wealth. The scale models developed starting from the hypothesized four-factor model were not parsimonious and not invariant across genders. The index model also did not show invariance across genders, but it provided a way of modeling participation as a unitary construct, when specified in gender-specific ways. A single-factor ad hoc scale model of social participation for females showed some promise, suggesting that the scree plot finding of one major factor should be further explored.

The lack of predictive validity for both models when covariates were included may reflect both a relatively weak relationship across participation and the outcomes of interest and could also reflect model misspecification. While a range of studies found evidence in favor of moderate to strong associations between participation, well-being, and social support or connection (e.g., Harlow & Cantor, 1996; Wahrendorf et al., 2008; Warr et al., 2004), other studies found only weak or qualified associations (Levasseur et al., 2004; Wahrendorf et al., 2008); the lack of predictive power of both the scale and index model may reflect limited underlying associations between participation, social connectedness, and well-being. Model misspecification may also be a factor. When the full structural model was estimated with covariates and with the scale or index measurement model of participation, covariates were only used to control for differences in the outcomes of life satisfaction and social connectedness. But some of the covariates likely influenced participation levels; for example, less healthy people may not be able to participate as much as they'd like, and this could impact their feelings of social connectedness and life satisfaction. Leaving out these relationships would tend to bias the estimates of the relationships between participation and the outcomes. Adding these predictors into the index models and the single-factor scale model did not appreciably change results, but improved model fit.

The potential for misspecification points to the need for more sophisticated modeling when constructing indexes (and perhaps when constructing scale models as well). The approach used here for index construction was similar to a multivariate multiple regression (multiple regression with multiple outcome variables) once a single-index model was settled upon. In any regression analysis problem, the researcher must develop a properly specified model; otherwise coefficient estimates will be biased. In scale construction measures are developed without reference to any confounding variables; the assumption is that correlations across items will solely arise from latent common factors, although addition of residual correlations may address violations of this assumption. In some scale development settings, third variables can and do confound measurement models. For example, consider the life satisfaction item, "the conditions of my life are excellent." For someone who suffers from poor health but otherwise is happy with their life, the item may tap into two separate latent constructs: perceived health and perceived life satisfaction. Thus the problem of confounding does not apply only to index construction. It seems more pressing in that setting since regression analysis almost always considers confounders while factor analysis rarely does. Since factor analysis for the purposes of scale construction typically takes place without reference to possible confounders, it made sense here to pursue index construction in a similar fashion, that is, without using covariates to control for differences across respondents.

Another way in which the index model may be misspecified is by a lack of interaction terms or other model features which might reflect how participation in different activities could either reinforce or detract from effects on life satisfaction. The use of nonlinear frequency response points incorporated a flattening of effects at higher levels of participation, but the model could be further manipulated to represent additional theory about how levels of participation in different kinds of activities might interact. Some participation researchers have suggested that absolute levels of participation matter less than a balance across diverse roles (e.g., Bogner et al., 2011). A term specifying the number of roles in which a person participates could be included in the model in order to allow diversity of roles to contribute to the participation composite's predictive power. While a linear regression model with first-order terms is a logical starting point for index construction, it is by no means the only or best way to represent how different items contribute to a composite variable.

More sophisticated modeling might help achieve greater predictive power for an index model, but another possible way to address this problem would be to specify outcomes that are closer to the construct of interest. Life satisfaction and perceived social connectedness are far downstream from participation and many demographic and other variables are likely to confound the relationships that hold across participation and these variables. An index measure of participation designed from scratch could include items that directly reflect the researcher's conception of what social participation is. The Community Integration Measure (CIM; McColl et al., 2001) is a subjective measure of community integration with items that might serve this purpose. Its items include "I feel like part of this community, like I belong here" and "I have something to do in this community during the main part of my day that is useful and productive." Using such items might help the index model function better since the associations between participation levels and these subjective outcome items should be relatively strong while at the same time less confounded by respondent differences compared to associations between participation levels and the outcomes used in this study. Also, additional subjective items could be used to better delineate different dimensions of participation. For example, if domestic participation is conceived of as productive activity around the house, items such as "I have something to do around the house during the main part of my day that is useful and productive" could be used to distinguish domestic participation from community participation defined by items such as those from the CIM. Adding such items would better identify the model and, if multiple such items were used, could provide for estimation of an error term for the composite variable. An example of this approach is shown in Figure 36. This is essentially a multiple-indicator, multiple-cause (MIMIC) model (Jöreskog and Goldberger, 1975); it doesn't suffer from the same underidentification issues that formative index models using correlated outcome constructs must address.

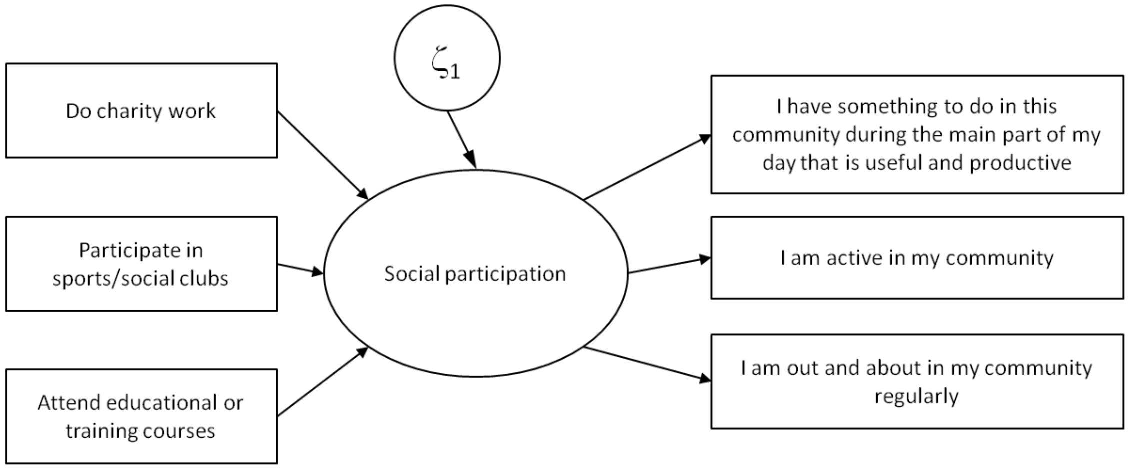


Figure 36. MIMIC model for index construction

The ad hoc analysis using a refined single-factor model of social participation for females suggested that reflective approaches could be useful in developing parsimonious models of social participation. But some conventional practices in exploratory and confirmatory factor analysis would tend to lead researchers away from such a single-dimensional model and consequently away from a model which might prove to have good empirical fit, admirable parsimony, and acceptable predictive validity. For example, one rough rule of thumb says that a factor solution should account for at least 50% of variance in the items (Meyers, Gamst, & Guarino, 2006). This would be achieved in the present case only with solutions having five or more factors. A one-factor solution explained just 20% of the variance, so it would not be a natural or obvious choice for a researcher, even though it might provide, in the end, a parsimonious solution with good fit, once ill-fitting items were discarded. When CFA techniques were used, chi-square difference tests called for more and more factors and CFI values said that the more parsimonious solutions were inadequate; they pointed toward a solution that fit the data but did not have the ability to address the question at issue, which was formulated in terms of a general participation construct rather than in terms of specific participatory activities.

The scale model might have been usefully specified as hierarchical factor model with a general second-order participation construct measured by first-order specific participation types such as community participation, home-based participation, sports/exercise participation and so forth. Just as with the formative model, a reflective model needs careful thought and specification. In neither case can simple algorithmic rules or rule of thumb cutoffs determine an ideal model.

In conclusion, the procedures here did not result in a definitive answer as to whether the scale approach or index approach was clearly better than the other for modeling participation. Scale construction approaches were problematic, as participation instrument developers have already realized. The index construction process might be criticized as ad hoc and the indexes that resulted showed some evidence of interpretational confounding, where regression coefficient signs changed direction depending on what else was in the model. Yet both approaches showed some potential in providing a model of the construct of interest: social participation at a general and abstract level.

Research question two: Dimensions of social participation. The initial scale construction procedures hypothesized a four-factor model and then used hypothesized additional factors as well as the results of EFA to respecify the model until there was adequate fit. This process identified many more dimensions of participation than seemed useful, especially in the case of female respondents, where eight factors (two defined by single indicators) were used to explain just 16 items. Even if four- or five-factor solutions had been developed, this doesn't map onto many researchers' sense that participation activities of the sort analyzed here can be captured as a unitary concept or perhaps a concept with just two or three dimensions. Given past empirical research, however, it is not surprising that the common factor approach identified so many dimensions. Harlow and Cantor (1996), for example, undertook a cluster analysis on a set of 33 participation items gathered in 1977 from a sample of older adults to identify domains of participation. After extracting eight multi-item domains, they still had five items that did not belong to any domain, so they found 13 domains in total, resulting in an average of 2.5 items per domain.

This is many more domains than participation instrument developers have theorized. For example, the Maastricht Social Participation Profile (MSPP; Mars et al., 2009) includes four sub-indexes covering consumptive participation, formal social participation, informal social participation with acquaintances, and informal social participation with family. The Frenchay Activities Index (FAI) uses three subscales: domestic, leisure/work, and outdoors (Schuling et al., 1993). The Community Integration Questionnaire (CIQ; Willer et al., 1993) organizes items into three dimensions: home integration, social integration, and productive activity. The discrepancy between the numbers of dimensions hypothesized by researchers versus the number of dimensions identified by factor analysis suggests that researchers are not developing their ideas about dimensionality of participation based upon intercorrelation of activities. Factor analysis identifies subscales by finding sets of items with high intercorrelations while researchers appear to group items based on substantive item content on various characteristics of those items such as where they take place, what sort of activities they involve (e.g., consumptive vs. productive), and who they involve (family vs. acquaintances, for example, in the case of the MSPP).

It's not clear, however, that using a formative approach to identify dimensions matches how researchers and instrument developers think about domains of participation or that it will result in intuitive appealing dimensions but it seems likely to get closer than the factor analysis approach. Activities that are similar in characteristics such as where they take place, whether they are productive or consumptive or purely social and so forth, do seem likely to produce relatively similar outcomes in terms of well-being, social connectedness, and other consequences that might serve to identify a formative model. In this project, the outcomes may have been too far downstream of the activities and too confounded with covariates to distinguish different dimensions among participation activities. Only one index was developed and most participatory items did not significantly contribute to this index when entered with all the other items. But the resultant measure did have the benefit of representing participation unidimensionally. This seems closer to theoretical conceptions of participation than the many-factor models produced by the initial scale construction process.

This project cannot offer a definitive answer to the question of the dimensionality of participation but it did suggest that researchers ought to consider unidimensional models among their candidates. EFA criteria were equivocal as to the number of factors: the scree plot found one factor, parallel analysis identified seven, and the Kaiser criterion pointed to four. But seven factors seems like too many for usefulness, and the factor solutions with many factors (six for men and eight for women) did not provide interesting evidence when embedded into a structural model. The index model functioned well with just one composite and the post-hoc single-factor model for females showed good potential, even though it accounted for only about 20% of variance across the sixteen activity items. Single-index or single-factor approaches have the benefit of allowing researchers to formulate theory about social participation in very general terms rather than in specific terms. This explanatory power is important and is, in fact, one of the reasons the present study was undertaken.

Research question three: Activities to include in definition of social participation. The scale model, in its multi-factor versions, gave little guidance as to what activities should be included in a definition of the construct of social participation; it merely identified correlated items. A researcher might pick one or a few subscales out of the six (male) or eight (female) identified as more representative of social participation than the other subscales, but this would be based not on empirical analysis but on researcher theory and speculation. The index model, on the other hand, allowed for an explicit definition of what social participation means, formulated in terms of its associations with outcomes. The index model represented social participation as those activities that significantly predicted life satisfaction and perceived social connectedness. It suggested which items might be most important to include in an index measuring this construct. Yet the index model seemed inadequate in at least two respects: first, because covariates weren't used during the construction process, the index didn't behave as expected in the validation models that did use covariates; and second, the approach used to select items (using statistical significance only) seemed ad hoc and likely to give non-reproducible results. The first problem, failing to address confounders in the index construction process, was discussed in relation to research question one, above. The second problem, how to select items for a composite, is addressed here.

Selection of regression predictors is an area of ongoing research and controversy (Beyene, Atenafu, Hamid, To, & Sung, 2009). The approach used in this project was somewhat crude: running the initial regression and then selecting any predictors across the three models (entire sample, male respondents only, female respondents only) that had significance at the $p < .10$ level. The predictors selected this way were conceptually appealing in that they spanned the range of subscales identified in factor analysis and they were generally the most broadly written items. However, more sophisticated predictor selection techniques might improve the index model. One algorithmic approach to regression predictor selection, stepwise regression, selects predictors to include by successively adding predictors (forward selection), by successively subtracting them (backward elimination), or by alternating between forward and backward steps using F tests or other criteria for deciding which variables to retain (Tabachnick & Fidell, 2007). Stepwise approaches, however, have been criticized for producing biased parameters, for selecting models that do not predict well outside of the sample in which they were developed, and for failing to control alpha levels (Whittingham, Stephens, Bradbury, & Freckleton, 2006). A variety of alternatives to stepwise regression for predictor selection are available (Flom & Cassell, 2007); future research into index construction might benefit from incorporating such approaches. For example, lasso regression minimizes the residual sum of squares subject to a constraint in order to produce models that are more interpretable and stable than OLS regression using stepwise or other algorithmic methods to select predictors (Tibshirani, 1996).

In any case, the selection of items for an index should also be guided by theoretical considerations. Structural equation modeling techniques allow a researcher to test the fit of an a priori model to empirical data; they can (and usually are) used in a confirmatory rather than exploratory mode. The approach chosen by this project for index construction was more exploratory than confirmatory once the two-composite model was shown to be no better than a one-composite model. However, as index development techniques advance, confirmatory approaches may become more commonly used.

Recommendations for the Use of Formative Models

Unlike reflective measurement models, which use only indicator information to define models, formative measurement models require either reflective indicators or consequential outcome constructs (or a mix of both) to achieve identification. Substantively, what this implies is that a formative construct is in part defined by whatever items or constructs are used to identify the model. This may limit the generality of the derived models, as outcomes of interest in one setting may not be the outcomes of interest in another. If reflective indicators are used alongside formative indicators in order to use a MIMIC-style model, the derived construct would appear to be broadly useful as a generic instrument to be used in measuring the construct of interest. When far-downstream constructs are used to identify the model, however, the index so defined would be useful only in specific, applied settings where the outcome constructs were specifically the ones of interest.

One alternative to a formative model for use with behavioral frequency data is simply summing frequencies of the activities of interest. The benefit of using a formative model over a simple additive score is that the formative model can give some insight into which activities are most important. With more sophisticated specifications, a formative model can suggest whether different activities may interact in defining the construct of interest by predicting the outcomes used to identify the construct. However, this is only useful if there are reflective indicators or outcome constructs available that appropriately delineate the construct. To the extent that reflective indicators or outcome constructs can select items according to the researcher's definition of a construct, a formative model may provide some useful empirical evidence about exactly which activities should be included in the operationalization of the construct and how they should be combined. To the extent, however, that the definition is less than accurately defined (as here, using only relatively downstream outcomes such as perceived social connectedness and life satisfaction), formative approaches may not be so useful in defining index measures. In that case, a researcher may instead consider simply choosing and combining items using theoretical considerations.

Limitations of the Study

In chapter one, four delimitations were discussed. Those delimitations are summarized here and then additional limitations are discussed. First, this was an observational study and as such, it gives no firm ground upon which to infer causality. It is likely that the relationships between participation and the outcomes work in both directions. However, the focus here was to demonstrate and compare practical scale and index construction techniques, not to make judgments about the true causal relationships that hold among participation, social connectedness and life satisfaction. Second, this study was based upon secondary data analysis. The participation items were taken as is, not custom-developed based upon a particular theoretical understanding of participation or its dimensions. There were not any subjective participation items available that might have been used to construct a MIMIC-style model to better identify the index version of social participation. The items didn't appear to have been designed from a common factor model perspective. However, often researchers are faced with a set of pre-existing items and they must construct some sort of model or measure from it, so this project mirrored problems that researchers often face. Third, formative models for the outcome constructs were not considered even though arguments could be made that they would be better modeled that way. Finally, it was noted that formative constructs cannot easily be incorporated as endogenous variables in a model, since different items participating in the construct might very well have different antecedents. This limitation became clearer during the analysis process, when it was found that certain covariates predicted some of the participation items but not others.

In addition to these four delimitations already identified in Chapter I, a number of additional limitations became clear during the analysis and interpretation stages of the project. Five in particular deserve mention: first, the problem of when to use covariates; second, the issue of how to choose index items; third, the use of approximate fit indexes to evaluate models; fourth, the specification of measurement error as a fixed percentage of indicator variance in the index model; and fifth, the dependence of index model results on choice of outcome. The issue of using covariates in index construction was discussed above in the context of the findings for research question one. It was not within scope of this study to consider how best to incorporate covariates; index construction was implemented in a way to make it parallel to scale construction, which, to this author's knowledge, rarely incorporates control variables. The issue of selection of index items was discussed in the context of the findings for research question three. The purpose of the study was not to develop sophisticated index construction techniques using the latest approaches to regression predictor selection but rather to demonstrate the feasibility, at a relatively simple level, of using formative indicators to construct indexes.

The complications arising from the use of approximate fit indexes to evaluate and select models became most obvious during the scale construction process. Chi-square difference tests along with rule-of-thumb cutoffs for RMSEA and CFI values were used in a kind of automated algorithmic model selection process, somewhat similar in flavor to stepwise regression or perhaps most similar to hierarchical regression in which the researcher decides which regression predictors to enter at which step in the process. Unfortunately, the search for an adequate CFI value led to models with too many factors for usefulness (eight in the case of females, including two single-item factors). SEM theorists have pointed out drawbacks of using approximate fit indexes to evaluate models (e.g., Barrett, 2007; McIntosh, 2007). Interestingly, Barrett (2007) called for researchers to eliminate their dependence on the use of approximate fit indexes by instead selecting models based on predictive accuracy:

> If the SEM model includes real world or ‘‘measurable’’ criterion classes or outcomes of one form or another, then strategies for determining cross-validated predictive accuracy and model parsimony via AIC/BIC indexes might prove most effective. Here the argument is that ‘‘good enough/empirically adequate for theory-purposes or pragmatic use’’ multi-facet cross-validated predictive accuracy might supersede the use of a single global statistical discrepancy test. (Barrett, 2007, p. 822)

The approach Barrett described agrees with the approach deployed during the model comparison analyses in the present study. Greater attention to predictive accuracy and what Barrett calls “empirical adequacy” rather than a focus on fit index values would lead away from the highly complex initial factor models, perhaps toward a single-factor model or perhaps toward a two- or three-dimensional model that closer corresponds to researcher’s theoretical conceptions about the dimensions of participation. It’s notable that the single-factor model for females, once trimmed, had quite good fit index values so a focus on predictive accuracy and empirical adequacy doesn’t necessarily mean poor fit. It does mean placing theoretical and explanatory concerns higher on the priority list.

The fourth limitation to be addressed is the specification of residual variance in indicators in the index value as a fixed ten percent of observed indicator variance. One of the main benefits of structural equation modeling is that it can account for measurement error by estimating it from the data at hand. This approach is only accurate, however, if the model used to estimate the error – the common factor model – actually represents the data. So while the scale models estimated measurement error using multiple indicators, this did not necessarily produce a more accurate estimate of measurement error than the approach taken in the index construction models. In the most basic unidimensional scale models, error variances are modeled as uncorrelated but in actuality the residuals of the participation items in the scale model were likely correlated due to participant differences such as age or health status. These do not represent unique variance but may differentially affect groups of participation items. For example, years of education may predict higher computer usage and higher educational participation, but these individual indicators were not specified with correlated residuals. The measurement error model of the scale models may have been incorrectly specified, in other words. Without additional extensive analysis, it is not clear whether the fixed residual variance approach or the multiple-indicator approach to estimating measurement error was a more realistic representation of the participation indicator data.

Finally, the use of index models for defining constructs such as social participation produces construct definitions that are specific to the outcome constructs chosen to identify the model. In the present case, perceived social connectedness and life satisfaction were the outcomes used to identify or form the social participation model. Participation items were selected by the analysis to the extent that they were predictive of these two outcomes. Had different outcome constructs been selected, the participation items included in the definition might have changed entirely. This limitation of the index approach means that constructs defined by formative measurement are only useful insofar as the outcome constructs used to identify the model are specifically the ones of interest in a certain application. Constructs defined formatively may be less generically useful than constructs defined in some more generalizable way, without the contextualization that formative measurement requires.

Recommendations for Future Study

The suggestions for future study come under two headings: those directed towards better measurement and modeling of social participation and those directed towards the development of index construction techniques based on formative measurement.

Measurement of participation. It would be useful to explore whether MIMIC-style models such as that shown in Figure 36 provide a better means of modeling participation. Researchers who wish to approach measurement of participation from a formative perspective using objective responses indicating frequency of participation in different activities should consider adding subjective items that might allow MIMIC-style identification of their index models. Such items would allow a researcher to explicitly and empirically define different dimensions of participation based upon theory about what perceptions a person might have about participation of different types. They would avoid the underidentification and interpretational confounding problems that arise when correlated outcomes must be used to define formative measures.

As previously mentioned, this study used a very simple regression model in defining its participation composite, but more sophisticated models that express theoretical notions of how participation works to improve outcomes such as well-being could be developed and tested. Models could include terms that capture diversity of roles, reflecting the theoretical notion that balance across roles is as or more important than absolute levels of participation. They could add interaction terms or other nonlinearities that might better predict well-being and social connectedness outcomes than first-order linear models. A benefit of the formative approach over the factor analysis approach is that it can build upon the rich and powerful techniques available in regression modeling. Researchers modeling participation should take advantage of these techniques.

The one-factor model of social participation for females considered at the end suggests an additional direction of study: towards whether existing participation instrument data might be usefully modeled with just one dimension rather than the three or four that the instrument developers hypothesized. Alternatively, researchers might consider second-order factor models with a general participation factor specialized into dimensions of participation. I suggested earlier that index models might be improved by considering more sophisticated mathematical models that used interaction terms or additional predictors such as count of roles to capture the full richness of our theoretical conceptions of participation. More sophisticated factor models built up around a unidimensional core of participation could perhaps lead towards the valid and reliable participation scales and subscales that have so far eluded us.

Advancing index construction techniques. Social participation is just one construct used in psychosocial research that might usefully be modeled formatively. Some other constructs that might be explored formatively include socio-economic status (measured, for example, by education levels, income, occupational prestige, and home location), employee job satisfaction (measured by satisfaction with different aspects of one's job), objective social connectedness (measured by number and quality of contacts), and quality of life (measured by summing measures of quality of various aspects of a person's life). Use of index construction techniques such as those employed in this project might at the same time shed light on those constructs and their association with other constructs of interest while showing the usefulness of formative approaches in different domains.

It may also be useful for researchers interested in formative measurement to undertake simulation studies to investigate under what conditions different dimensions might be uncovered using formative techniques. This study hypothesized two dimensions of participation in the sixteen activity items, but found only one. Is this because the techniques used were too crude to identify different patterns of association with outcomes that held for community participation items versus domestic participation items? Or were there, in fact, no underlying predictive dimensions – did the different kinds of items act similarly enough on outcomes that there was only one dimension to find? Might dimensions have been uncovered if subjective, reflective items were added as in the MIMIC-style model suggested earlier? A simulation study could demonstrate the power or lack thereof of composite modeling to identify useful predictive dimensions among a set of items.

Conclusion

Some constructs that researchers wish to model at a summary level do not fit neatly into the common factor model approach of classical psychometrics. Formative measurement in which items sum together in some way to create a high-level measure of the quantity of interest has been proposed as an alternative to CTT-based measurement for these cases. But formative approaches have been criticized on philosophical and practical grounds and they have not been rigorously applied to index construction. This study compared reflective (scale model) and formative (index model) approaches to modeling social participation in the context of its associations with life satisfaction and perceived social connectedness. Results did not unequivocally point to the scale or index approach as superior, though the initial index construction process produced a more parsimonious model of social participation with less equivocation than the scale construction process offered. The scale model was not parsimonious, did not show invariance across gender, and did not give results conforming to theory and past research when embedded into a structural model that included the outcomes of life satisfaction and perceived social connectedness. The index model identified one theoretically plausible index of participation that, at least in the female subsample, showed hypothesized associations with outcomes and in the male subsample offered evidence for an alternative model in which the relationship between life satisfaction and social participation is fully mediated by participation's effects on social connectedness. While much work remains to be done to make index techniques practical and rigorous, these results suggest that they may allow a way forward for modeling social participation and other important constructs in psychosocial and health research that do not conform to the common factor model. Perhaps more importantly, this study suggested that unidimensional models of participation show merit and should be further explored by participation instrument developers and other researchers wanting to incorporate models of participation into their analyses.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423.

Aquino, J. A., Russell, D. W., Cutrona, C. E., & Altmaier, E. M. (1996). Employment status, social support, and life satisfaction among the elderly. *Journal of Counseling Psychology, 43*(4), 480-489.

Bagozzi, R. P. (2007). On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilcox (2007). *Psychological methods*, *12*(2), 229-237.

Baker, L. A, Cahalin, L. P., Gerst, K., & Burr, J. A. (2005). Productive activities and subjective well-being among older adults: The influence of number of activities and time commitment. *Social Indicators Research*, *73*(3), 431-458.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815-824.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606.

Berkman, L. F., Berkman, C. S., Kasl, S., Freeman Jr., D. H., Leo, L., Ostfeld, A. M., Cornoni-Huntley, J., & Brody, J. A. (1986). Depressive symptoms in relation to physical health and functioning in the elderly. *American Journal of Epidemiology, 124*(3), 372-388.

Beyene, J., Atenafu, E. G., Hamid, J. S., To, T., & Sung, L. (2009). Determining relative importance of variables in developing and validating predictive models. *BMC Medical Research Methodology, 9*(64).

Blalock, H. M. (1964). *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

Bogner, J. A., Whiteneck, G., Corrigan, J. D., Lai, J.-S., Dijkers, M. P., & Heinemann, A. W. (2011). Comparison of scoring methods for the Participation Assessment with Recombined Tools-Objective. *Archives of Physical Medicine and Rehabilitation, 92*(4), 552-563.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605-634.

Bollen, K. A. (2007). Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods, 12*(2), 219-228.

Bollen, K. A., & Davis, W. R. (2009). Causal indicator models: Identification, estimation, and testing. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 498-522. (Reprinted from Bollen, K. A., & Davis, W. R. (1994). Causal indicator models: Identification, estimation, and testing. Paper presented at the 1993 American Sociological Association Convention, Miami, FL.)

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305-314.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* Mahwah, N.J: Lawrence Erlbaum Associates.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110*(2), 203-219.

Brault, M. W. (2008). Americans with Disabilities: 2005 (Current Population Reports, P70-117). Washington, DC: U.S. Census Bureau.

Brock, J. K.-U., & Zhou, Y. (2005). Organizational use of the internet: Scale development and validation. *Internet Research*, *15*(1), 67-87.

Brown, M. (2006). Participation Objective, Participation Subjective. San Jose, CA: The Center for Outcome Measurement in Brain Injury. Retrieved from http://www.tbims.org/combi/pops.

Brown, M., Dijkers, M. P., Gordon, W. A., Ashman, T., Charatz, H., & Cheng, Z. (2004). Participation objective, participation subjective: A measure of participation combining outsider and insider perspectives. *The Journal of Head Trauma Rehabilitation*, *19*(6), 459-481.

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Newbury Park, CA: Sage.

Bruhn, M., Georgi, D., & Hadwich, K. (2008). Customer equity management as formative second-order construct. *Journal of Business Research*, *61*(12), 1292-1301.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods & Research*, *5*(1), 3-52.

Busemeyer, J., & Wang, Y. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, *44*(1), 171-189.

Cadogan, J. W. & Lee, N. (in press). Improper use of endogenous formative variables. *Journal of Business Research.* Retrieved from http://www.research-training.com/index_htm_files/NEW_CADOGAN_LEE_15_OCT_2010.pdf

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

Chi-square difference testing using the Satorra-Bentler scaled chi-square. (n.d.). Retrieved August 22, 2011 from http://www.statmodel.com/chidiff.shtml.

Chin, W. W., Peterson, R. A., & Brown, S. P. (2008). Structural equation modeling in marketing: Some practical reminders. *The Journal of Marketing Theory and Practice*, *16*(4), 287-298.

Clarke, P., Fisher, G., House, J., Smith, J., & Weir, D. (2008). Guide to content of the HRS psychosocial leave-behind participant lifestyle questionnaires: 2004 & 2006. Retrieved from http://hrsonline.isr.umich.edu/sitedocs/userg/HRS2006LBQscale.pdf

Collier, J. E. (2006). Measuring service quality in e-retailing. *Journal of Service Research*, *8*(3), 260-275.

Cornwell, B., Laumann, E. O., & Schumm, L. P. (2008). The social connectedness of older adults: A national profile. *American Sociological Review*, *73*(2), 185-203.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of a test. *Psychometrika, 16*, 297-334.

de Jong Gierveld, J., & Kamphuls, F. (1985). The development of a Rasch-type loneliness scale. *Applied Psychological Measurement*, *9*(3), 289-299.

de Jong Gierveld, J., & van Tilburg, T. (2006). A 6-item scale for overall, emotional, and social loneliness: Confirmatory tests on survey data. *Research on Aging*, *28*(5), 582-598.

Dench, S., & Regan, J. (2000). *Learning in Later Life: Motivation and Impact*. Institute for Employment Studies. Report commissioned by the Department for Education and Employment. Research Report Number 183. Retrieved from http://www.employment-studies.co.uk/summary/summary.php?id=rr183

Diamantopoulos, A., Riefler, P., & Roth, K. (2008). Advancing formative measurement models. *Journal of Business Research, 61*(12), 1203-1218.

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, *17*(4), 263-282.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*(2), 269-277.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71-75.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*(2), 276-302.

Dijkers, M. P. (2010). Issues in the conceptualization and measurement of participation: An overview. *Archives of Physical Medicine and Rehabilitation, 91*(9 Suppl), S5-16.

Edwards, J. R. (2010). The fallacy of formative measurement. *Organizational Research Methods, 14*(2), 370-388.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*(2), 155-174.

Flom, P. L., & Cassell, D. L. (2007). Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. *NESUG 2007 Proceedings.* Retrieved from http://www.nesug.org/Proceedings/nesug07/sa/sa07.pdf

Forster, M. R. (2000). Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, *44*(1), 205-231.

Franke, G., Preacher, K., & Rigdon, E. (2008). Proportional structural effects of formative indicators. *Journal of Business Research*, *61*(12), 1229-1237.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

Grace, J.B., & Bollen, K.A. (2006). The interface between theory and data in structural equation models. USGS Open-File Report 2006-1363. U.S. Geological Survey: Reston, VA.

Grace, J. B., & Bollen, K. A. (2008). Representing general theoretical concepts in structural equation models: The role of composite variables. *Environmental and Ecological Statistics*, *15*(2), 191-213.

Hao, Y. (2008). Productive activities and psychological well-being among older adults. *Journal of Gerontology: Social Sciences, 63B*(2), S64-72.

Harlow, R. E., & Cantor, N. (1996). Still participating after all these years: A study of life task participation in later life. *Journal of Personality and Social Psychology, 71*(6), 1235-1249.

Headey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*(4), 731-739.

Health and Retirement Study. (2010). *Health and Retirement Study: Tracker 2008 Final, Version 1.0*. Retrieved from http://hrsonline.isr.umich.edu/modules/meta/tracker/desc/trk2008.pdf

Herero, V. G., & Extremera, N. (2010). Daily life activities as mediators of the relationship between personality variables and subjective well-being among older adults. *Personality and Individual Differences, 49*(2), 124-129.

Hinterlong, J. E., Morrow-Howell, N., & Rozario, P. A. (2007). Productive engagement and late life physical and mental health: Findings from a nationally representative panel study. *Research on Aging, 29*(4), 348-370.

Holbrook, M. & Skilbeck, C. E. (1983). An activities index for use with stroke patients. *Age and Aging 12*(2), 166-170.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Pyschometrika, 30*, 179-185.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods*, *12*(2), 205-218.

Hookway, C. (2010). Pragmatism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition).* Retrieved from http://plato.stanford.edu/archives/spr2010/entries/pragmatism/.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199-218.

Jette, A. M., Haley, S. M., & Kooyoomjian, J. T. (2003). Are the ICF activity and participation dimensions distinct? *Journal of Rehabilitation Medicine*, *35*(3), 145-149.

Johnson, G., Bruner, G., & Kumar, A. (2006). Interactivity and its facets revised: Theory and empirical test. *Journal of Advertising*, *35*(4), 35-52.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, *70*(351), 631–639.

Jung, Y., Gruenewald, T. L., Seeman, T. E., & Sarkisian, C. A. (2010). Productive activities and development of frailty in older adults. *Journal of Gerontology: Social Sciences*, *65B*(2), 256-261.

Kabacoff, R. (2002). Determining the dimensionality of data: A SAS® macro for parallel analysis. *Proceedings of the 27th Annual Meeting of the SAS Users Group International*. Cary, NC: SAS Institute.

Kenny, D. A. (2011, September 4). Measuring model fit. Retrieved from http://www.davidakenny.net/cm/fit.htm

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press.

Leacock, C., Ed. (2006). *Getting Started with the Health and Retirement Study, Version 1.0.* Ann Arbor, MI: University of Michigan. Retrieved from http://hrsonline.isr.umich.edu/docs/dmgt/

Lemon, B.W., Bengtson, V. L., & Peterson, J. A. (1972). An exploration of the activity theory of aging: Activity types and life satisfaction among in-movers to a retirement community. *Journal of Gerontology*, *27*(5), 511-523.

Levasseur, M., Desrosiers, J., & Noreau, L. (2004). Is social participation associated with quality of life of older adults with physical disabilities? *Disability and Rehabilitation*, *26*(20), 1206-1213.

Lin, C.-H., Sher, P. J., & Shih, H.-Y. (2005). Past progress and future directions in conceptualizing customer perceived value. *International Journal of Service Industry Management*, *16*(4), 318-336.

Lubben, J. & Gironda, M. (2003). Centrality of social ties to the health and well-being of older adults. In Berkman, B., & Harootyan, L. K., (Eds.) *Social Work and Health Care in an Aging Society: Education, Policy, Practice, and Research*. New York: Springer.

MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *The Journal of Applied Psychology, 90*(4), 710-730.

Mars, G. M. J., Kempen, G. I. J. M., Post, M. W. M., Proot, I. M., Mesters, I., & van Eijk, J. T. M. (2009). The Maastricht social participation profile: Development and clinimetric properties in older adults with a chronic physical illness. *Quality of Life Research 8*(9), 1207-1218.

McIntosh, C. (2007). Rethinking fit assessment in structural equation modeling: A commentary and elaboration on Barrett (2007). *Personality and Individual Differences*, *42*(5), 859-867.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied Multivariate Research: Design and Interpretation*. Thousand Oaks: Sage Publications.

Missing Data Modeling. (n.d.). Retrieved from http://www.statmodel.com/discussion/messages/22/22.html.

Morrow-Howell, N. (2010). Volunteering in later life: Research frontiers. *Journal of Gerontology: Social Sciences*, *65B*(4), 461-469.

Mulaik, S., & Millsap, R. (2000). Doing the four-step right. *Structural Equation Modeling: A Multidisciplinary Journal*, *7*(1), 36-73.

Muthén, B. O. (2006). Natural log transformation in growth model (discussion board posting). Retrieved from http://www.statmodel.com/discussion/messages/14/1081.html?1255823782

Muthén, L.K. & Muthén, B.O. (2010). *Mplus User's Guide. Sixth Edition.* Los Angeles, CA: Muthén & Muthén.

Nagi, S. (1991). Disability concepts revisited: Implications for prevention. In Pope, A. M. & Tarlov, A. R. (Eds.). *Disability in America: Towards a National Agenda for Prevention*. Washington, D.C.: National Academies Press.

Newcomb, M. D., & Bentler, P. M. (1986). Loneliness and social support: A confirmatory hierarchical analysis. *Personality and Social Psychology Bulletin*, *12*(4), 520-535.

Newsom, J. T., & Schulz, R. (1996). Social support as a mediator in the relation between functional status and quality of life in older adults. *Psychology and Aging, 11*(1), 34-44.

Newall, N. E., Chipperfield, J. G., Clifton, R. A, Perry, R. P., Swift, A. U., & Ruthig, J. C. (2009). Causal beliefs, social participation, and loneliness among older adults: A longitudinal study. *Journal of Social and Personal Relationships*, *26*(2-3), 273-290.

Noreau, L., Desrosiers, J., Robichaud, L., Fougeyrollas, P., Rochette, A., & Viscogliosi, C. (2004). Measuring social participation: Reliability of the LIFE-H in older adults with disabilities. *Disability and Rehabilitation, 26*(6), 346-352.

Pavlou, P. A., & Gefen, D. (2005). Psychological contract violation in online marketplaces: Antecedents, consequences, and moderating role. *Information Systems Research*, *16*(4), 372-399.

Peirce, C.S. (1878). Illustrations of the logic of science: How to make our ideas clear. *Popular Science Monthly 12*, 286-302.

Peplau, L. A., & Perlman, D. (1982). Perspectives on loneliness. In L. A. Peplau & D. Perlman (Eds.), *Loneliness: A Sourcebook of Current Theory, Research, and Therapy* (pp. 1-20). New York: Wiley-Interscience.

Post, M. W. M., de Witte, L. P., Reichrath, E., Verdonschot, M. M., Wijlhuizen, G. J., & Perenboom, R. J. M. (2008). Development and validation of IMPACT-S, an ICF-based questionnaire to measure activities and participation. *Journal of Rehabilitation Medicine*, *40*(8), 620-627.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385-401.

Raftery, A. E. (1993). Bayesian model selection in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 163-180). Newbury Park, CA: Sage.

RAND Center for the Study of Aging. (2011). *Data Products*. Retrieved from http://www.rand.org/labor/aging/dataprod.html.

Reistetter, T. A., Spencer, J. C., Trujillo, L., & Abreu, B. C. (2005). Examining the Community Integration Measure (CIM): A replication study with life satisfaction. *NeuroRehabilitation*, *20*(2), 139-148.

Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *The American Psychologist*, *65*(1), 1-12.

Rowe, J. W., & Kahn, R. L. (1997). Successful aging. *The Gerontologist, 37*(4), 433-440.

Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment 66,* 20-40.

Sander, A. M., Fuchs, K. L., High, W. M., Hall, K. M., Kreutzer, J. S., & Rosenthal, M. (1999). The Community Integration Questionnaire revisited: An assessment of factor structure and validity. *Archives of Physical Medicine and Rehabilitation, 80*(10), 1303-1308.

Sarason, B. R., Levine, H. M. Basham, R. B., & Sarason, B. R. (1983). Assessing social support: The social support questionnaire. *Journal of Personality and Social Psychology, 44*(1), 127-139.

Schuling, J., de Haan, R., Limburg, M., & Groenier, K. H. (1993). The Frenchay Activities Index: Assessment of functional status in stroke patients. *Stroke, 24*(8), 1173-1177.

St. Clair, P., Blake, D., Bugliari, D., Chien, S., Hayden, O., Hurd, M., Ilchuk, S., … Zissimopoulos, J. (2010). *Rand HRS Data Documentation, Version J*. Santa Monica, CA: RAND Center for the Study of Aging.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677-680.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics (5th Edition).* Boston: Pearson/Allyn & Bacon.

Thompson, B. (1984). *Canonical Correlation Analysis: Uses and Interpretation.* Thousand Oaks, CA: Sage Publications, Quantitative Applications in the Social Sciences Series, No. 47.

Thurstone, L. L. (1947). *Multiple-Factor Analysis.* Chicago: University of Chicago Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Societ, 58*(1), 267-288.

Veroff, J., & Veroff, J. B. (1980). *Social incentives: A life-span developmental approach*. New York: Academic Press

Wahrendorf, M., Ribet, C., Zins, M., & Siegrist, J. (2008). Social productivity and depressive symptoms in early old age: Results from the GAZEL study. *Aging & Mental Health, 12*(3), 310-316.

Warr, P., Butcher, V., & Robertson, I. (2004). Activity and psychological well-being in older people. *Aging and Mental Health, 8*(2), 172-83.

Whiteneck, G. (2010). Issues affecting the selection of participation measurement in outcomes research and clinical trials. *Archives of Physical Medicine and Rehabilitation*, *91*(9 Suppl), S54-9.

Whiteneck, G., & Dijkers, M. P. (2009). Difficult to measure constructs: Conceptual and methodological issues concerning participation and environmental factors. *Archive of Physical Medicine and Rehabilitation, 90*(11 Suppl), S22-35.

Whittingham, M. J., Stephens, P. A., Bradbury, R. B., & Freckleton, R. P. (2006). Why do we still use stepwise modeling in ecology and behavior? *Journal of Animal Ecology*, *75*, 1182-1189.

Wilkie, R., Peat, G., Thomas, E., Hooper, H., & Croft, P. R. (2005). The Keele Assessment of Participation: A new instrument to measure participation restriction in population studies. Combined qualitative and quantitative examination of its psychometric properties. *Quality of Life Research*, *14*(8), 1889-1899.

Willer, B., Rosenthal, J.S., Kreutzer, W.A., Gordon, W.A. and Rempel, R. (1993) Assessment of community integration following rehabilitation for traumatic brain injury. *Journal of Head Trauma Rehabilitation 8*(2), 75-87.

Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach (4th Edition).* Mason, Ohio: South-Western Cengage Learning.

Zucchini, W. (2000). An introduction to model selection. *Journal of Mathematical Psychology*, *44*(1), 41-61.

Appendices

Appendix A: Descriptive Statistics

Table A.1

Percentages of respondents reporting different levels of participation in 16 activities, ordered by percentage reporting daily participation

	Daily	***Several times a week***	***Once a week***	***Several times a month***	***At least once a month***	***Not in the last month***	***Missing***
sp6. Read	68	14	4	3	2	5	3
sp10. Computer	24	8	2	2	2	52	9
sp7. Word games	21	9	5	6	6	48	6
sp16. Walk	20	22	8	10	8	28	4
sp11. Home maintenance/ gardening	18	19	11	11	10	26	5
sp15. Sports/exercise	14	20	7	8	6	39	6
sp14. Hobby	11	16	7	10	11	38	7
sp12. Bake or cook	10	15	13	12	14	31	4
sp8. Cards/chess/other games	6	8	6	7	9	58	7
sp9. Writing	5	7	5	8	13	56	6
sp13. Sew or knit	3	4	1	4	4	77	8
sp1. Volunteer with youth	2	2	3	2	4	80	8
sp2. Volunteer – other	1	5	6	6	9	64	8
sp4. Sports/social club	1	7	6	7	13	58	8
sp3. Education	0	1	2	1	5	81	9
sp5. Non-religious organizations	0	1	3	4	12	72	8

Table A.2

Percentages of respondents reporting different levels of participation in 16 activities, by gender

		Daily	***Several times a week***	***Once a week***	***Several times a month***	***At least once a month***	***Not in the last month***	***Missing***
sp1. Volunteer with youth	male	1.2	1.2	2.2	1.9	4.1	***83.3***	6.0
	female	2.0	2.4	3.1	2.3	3.9	***77.1***	9.3
	difference	-0.8	-1.2	-0.9	-0.4	0.2	***6.2***	-3.3
sp2. Volunteer – other	male	1.5	4.6	5.8	6.8	8.9	66.4	5.9
	female	1.4	4.9	6.5	6.2	9.2	63.1	8.8
	difference	0.1	-0.3	-0.7	0.6	-0.3	3.3	-2.9
sp3. Education	male	0.4	0.8	2.0	1.3	5.8	83.1	6.7
	female	0.6	1.3	2.2	1.5	4.6	79.4	10.5
	difference	-0.2	-0.5	-0.2	-0.2	1.2	3.7	-3.8
sp4. Sports/social club	male	1.7	7.7	7.7	6.7	14.2	56.2	5.8
	female	0.7	6.4	5.3	7.3	12.3	58.7	9.4
	difference	1.0	1.3	2.4	-0.6	1.9	-2.5	-3.6
sp5. Non-religious organizations	male	0.3	1.3	2.4	3.7	12.4	74.2	5.7
	female	0.4	1.3	2.7	3.8	11.5	71.2	9.1
	difference	-0.1	0.0	-0.3	-0.1	0.9	3.0	-3.4
sp6. Read	male	66.1	13.1	4.7	4.2	2.6	6.3	2.9
	female	68.8	15.5	3.6	2.8	1.6	4.8	2.9
	difference	-2.7	-2.4	1.1	1.4	1.0	1.5	0.0
sp7. Word games	male	***13.7***	***5.4***	3.8	3.8	6.5	***61.8***	5.0
	female	***25.7***	***11.5***	5.7	7.7	5.4	***38.0***	6.0
	difference	***-12.0***	***-6.1***	-1.9	-3.9	1.1	***23.8***	-1.0
sp8. Cards/chess/other games	male	4.8	7.5	5.2	7.0	8.6	***61.6***	5.4
	female	6.9	7.8	5.8	6.4	9.8	***54.9***	8.3
	difference	-2.1	-0.3	-0.6	0.6	-1.2	***6.7***	-2.9
sp9. Writing	male	3.7	5.7	3.3	***5.4***	10.0	***66.1***	5.7
	female	5.9	7.3	6.3	***10.5***	14.5	***48.8***	6.8
	difference	-2.2	-1.6	-3.0	***-5.1***	-4.5	***17.3***	-1.1
sp10. Computer	male	26.8	7.9	2.1	2.1	2.8	51.9	6.3
	female	22.3	8.1	2.6	2.5	2.3	52.0	10.2
	difference	4.5	-0.2	-0.5	-0.4	0.5	-0.1	-3.9

sp11. Home maintenance/ gardening	male	16.2	***23.0***	13.2	12.4	12.0	***19.9***	3.4
	female	18.5	***16.5***	9.6	9.9	8.7	***30.9***	5.9
	difference	-2.3	***6.5***	3.6	2.5	3.3	***-11.0***	-2.5
sp12. Bake or cook	male	***5.6***	***9.1***	***8.2***	***7.9***	***11.3***	***53.0***	5.0
	female	***13.2***	***19.9***	***15.8***	***14.4***	***16.6***	***16.0***	4.0
	difference	***-7.6***	***-10.8***	***-7.6***	***-6.5***	***-5.3***	***37.0***	1.0
sp13. Sew or knit	male	0.3	***0.4***	0.1	***0.3***	1.0	***90.0***	7.9
	female	4.9	***6.0***	2.1	***5.9***	5.5	***67.0***	8.5
	difference	-4.6	***-5.6***	-2.0	***-5.6***	-4.5	***23.0***	-0.6
sp14. Hobbies	male	10.9	17.5	8.9	10.2	12.1	35.0	5.4
	female	10.7	15.1	6.3	9.6	10.8	39.6	7.9
	difference	0.2	2.4	2.6	0.6	1.3	-4.6	-2.5
sp15. Sports/exercise	male	15.7	21.9	6.9	8.6	6.5	***35.5***	5.0
	female	12.0	19.1	6.7	7.3	6.0	***41.8***	7.2
	difference	3.7	2.8	0.2	1.3	0.5	***-6.3***	-2.2
sp16. Walk	male	***22.7***	23.0	8.3	9.7	7.7	25.8	2.8
	female	***17.7***	21.7	8.5	9.9	8.1	29.4	4.7
	difference	***5.0***	1.3	-0.2	-0.2	-0.4	-3.6	-1.9

Note: Shaded cells indicate percentage differences in participation across male and female at least five percent or greater

Table A.3

Descriptive statistics for continuous and ordinal variables

	N	*Min*	*Max*	*Mean*	*SD*	*Skew*	*Kurt*
Social participation							
sp1. Volunteer with youth	4002	0	5	0.35	1.03	3.16	9.23
sp2. Volunteer – other	4017	0	5	0.72	1.30	1.71	1.76
sp3. Education	3958	0	5	0.23	0.77	3.92	15.79
sp4. Sports/social club	4003	0	5	0.86	1.35	1.44	0.82
sp5. Non-religious organizations	4011	0	5	0.37	0.85	2.80	8.18
sp6. Read	4220	0	5	4.30	1.37	-2.14	3.51
sp7. Word games	4103	0	5	1.82	2.11	0.53	-1.49
sp8. Cards/chess/other games	4038	0	5	1.08	1.64	1.28	0.16
sp9. Writing	4070	0	5	1.02	1.55	1.37	0.58
sp10. Computer	3972	0	5	1.83	2.24	0.52	-1.62
sp11. Home maintenance/ gardening	4135	0	5	2.41	1.90	-0.03	-1.53
sp12. Bake or cook	4154	0	5	1.96	1.79	0.32	-1.34
sp13. Sew or knit	3987	0	5	0.48	1.24	2.62	5.64
sp14. Hobby	4047	0	5	1.84	1.88	0.44	-1.38
sp15. Sports/Exercise	4073	0	5	2.04	1.99	0.24	-1.63
sp16. Walk 20 minutes	4176	0	5	2.51	1.96	-0.12	-1.59
Life satisfaction							
ls1. Life is close to ideal	4210	1	7	4.85	1.81	-0.72	-0.56
ls2. Conditions of life excellent	4204	1	7	4.82	1.85	-0.68	-0.72
ls3. Satisfied with life	4245	1	7	5.50	1.69	-1.26	0.65
ls4. Have important things to do in life	4242	1	7	5.54	1.63	-1.30	0.87
ls5. Would change nothing	4248	1	7	4.62	2.04	-0.51	-1.11
Social connectedness							
sc1. In tune with others	4199	1	3	1.63	0.67	0.60	-0.69
sc2. Have I can talk to	4229	1	3	1.42	0.59	1.10	0.19
sc3. Have people I can turn to	4237	1	3	1.42	0.59	1.11	0.21
sc4. There are people who understand me	4239	1	3	1.57	0.61	0.55	-0.61
sc5. There are people I feel close to	4231	1	3	1.36	0.56	1.31	0.73
sc6. Feel part of a group	4203	1	3	1.65	0.70	0.61	-0.80
sc7. Have a lot in common with friends	4247	1	3	1.57	0.63	0.62	-0.57
Covariates							
Years education	4335	0	17	12.32	3.13	-0.85	1.64
Age	4346	65	100	74.70	6.97	0.71	-0.17
Self-report of health	4344	1	5	2.93	1.07	0.16	-0.62
Depression score	4317	0	8	1.32	1.85	1.70	2.35

Wealth, household ($100,000s)	4346	-5.5	306.6	5.6	13.1	9.40	134.96
Income, household ($1,000s)	4346	0.00	1,619.7	51.8	70.7	7.89	109.48

Table A.4

Correlations – Life satisfaction items – Both genders

		ls1 Q03A. LIFE IS CLOSE TO IDEAL	***ls2 Q03B. CONDITIONS OF LIFE ARE EXCELLENT***	***ls3 Q03C. SATISFIED WITH LIFE***	***ls4 Q03D. HAVE IMPORTANT THINGS IN LIFE***	***ls5 Q03E. CHANGE NOTHING IF LIVED LIFE OVER***
ls1 Q03A. LIFE IS CLOSE TO IDEAL	*Pearson Correlation*	1	.740**	.632**	.503**	.450**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001
	N	4210	4162	4182	4182	4182
ls2 Q03B. CONDITIONS OF LIFE ARE EXCELLENT	*Pearson Correlation*	.740**	1	.710**	.543**	.460**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001
	N	4162	4204	4181	4185	4181
ls3 Q03C. SATISFIED WITH LIFE	*Pearson Correlation*	.632**	.710**	1	.627**	.482**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001
	N	4182	4181	4245	4219	4219
ls4 Q03D. HAVE IMPORTANT THINGS IN LIFE	*Pearson Correlation*	.503**	.543**	.627**	1	.511**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001
	N	4182	4185	4219	4242	4221
ls5 Q03E. CHANGE NOTHING IF LIVED LIFE OVER	*Pearson Correlation*	.450**	.460**	.482**	.511**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	
	N	4182	4181	4219	4221	4248

**. Correlation is significant at the 0.01 level (2-tailed).

Table A.5

Correlations – Perceived social connectedness items – Both genders

		sc1	sc2	sc3	sc4	sc5	sc6	sc7
sc1 Q20D. IN TUNE WITH OTHERS	*Pearson Correlation*	1	.365**	.355**	.357**	.345**	.334**	.380**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001	<.001	<.001
	N	4199	4171	4179	4180	4172	4154	4185
sc2 Q20F. PEOPLE CAN TALK TO	*Pearson Correlation*	.365**	1	.719**	.562**	.553**	.441**	.457**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001	<.001	<.001
	N	4171	4229	4209	4209	4200	4174	4214
sc3 Q20G. PEOPLE CAN TURN TO	*Pearson Correlation*	.355**	.719**	1	.603**	.616**	.465**	.478**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001	<.001	<.001
	N	4179	4209	4237	4219	4211	4184	4222
sc4 Q20H. PEOPLE UNDERSTAND YOU	*Pearson Correlation*	.357**	.562**	.603**	1	.605**	.494**	.482**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001	<.001	<.001
	N	4180	4209	4219	4239	4211	4188	4225
sc5 Q20I. PEOPLE FEEL CLOSE TO	*Pearson Correlation*	.345**	.553**	.616**	.605**	1	.507**	.509**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001		<.001	<.001
	N	4172	4200	4211	4211	4231	4180	4221
sc6 Q20J. FEEL PART OF GROUP	*Pearson Correlation*	.334**	.441**	.465**	.494**	.507**	1	.608**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001		<.001
	N	4154	4174	4184	4188	4180	4203	4198
sc7 Q20K. A LOT IN COMMON WITH FRIENDS	*Pearson Correlation*	.380**	.457**	.478**	.482**	.509**	.608**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	
	N	4185	4214	4222	4225	4221	4198	4247

**. Correlation is significant at the 0.01 level (2-tailed).

Table A.6

Correlations – Social participation items – Male respondents

		BI	sp2	sp3	sp4	sp5	sp6	sp7	sp8	sp9	sp10	sp11	sp12	sp13	sp14	sp15	sp16
sp1 Q1B. OFTEN VOLUNTEER YOUTH	*Pearson Correlation*	1	.265**	.212**	.091**	.225**	.021	.058*	.097**	.105**	.019	.094**	.129**	.117**	.057*	.060*	.092**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001	.390	.018	<.001	<.001	.444	<.001	<.001	<.001	.021	.015	<.001
	N	1703	1688	1683	1683	1689	1695	1692	1683	1687	1675	1693	1681	1660	1657	1669	1689
sp2 Q1C. OFTEN VOLUNTEER - OTHER	*Pearson Correlation*	.265**	1	.280**	.182**	.341**	.126**	.030	.049*	.216**	.208**	.125**	.051*	.083**	.171**	.159**	.094**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001	<.001	.210	.045	<.001	<.001	<.001	.035	.001	<.001	<.001	<.001
	N	1688	1705	1680	1680	1687	1697	1691	1681	1684	1676	1694	1683	1656	1658	1670	1688
sp3 Q1D. OFTEN EDUCATION	*Pearson Correlation*	.212**	.280**	1	.201**	.280**	.076**	.058*	.076**	.222**	.144**	.047	.090**	.171**	.124**	.171**	.117**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001	.002	.018	.002	<.001	<.001	.054	<.001	<.001	<.001	<.001	<.001
	N	1683	1680	1691	1671	1682	1684	1681	1674	1680	1668	1682	1673	1654	1647	1660	1676
sp4 Q1E. OFTEN ATTEND SPORTS/SOCIAL/CLUB	*Pearson Correlation*	.091**	.182**	.201**	1	.276**	.196**	.113**	.164**	.151**	.153**	.080**	.068**	.076**	.095**	.298**	.116**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001	<.001	.001	.005	.002	<.001	<.001	<.001
	N	1683	1680	1671	1707	1683	1696	1686	1683	1681	1669	1694	1679	1649	1655	1669	1687
sp5 Q1F. OFTEN ATTEND NON RELIGIOUS ORGS	*Pearson Correlation*	.225**	.341**	.280**	.276**	1	.088**	.059*	.087**	.232**	.176**	.060*	.089**	.151**	.089**	.121**	.108**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001		<.001	.016	<.001	<.001	<.001	.013	<.001	<.001	<.001	<.001	<.001
	N	1689	1687	1682	1683	1708	1699	1695	1689	1689	1680	1696	1686	1661	1661	1673	1686
sp6 Q1H. OFTEN READ	*Pearson Correlation*	.021	.126**	.076**	.196**	.088**	1	.212**	.076**	.176**	.192**	.157**	.099**	-.022	.183**	.247**	.150**
	Sig. (2-tailed)	.390	<.001	.002	<.001	<.001		<.001	.002	<.001	<.001	<.001	<.001	.376	<.001	<.001	<.001
	N	1695	1697	1684	1696	1699	1759	1712	1707	1700	1689	1738	1710	1663	1690	1702	1730
sp7 Q1I. OFTEN DO WORD GAMES	*Pearson Correlation*	.058*	.030	.058*	.113**	.059*	.212**	1	.222**	.132**	.152**	.090**	.056*	.042	.147**	.067**	.080**
	Sig. (2-tailed)	.018	.210	.018	<.001	.016	<.001		<.001	<.001	<.001	<.001	.021	.084	<.001	.006	.001

	N	1692	1691	1681	1686	1695	1712	1721	1698	1701	1685	1710	1697	1667	1671	1682	1702
sp8 Q1J. OFTEN PLAY CARDS/CHESS/OTHRS	*Pearson Correlation*	.097**	.049*	.076**	.164**	.087**	.076**	.222**	1	.084**	.113**	.064**	.076**	.094**	.084**	.055*	.047
	Sig. (2-tailed)	<.001	.045	.002	<.001	<.001	.002	<.001		.001	<.001	.008	.002	<.001	.001	.025	.054
	N	1683	1681	1674	1683	1689	1707	1698	1715	1695	1681	1705	1689	1660	1667	1676	1694
sp9 Q1K. OFTEN DO WRITING	*Pearson Correlation*	.105**	.216**	.222**	.151**	.232**	.176**	.132**	.084**	1	.312**	.074**	.119**	.108**	.264**	.181**	.155**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.001		<.001	.002	<.001	<.001	<.001	<.001	<.001
	N	1687	1684	1680	1681	1689	1700	1701	1695	1708	1683	1699	1694	1666	1663	1672	1690
sp10 Q1L. OFTEN USE COMPUTER	*Pearson Correlation*	.019	.208**	.144**	.153**	.176**	.192**	.152**	.113**	.312**	1	.127**	.068**	.034	.213**	.173**	.121**
	Sig. (2-tailed)	.444	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	.005	.165	<.001	<.001	<.001
	N	1675	1676	1668	1669	1680	1689	1685	1681	1683	1697	1689	1684	1662	1654	1660	1678
sp11 Q1M. OFTEN MAINTENANCE/GARDENING	*Pearson Correlation*	.094**	.125**	.047	.080**	.060*	.157**	.090**	.064**	.074**	.127**	1	.213**	.069**	.402**	.210**	.244**
	Sig. (2-tailed)	<.001	<.001	.054	.001	.013	<.001	<.001	.008	.002	<.001		<.001	.005	<.001	<.001	<.001
	N	1693	1694	1682	1694	1696	1738	1710	1705	1699	1689	1750	1708	1665	1685	1694	1724
sp12 Q1N. OFTEN BAKE OR COOK	*Pearson Correlation*	.129**	.051*	.090**	.068**	.089**	.099**	.056*	.076**	.119**	.068**	.213**	1	.097**	.131**	.087**	.111**
	Sig. (2-tailed)	<.001	.035	<.001	.005	<.001	<.001	.021	.002	<.001	.005	<.001		<.001	<.001	<.001	<.001
	N	1681	1683	1673	1679	1686	1710	1697	1689	1694	1684	1708	1722	1662	1667	1677	1699
sp13 Q1O. OFTEN SEW OR KNIT	*Pearson Correlation*	.117**	.083**	.171**	.076**	.151**	-.022	.042	.094**	.108**	.034	.069**	.097**	1	.069**	.004	-.008
	Sig. (2-tailed)	<.001	.001	<.001	.002	<.001	.376	.084	<.001	<.001	.165	.005	<.001		.005	.885	.760
	N	1660	1656	1654	1649	1661	1663	1667	1660	1666	1662	1665	1662	1669	1632	1639	1656
sp14 Q1P. OFTEN DO HOBBY	*Pearson Correlation*	.057*	.171**	.124**	.095**	.089**	.183**	.147**	.084**	.264**	.213**	.402**	.131**	.069**	1	.237**	.170**
	Sig. (2-tailed)	.021	<.001	<.001	<.001	<.001	<.001	<.001	.001	<.001	<.001	<.001	<.001	.005		<.001	<.001
	N	1657	1658	1647	1655	1661	1690	1671	1667	1663	1654	1685	1667	1632	1714	1681	1698
sp15 Q1Q. OFTEN PLAY SPORT/EXERCISE	*Pearson Correlation*	.060*	.159**	.171**	.298**	.121**	.247**	.067**	.055*	.181**	.173**	.210**	.087**	.004	.237**	1	.456**
	Sig. (2-tailed)	.015	<.001	<.001	<.001	<.001	<.001	.006	.025	<.001	<.001	<.001	<.001	.885	<.001		<.001
	N	1669	1670	1660	1669	1673	1702	1682	1676	1672	1660	1694	1677	1639	1681	1722	1711

sp16 Q1R. OFTEN WALK FOR 20 MINS	*Pearson Correlation*	.092**	.094**	.117**	.116**	.108**	.150**	.080**	.047	.155**	.121**	.244**	.111**	-.008	.170**	.456**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	.001	.054	<.001	<.001	<.001	<.001	.760	<.001	<.001	
	N	1689	1688	1676	1687	1686	1730	1702	1694	1690	1678	1724	1699	1656	1698	1711	1761

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table A.7

Correlations – Social participation items – Female respondents

		sp1	***sp2***	***sp3***	***sp4***	***sp5***	***sp6***	***sp7***	***sp8***	***sp9***	***sp10***	***sp11***	***sp12***	***sp13***	***sp14***	***sp15***	***sp16***
sp1 Q1B. OFTEN VOLUNTEER YOUTH	*Pearson Correlation*	1	.262**	.183**	.089**	.138**	.030	.010	-.007	.093**	.012	.071**	.152**	.093**	.094**	.119**	.102**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001	.149	.644	.735	<.001	.582	.001	<.001	<.001	<.001	<.001	<.001
	N	2299	2275	2240	2251	2262	2290	2279	2264	2272	2231	2274	2287	2258	2229	2238	2269
sp2 Q1C. OFTEN VOLUNTEER - OTHER	*Pearson Correlation*	.262**	1	.305**	.203**	.299**	.147**	.047*	.071**	.188**	.184**	.099**	.069**	.138**	.221**	.177**	.142**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001	<.001	.023	.001	<.001	<.001	<.001	.001	<.001	<.001	<.001	<.001
	N	2275	2312	2249	2263	2270	2302	2291	2271	2285	2235	2281	2300	2269	2238	2250	2280
sp3 Q1D. OFTEN EDUCATION	*Pearson Correlation*	.183**	.305**	1	.215**	.309**	.048*	-.022	.038	.178**	.146**	.085**	.032	.082**	.131**	.167**	.128**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001	.021	.303	.075	<.001	<.001	<.001	.125	<.001	<.001	<.001	<.001
	N	2240	2249	2267	2233	2236	2261	2254	2239	2251	2210	2248	2257	2232	2203	2219	2236
sp4 Q1E. OFTEN ATTEND SPORTS/SOCIAL/CLUB	*Pearson Correlation*	.089**	.203**	.215**	1	.358**	.129**	.094**	.194**	.148**	.208**	.171**	.107**	.106**	.234**	.326**	.181**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	N	2251	2263	2233	2296	2258	2288	2272	2260	2269	2220	2264	2284	2247	2221	2238	2261

sp5 Q1F. OFTEN ATTEND NON RELIGIOUS ORGS	*Pearson Correlation*	.138**	.299**	.309**	.358**	1	.041	.059**	.140**	.180**	.142**	.105**	.096**	.141**	.195**	.166**	.115**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001		.050	.005	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	N	2262	2270	2236	2258	2303	2294	2278	2267	2275	2226	2278	2291	2257	2224	2244	2268
sp6 Q1H. OFTEN READ	*Pearson Correlation*	.030	.147**	.048*	.129**	.041	1	.277**	.103**	.215**	.207**	.165**	.176**	.051*	.186**	.134**	.106**
	Sig. (2-tailed)	.149	<.001	.021	<.001	.050		<.001	<.001	<.001	<.001	<.001	<.001	.013	<.001	<.001	<.001
	N	2290	2302	2261	2288	2294	2461	2370	2317	2352	2268	2371	2412	2311	2302	2317	2374
sp7 Q1I. OFTEN DO WORD GAMES	*Pearson Correlation*	.010	.047*	-.022	.094**	.059**	.277**	1	.277**	.146**	.126**	.081**	.086**	.132**	.170**	.069**	.036
	Sig. (2-tailed)	.644	.023	.303	<.001	.005	<.001		<.001	<.001	<.001	<.001	<.001	<.001	<.001	.001	.080
	N	2279	2291	2254	2272	2278	2370	2382	2298	2326	2253	2324	2355	2293	2273	2294	2324
sp8 Q1J. OFTEN PLAY CARDS/CHESS/OTHRS	*Pearson Correlation*	-.007	.071**	.038	.194**	.140**	.103**	.277**	1	.183**	.198**	.096**	.092**	.149**	.194**	.111**	.069**
	Sig. (2-tailed)	.735	.001	.075	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001	<.001	<.001	.001
	N	2264	2271	2239	2260	2267	2317	2298	2323	2291	2240	2290	2308	2267	2243	2257	2279
sp9 Q1K. OFTEN DO WRITING	*Pearson Correlation*	.093**	.188**	.178**	.148**	.180**	.215**	.146**	.183**	1	.216**	.127**	.193**	.176**	.297**	.204**	.142**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001	<.001	<.001
	N	2272	2285	2251	2269	2275	2352	2326	2291	2362	2250	2313	2338	2283	2256	2276	2306

sp10 Q1L. OFTEN USE COMPUTER	*Pearson Correlation*	.012	.184**	.146**	.208**	.142**	.207**	.126**	.198**	.216**	1	.153**	.107**	.102**	.266**	.169**	.059**
	Sig. (2-tailed)	.582	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001	.005
	N	2231	2235	2210	2220	2226	2268	2253	2240	2250	2275	2261	2267	2246	2212	2230	2244
sp11 Q1M. OFTEN MAINTENANCE/GARDENING	*Pearson Correlation*	.071**	.099**	.085**	.171**	.105**	.165**	.081**	.096**	.127**	.153**	1	.356**	.148**	.272**	.235**	.256**
	Sig. (2-tailed)	.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001
	N	2274	2281	2248	2264	2278	2371	2324	2290	2313	2261	2385	2366	2293	2268	2294	2335
sp12 Q1N. OFTEN BAKE OR COOK	*Pearson Correlation*	.152**	.069**	.032	.107**	.096**	.176**	.086**	.092**	.193**	.107**	.356**	1	.183**	.234**	.153**	.204**
	Sig. (2-tailed)	<.001	.001	.125	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001
	N	2287	2300	2257	2284	2291	2412	2355	2308	2338	2267	2366	2432	2306	2294	2311	2363
sp13 Q1O. OFTEN SEW OR KNIT	*Pearson Correlation*	.093**	.138**	.082**	.106**	.141**	.051*	.132**	.149**	.176**	.102**	.148**	.183**	1	.492**	.118**	.093**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	.013	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001
	N	2258	2269	2232	2247	2257	2311	2293	2267	2283	2246	2293	2306	2318	2250	2260	2278
sp14 Q1P. OFTEN DO HOBBY	*Pearson Correlation*	.094**	.221**	.131**	.234**	.195**	.186**	.170**	.194**	.297**	.266**	.272**	.234**	.492**	1	.277**	.186**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001
	N	2229	2238	2203	2221	2224	2302	2273	2243	2256	2212	2268	2294	2250	2333	2272	2296

sp15 Q1Q. OFTEN PLAY SPORT/EXERCIZE	*Pearson Correlation*	.119**	.177**	.167**	.326**	.166**	.134**	.069**	.111**	.204**	.169**	.235**	.153**	.118**	.277**	1	.449**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001
	N	2238	2250	2219	2238	2244	2317	2294	2257	2276	2230	2294	2311	2260	2272	2351	2316
sp16 Q1R. OFTEN WALK FOR 20 MINS	*Pearson Correlation*	.102**	.142**	.128**	.181**	.115**	.106**	.036	.069**	.142**	.059**	.256**	.204**	.093**	.186**	.449**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	.080	.001	<.001	.005	<.001	<.001	<.001	<.001	<.001	
	N	2269	2280	2236	2261	2268	2374	2324	2279	2306	2244	2335	2363	2278	2296	2316	2415

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix B: Scale Construction Results

Table B.1

Parameters for initial six-factor scale model of social participation – Male respondents – Measure construction subsample

	Unstandardized			*Standardized*		
	Est.	SE	p	Est.	SE	p
Factor Loadings						
Volunteering by						
Volunteer work with children or young people	1.00	0.00	999	0.33	0.07	<.001
Other volunteer or charity work	4.41	1.44	0.002	0.87	0.14	<.001
Social by						
Attend educational or training course	1.00	0.00	999	0.42	0.06	<.001
Go to a sport, social or other club	2.49	0.46	<.001	0.49	0.06	<.001
Attend meetings of non-religious organizations	1.58	0.33	<.001	0.57	0.06	<.001
Games by						
Do word games	1.00	0.00	999	0.61	0.08	<.001
Play cards or games such as chess	0.48	0.12	<.001	0.37	0.06	<.001
Intellectual by						
Read books, magazines, or newspapers	1.00	0.00	999	0.49	0.04	<.001
Do word games	0.89	0.15	<.001	0.51	0.04	<.001
Use a computer for e-mail, Internet, or other tasks	1.44	0.19	<.001	0.51	0.04	<.001
Home and Hobbies by						
Do home or car maintenance or gardening	1.00	0.00	999	0.63	0.04	<.001
Bake or cook something special	0.46	0.06	<.001	0.33	0.05	<.001
Make clothes, knit, embroider, etc.	0.01	0.01	0.475	0.03	0.04	0.510
Work on a hobby or project	1.14	0.12	<.001	0.72	0.04	<.001
Sports and Exercise by						
Play sports or exercise	1.00	0.00	999	0.73	0.04	<.001
Walk for 20 minutes or more	0.89	0.08	<.001	0.66	0.04	<.001
Factor Covariances						
Social with Volunteering	0.04	0.01	0.010	0.52	0.12	<.001
Games with Volunteering	0.01	0.02	0.666	0.03	0.07	0.666
Games with Social	0.10	0.03	0.004	0.29	0.11	0.009
Intellectual with Volunteering	0.09	0.03	0.003	0.47	0.10	<.001
Intellectual with Social	0.16	0.03	<.001	0.72	0.09	<.001
Intellectual with Games	0.58	0.12	<.001	0.62	0.10	<.001

Home and Hobbies with Volunteering	0.09	0.03	0.012	0.29	0.08	<.001
Home and Hobbies with Social	0.10	0.03	0.001	0.31	0.07	<.001
Home and Hobbies with Games	0.64	0.11	<.001	0.45	0.08	<.001
Home and Hobbies with Intellectual	0.67	0.12	<.001	0.70	0.06	<.001
Sports and Exercise with Volunteering	0.09	0.04	0.014	0.25	0.07	<.001
Sports and Exercise with Social	0.19	0.04	<.001	0.46	0.06	<.001
Sports and Exercise with Games	0.42	0.14	0.002	0.24	0.08	0.004
Sports and Exercise with Intellectual	0.82	0.13	<.001	0.70	0.06	<.001
Sports and Exercise with Home and Hobbies	0.95	0.13	<.001	0.54	0.06	<.001
Correlated Residuals						
"Go to a sport, social or other club" with "Play sports or exercise"	0.47	0.10	<.001	0.27	0.05	<.001
Factor Variances						
Volunteering	0.06	0.03	0.034			
Social	0.08	0.03	0.002			
Games	1.41	0.40	<.001			
Intellectual	0.64	0.14	<.001			
Home and Hobbies	1.43	0.20	<.001			
Sports and Exercise	2.16	0.23	<.001			
Residual Variances						
Volunteer work with children or young people	0.50	0.09	<.001	0.89	0.04	<.001
Other volunteer or charity work	0.38	0.39	0.334	0.24	0.25	0.341
Attend educational or training course	0.36	0.06	<.001	0.82	0.05	<.001
Go to a sport, social or other club	1.57	0.15	<.001	0.76	0.06	<.001
Attend meetings of non-religious organizations	0.41	0.06	<.001	0.68	0.07	<.001
Do word games	2.06	0.14	<.001	0.76	0.04	<.001
Play cards or games such as chess	2.36	0.37	<.001	0.63	0.10	<.001
Read books, magazines, or newspapers	2.09	0.16	<.001	0.87	0.04	<.001
Do word games	1.44	0.12	<.001	0.74	0.04	<.001
Use a computer for e-mail, Internet, or other tasks	3.84	0.22	<.001	0.74	0.04	<.001
Do home or car maintenance or gardening	2.12	0.19	<.001	0.60	0.05	<.001
Bake or cook something special	2.42	0.14	<.001	0.89	0.03	<.001
Make clothes, knit, embroider, etc.	0.12	0.06	0.038	>.99	0.00	<.001
Work on a hobby or project	1.73	0.20	<.001	0.48	0.06	<.001
Play sports or exercise	1.86	0.22	<.001	0.46	0.05	<.001
Walk for 20 minutes or more	2.24	0.20	<.001	0.57	0.05	<.001

Table B.2

Parameters for initial eight-factor scale model of social participation – Female respondents – Measure construction subsample

	Unstandardized			*Standardized*		
	Est.	SE	p	Est.	SE	p
Factor Loadings						
Volunteering by						
Volunteer work with children or young people	1.00	0.00	999	0.32	0.07	<.001
Other volunteer or charity work	2.95	0.94	0.002	0.78	0.10	<.001
Clubs by						
Go to a sport, social or other club	1.00	0.00	999	0.49	0.06	<.001
Attend meetings of non-religious organizations	0.73	0.18	<.001	0.60	0.08	<.001
Games by						
Do word games	1.00	0.00	999	0.54	0.06	<.001
Play cards or games such as chess	0.75	0.13	<.001	0.52	0.05	<.001
Intellectual by						
Read books, magazines, or newspapers	1.00	0.00	999	0.49	0.04	<.001
Do writing	1.08	0.17	<.001	0.47	0.04	<.001
Use a computer for e-mail, Internet or other tasks	1.53	0.24	<.001	0.49	0.04	<.001
Home by						
Do home or car maintenance or gardening	1.00	0.00	999	0.66	0.05	<.001
Bake or cook something special	0.78	0.10	<.001	0.59	0.05	<.001
Hobbies by						
Make clothes, knit, embroider, etc.	1.00	0.00	999	0.53	0.04	<.001
Work on a hobby or project	2.36	0.36	<.001	0.97	0.07	<.001
Sports by						
Play sports or exercise	1.00	0.00	999	0.65	0.05	<.001
Walk for 20 minutes or more	0.96	0.12	<.001	0.62	0.05	<.001
Factor Covariances						
Education with Volunteering	0.12	0.04	0.004	0.46	0.07	<.001
Education with Clubs	0.17	0.04	<.001	0.39	0.10	<.001
Education with Intellectual	0.17	0.03	<.001	0.32	0.05	<.001
Education with Home	0.10	0.04	0.021	0.10	0.04	0.013
Education with Hobbies	0.06	0.02	0.016	0.10	0.04	0.007
Education with Sports and Exercise	0.24	0.06	<.001	0.26	0.05	<.001
Clubs with Volunteering	0.10	0.04	0.008	0.49	0.12	<.001
Games with Volunteering	0.06	0.03	0.019	0.16	0.07	0.019

Games with Clubs	0.21	0.07	0.003	0.29	0.09	0.001
Intellectual with Volunteering	0.12	0.04	0.004	0.50	0.08	<.001
Intellectual with Clubs	0.22	0.06	<.001	0.51	0.09	<.001
Intellectual with Games	0.58	0.14	<.001	0.70	0.07	<.001
Home with Volunteering	0.09	0.05	0.074	0.20	0.08	0.007
Home with Clubs	0.24	0.07	0.001	0.30	0.07	<.001
Home with Games	0.37	0.12	0.001	0.24	0.07	<.001
Home with Intellectual	0.48	0.09	<.001	0.51	0.07	<.001
Hobbies with Volunteering	0.08	0.03	0.023	0.28	0.06	<.001
Hobbies with Clubs	0.16	0.04	<.001	0.33	0.06	<.001
Hobbies with Games	0.33	0.09	<.001	0.36	0.07	<.001
Hobbies with Intellectual	0.27	0.05	<.001	0.49	0.06	<.001
Hobbies with Home	0.39	0.08	<.001	0.38	0.06	<.001
Sports and Exercise with Volunteering	0.16	0.06	0.014	0.35	0.08	<.001
Sports and Exercise with Clubs	0.35	0.09	<.001	0.45	0.06	<.001
Sports and Exercise with Games	0.23	0.11	0.032	0.16	0.07	0.027
Sports and Exercise with Intellectual	0.42	0.09	<.001	0.46	0.07	<.001
Sports and Exercise with Home	0.82	0.12	<.001	0.49	0.07	<.001
Sports and Exercise with Hobbies	0.34	0.07	<.001	0.34	0.05	<.001
Correlated Residuals						
"Go to a sport, social or other club" with "Play sports or exercise"	0.42	0.08	<.001	0.26	0.04	<.001
Factor Variances						
Education	0.55	0.08	<.001			
Volunteering	0.12	0.06	0.035			
Clubs	0.37	0.10	<.001			
Games	1.36	0.29	<.001			
Intellectual	0.51	0.12	<.001			
Home	1.75	0.27	<.001			
Hobbies	0.61	0.12	<.001			
Sports and Exercise	1.61	0.25	<.001			
Residual Variances						
Volunteer work with children or young people	1.06	0.13	<.001	0.90	0.05	<.001
Other volunteer or charity work	0.67	0.28	0.018	0.39	0.16	0.017
Go to a sport, social or other club	1.13	0.12	<.001	0.76	0.06	<.001
Attend meetings of non-religious organizations	0.35	0.06	<.001	0.65	0.10	<.001
Do word games	1.57	0.12	<.001	0.76	0.04	<.001
Play cards or games such as chess	3.33	0.29	<.001	0.71	0.06	<.001
Read books, magazines, or newspapers	2.11	0.18	<.001	0.73	0.06	<.001
Do word games	2.02	0.13	<.001	0.78	0.04	<.001

Use a computer for e-mail, Internet, or other tasks	3.80	0.19	<.001	0.76	0.04	<.001
Do home or car maintenance or gardening	2.23	0.27	<.001	0.56	0.07	<.001
Bake or cook something special	1.94	0.17	<.001	0.65	0.06	<.001
Make clothes, knit, embroider, etc.	1.60	0.13	<.001	0.72	0.05	<.001
Work on a hobby or project	0.19	0.50	0.699	0.05	0.14	0.699
Play sports or exercise	2.22	0.26	<.001	0.58	0.07	<.001
Walk for 20 minutes or more	2.36	0.22	<.001	0.62	0.06	<.001

Table B.3

Parameter estimates from alternate four-factor scale model for females – Validation subsample

	Unstandardized			***Standardized***		
	Estimate	***SE***	***p***	***Estimate***	***SE***	***p***
Factor loadings						
Community by						
Volunteer - other	1.000	0.000	999.000	0.540	0.046	0.000
Education/training	0.657	0.104	0.000	0.524	0.059	0.000
Non-religious organizations	0.818	0.132	0.000	0.602	0.062	0.000
Intellectual by						
Reading	1.000	0.000	999.000	0.422	0.042	0.000
Do word games	1.774	0.272	0.000	0.446	0.052	0.000
Cards/chess/other games	1.367	0.254	0.000	0.432	0.043	0.000
Use computer	1.809	0.326	0.000	0.441	0.048	0.000
Home and hobbies by						
Home/car maintenance or gardening	1.000	0.000	999.000	0.455	0.044	0.000
Baking/cooking	0.781	0.079	0.000	0.414	0.046	0.000
Sewing/knitting	0.970	0.146	0.000	0.571	0.035	0.000
Hobby/project	1.640	0.213	0.000	0.777	0.034	0.000
Sports/exercise by						
Sport/social clubs	1.000	0.000	999.000	0.515	0.047	0.000
Sports/exercise	2.055	0.280	0.000	0.724	0.038	0.000
Walk for 20 minutes	1.686	0.233	0.000	0.608	0.037	0.000
Factor covariances						
Intellectual with Community	0.129	0.033	0.000	0.353	0.076	0.000
Home/hobbies with Community	0.282	0.057	0.000	0.468	0.058	0.000
Home/hobbies with intellectual	0.292	0.060	0.000	0.605	0.056	0.000
Sports/exercise with Community	0.264	0.062	0.000	0.558	0.068	0.000
Sports/exercise with Intellectual	0.168	0.038	0.000	0.442	0.067	0.000
Sports/exercise with Home/hobbies	0.354	0.072	0.000	0.565	0.050	0.000
Factor variances						
Community	0.455	0.088	0.000	1.000	0.000	999.000
Intellectual	0.292	0.076	0.000	1.000	0.000	999.000
Home and hobbies	0.797	0.156	0.000	1.000	0.000	999.000

Sports and exercise	0.493	0.101	0.000	1.000	0.000	999.000
Residual variances						
Other volunteer or charity work	1.104	0.088	0.000	0.708	0.050	0.000
Attend educational or training course	0.519	0.064	0.000	0.725	0.062	0.000
Go to a sport, social or other club	1.365	0.095	0.000	0.735	0.048	0.000
Attend meetings of non-religious organizations	0.535	0.070	0.000	0.637	0.075	0.000
Read books, magazines, or newspapers	1.350	0.108	0.000	0.822	0.036	0.000
Do word games	3.712	0.217	0.000	0.802	0.047	0.000
Play cards or games such as chess	2.377	0.139	0.000	0.813	0.037	0.000
Use a computer	3.967	0.213	0.000	0.806	0.042	0.000
Do home or car maintenance or gardening	3.055	0.154	0.000	0.793	0.040	0.000
Bake or cook something special	2.342	0.123	0.000	0.828	0.038	0.000
Make clothes, knit, embroider, etc.	1.552	0.100	0.000	0.674	0.040	0.000
Work on a hobby or project	1.403	0.187	0.000	0.396	0.054	0.000
Play sports or exercise	1.891	0.220	0.000	0.476	0.055	0.000
Walk for 20 minutes or more	2.391	0.171	0.000	0.630	0.044	0.000

Appendix C: Index Construction Results

Table C.1

Coefficient estimates from disaggregated model with different levels of measurement error in participation items specified

		No measurement error (N=1790)			***10% measurement error (N=1791)***			***20% measurement error (N=1791)***			***30% measurement error (N=1791)***		
		Coef	***SE***	***p***	***Coef***	***SE***	***p***	***Coef***	***SE***	***p***	***Coef***	***SE***	***p***
Life satisfaction on													
	Social connectedness	*1.156*	*0.147*	***0.000***	*1.147*	*0.147*	***0.000***	*1.137*	*0.148*	***0.000***	*1.129*	*0.149*	***0.000***
	sp1. Volunteer with youth	-0.002	0.042	0.964	-0.005	0.048	0.909	-0.011	0.057	0.846	-0.028	0.071	0.694
	sp2. Volunteer – other	*0.066*	*0.029*	***0.023***	*0.074*	*0.034*	***0.030***	*0.086*	*0.042*	***0.041***	*0.11*	*0.057*	*0.054*
	sp3. Education	-0.047	0.061	0.443	-0.059	0.07	0.395	-0.076	0.082	0.351	-0.089	0.102	0.384
	sp4. Sports/social club	0.051	0.028	0.068	0.056	0.033	0.089	0.063	0.041	0.118	0.085	0.056	0.128
	sp5. Non-religious organizations	0.01	0.049	0.834	0.006	0.056	0.921	-0.002	0.068	0.976	-0.013	0.087	0.879
	sp6. Read	0.024	0.031	0.442	0.023	0.036	0.527	0.02	0.043	0.641	0.022	0.055	0.695
	sp7. Word games	-0.009	0.018	0.619	-0.01	0.021	0.619	-0.012	0.025	0.638	-0.013	0.032	0.685
	sp8. Cards/chess/other games	-0.012	0.024	0.611	-0.016	0.027	0.551	-0.022	0.031	0.483	-0.025	0.037	0.509
	sp9. Writing	-0.027	0.029	0.352	-0.035	0.034	0.295	-0.048	0.04	0.238	-0.059	0.051	0.248
	sp10. Computer	*0.068*	*0.018*	***0.000***	*0.076*	*0.02*	***0.000***	*0.087*	*0.024*	***0.000***	*0.066*	*0.019*	***0.001***
	sp11. Home maintenance/ gardening	0.042	0.022	0.058	0.041	0.026	0.112	0.039	0.033	0.237	0.028	0.047	0.545
	sp12. Bake or cook	0.038	0.023	0.099	0.043	0.027	0.106	0.052	0.033	0.111	0.069	0.045	0.121
	sp13. Sew or knit	-0.012	0.033	0.708	-0.017	0.038	0.652	-0.026	0.047	0.583	-0.05	0.065	0.446
	sp14. Hobby	*0.059*	*0.024*	***0.013***	*0.065*	*0.029*	***0.023***	*0.076*	*0.038*	***0.046***	*0.105*	*0.059*	*0.076*
	sp15. Sports/Exercise	0.024	0.022	0.271	0.022	0.026	0.405	0.018	0.035	0.605	0.008	0.054	0.880
	sp16. Walk 20 minutes	0.035	0.022	0.102	0.038	0.026	0.140	0.042	0.032	0.193	0.051	0.046	0.264
Social connectedness on													

	sp1. Volunteer with youth	0.002	0.011	0.886	-0.001	0.013	0.950	-0.005	0.015	0.741	-0.013	0.018	0.461
	sp2. Volunteer – other	*0.024*	*0.008*	***0.002***	*0.028*	*0.009*	***0.003***	*0.033*	*0.012*	***0.005***	*0.042*	*0.015*	***0.007***
	sp3. Education	-0.003	0.013	0.828	-0.006	0.015	0.707	-0.01	0.018	0.570	-0.018	0.023	0.429
	sp4. Sports/social club	-0.003	0.009	0.737	-0.005	0.01	0.654	-0.007	0.012	0.563	-0.012	0.016	0.474
	sp5. Non-religious organizations	0.021	0.012	0.079	0.023	0.014	0.099	0.026	0.017	0.126	0.031	0.022	0.160
	sp6. Read	*0.015*	*0.007*	***0.035***	*0.016*	*0.008*	*0.054*	*0.017*	*0.01*	*0.090*	*0.019*	*0.013*	*0.152*
	sp7. Word games	0.002	0.005	0.753	0.001	0.006	0.845	0	0.007	0.949	-0.001	0.009	0.940
	sp8. Cards/chess/other games	0.006	0.007	0.391	0.006	0.007	0.411	0.007	0.009	0.428	0.009	0.01	0.413
	sp9. Writing	0.01	0.007	0.137	0.011	0.008	0.192	0.011	0.01	0.276	0.011	0.012	0.373
	sp10. Computer	0.006	0.005	0.212	0.006	0.006	0.259	0.007	0.007	0.319	0.004	0.005	0.399
	sp11. Home maintenance/ gardening	0.002	0.005	0.757	0.001	0.006	0.901	-0.001	0.008	0.935	-0.003	0.012	0.790
	sp12. Bake or cook	*0.024*	*0.006*	***0.000***	*0.027*	*0.007*	***0.000***	*0.032*	*0.008*	***0.000***	*0.039*	*0.011*	***0.000***
	sp13. Sew or knit	-0.002	0.008	0.748	-0.004	0.009	0.690	-0.005	0.011	0.643	-0.008	0.016	0.621
	sp14. Hobby	0.001	0.006	0.876	0	0.008	0.968	-0.001	0.01	0.942	-0.002	0.015	0.886
	sp15. Sports/Exercise	0.009	0.006	0.128	0.01	0.007	0.145	0.013	0.009	0.161	0.019	0.014	0.181
	sp16. Walk 20 minutes	0.003	0.006	0.641	0.002	0.007	0.784	0	0.008	0.975	-0.004	0.012	0.765

Note. p-values less than .05 shaded. Rows for predictors with at least one significant p-value indicated with italics. Sample size for model estimated with no measurement error smaller than that for other models because Mplus treats models with exogenous variables differently than models with endogenous variables.

Appendix D: Results of Post-Comparison Analyses

Table D.1

Estimated parameters – Final model of life satisfaction as related to social connectedness and social participation – Male

	Unstandardized			*Standardized*		
	Est.	***SE***	***p***	***Est.***	***SE***	***p***
Factor loadings						
Life satisfaction by						
LS1. Life is close to ideal	1.000	0.000	999.000	0.720	0.030	<.001
LS2. Conditions of life excellent	1.128	0.044	<.001	0.780	0.025	<.001
LS3. Satisfied with life	1.173	0.077	<.001	0.882	0.019	<.001
LS4. Have important things to do in life	0.940	0.074	<.001	0.736	0.032	<.001
LS5. Would change nothing	0.943	0.070	<.001	0.571	0.033	<.001
Social connectedness by						
SC1. In tune with others	1.000	0.000	999.000	0.476	0.037	<.001
SC2. Have people I can talk to	1.267	0.112	<.001	0.697	0.028	<.001
SC3. Have people I can turn to	1.430	0.130	<.001	0.778	0.021	<.001
SC4. There are people who understand me	1.426	0.118	<.001	0.769	0.021	<.001
SC5. There are people I feel close to	1.413	0.126	<.001	0.806	0.019	<.001
SC6. Feel part of a group	1.420	0.127	<.001	0.675	0.028	<.001
SC7. Have a lot in common with friends.	1.264	0.109	<.001	0.653	0.028	<.001
Single-item participation factors						
SP2L by SP2	1.000	0.000	999.000	0.949	0.004	<.001
SP4L by SP4	1.000	0.000	999.000	0.948	0.004	<.001
SP11L by SP11	1.000	0.000	999.000	0.948	0.002	<.001
SP14L by SP14	1.000	0.000	999.000	0.948	0.002	<.001
Structural coefficients						
Life satisfaction on						
Social connectedness	1.011	0.186	<.001	0.274	0.049	<.001
Participation	-0.005	0.037	0.887	-0.008	0.057	0.884
Social connectedness on						
Participation	0.030	0.012	0.012	0.173	0.045	<.001
Participation on						
SP2L. Volunteer – other	1.000	0.000	999.000	0.637	0.219	0.004
SP4L. Sports/social clubs	0.679	0.504	0.178	0.481	0.225	0.032
SP11L. Home maintenance/gardening	-0.030	0.370	0.936	-0.025	0.321	0.937
SP14L. Hobbies/projects	0.424	0.367	0.249	0.384	0.258	0.137

Life satisfaction on						
Age	0.000	0.007	0.973	0.001	0.040	0.973
Years education	-0.027	0.015	0.075	-0.072	0.040	0.074
Depression	-0.198	0.036	<.001	-0.264	0.046	<.001
Partnered	0.264	0.110	0.017	0.093	0.038	0.015
Self-report health	0.223	0.054	<.001	0.200	0.046	<.001
Log(wealth)	0.243	0.078	0.002	0.116	0.038	0.002
Social connectedness on						
Age	-0.001	0.002	0.742	-0.013	0.041	0.743
Years education	0.002	0.005	0.604	0.023	0.045	0.604
Depression	-0.027	0.010	0.008	-0.131	0.048	0.006
Partnered	-0.019	0.034	0.573	-0.025	0.044	0.574
Self-report health	0.023	0.015	0.123	0.076	0.049	0.120
Log(wealth)	0.004	0.026	0.865	0.008	0.046	0.865
SP4L. Sports/social clubs on						
Log(wealth)	0.327	0.118	0.005	0.141	0.051	0.006
SP11L. Home maintenance/gardening on						
Age	-0.036	0.010	<.001	-0.147	0.042	0.001
SP14L. Hobbies/projects on						
Years education	0.074	0.022	0.001	0.137	0.041	0.001
Factor covariances						
SP2L with SP4L	0.329	0.105	0.002	0.202	0.059	0.001
SP2L with SP11L	0.245	0.087	0.005	0.124	0.043	0.005
SP2L with SP14L	0.353	0.109	0.001	0.169	0.049	0.001
SP4L with SP11L	0.021	0.104	0.839	0.010	0.048	0.839
SP4L with SP14L	0.127	0.114	0.265	0.055	0.049	0.262
SP11L with SP14L	1.220	0.117	<.001	0.438	0.038	<.001
Residual correlations						
SC2 with SC3	0.062	0.011	<.001	0.374	0.055	<.001
SC6 with SC7	0.097	0.014	<.001	0.391	0.043	<.001
LS1 with LS2	0.588	0.120	<.001	0.452	0.063	<.001
LS4 with LS5	0.360	0.104	0.001	0.206	0.053	<.001
Factor variances						
SP2L	1.484	0.131	<.001	1.000	0.000	999.000
Residual variances						
SP2. Volunteer – other	0.165	0.000	999.000	0.100	0.008	<.001
SP4. Sports/social clubs	0.205	0.000	999.000	0.101	0.007	<.001
SP11. Home maintenance/gardening	0.306	0.000	999.000	0.102	0.003	<.001
SP14. Hobbies/projects	0.335	0.000	999.000	0.101	0.003	<.001
LS1. Life is close to ideal	1.388	0.130	<.001	0.482	0.043	<.001
LS2. Conditions of life excellent	1.220	0.127	<.001	0.391	0.040	<.001

LS3. Satisfied with life	0.586	0.080	<.001	0.222	0.034	<.001
LS4. Have important things to do in life	1.117	0.110	<.001	0.459	0.047	<.001
LS5. Would change nothing	2.741	0.170	<.001	0.674	0.038	<.001
SC1. In tune with others	0.373	0.023	<.001	0.774	0.035	<.001
SC2. Have people I can talk to	0.186	0.016	<.001	0.515	0.039	<.001
SC3. Have people I can turn to	0.145	0.013	<.001	0.394	0.033	<.001
SC4. There are people who understand me	0.153	0.011	<.001	0.408	0.032	<.001
SC5. There are people I feel close to	0.118	0.010	<.001	0.351	0.031	<.001
SC6. Feel part of a group	0.264	0.019	<.001	0.545	0.038	<.001
SC7. Have a lot in common with friends.	0.235	0.015	<.001	0.574	0.036	<.001
Life satisfaction	1.040	0.128	<.001	0.698	0.042	<.001
Social connectedness	0.102	0.016	<.001	0.934	0.020	<.001
Participation	0.000	0.000	999.000	0.000	999.000	999.000
SP4L	1.797	0.136	<.001	0.980	0.014	<.001
SP11L	2.647	0.102	<.001	0.978	0.012	<.001
SP14L	2.933	0.108	<.001	0.981	0.011	<.001

Table D.2

Estimated parameters – Final model of life satisfaction as related to social connectedness and social participation – Female

	Unstandardized			***Standardized***		
	Estimate	***S.E.***	***p***	***Estimate***	***S.E.***	***p***
Factor loadings						
Life satisfaction by						
LS1. Life is close to ideal	1.000	0.000	999.000	0.755	0.023	<.001
LS2. Conditions of life excellent	1.115	0.037	<.001	0.826	0.019	<.001
LS3. Satisfied with life	1.095	0.048	<.001	0.871	0.015	<.001
LS4. Have important things to do in life	0.821	0.057	<.001	0.691	0.030	<.001
LS5. Would change nothing	0.860	0.059	<.001	0.576	0.032	<.001
Social connectedness by						
SC1. In tune with others	1.000	0.000	999.000	0.444	0.033	<.001
SC2. Have people I can talk to	1.431	0.122	<.001	0.732	0.026	<.001
SC3. Have people I can turn to	1.405	0.124	<.001	0.729	0.030	<.001
SC4. There are people who understand me	1.550	0.119	<.001	0.743	0.022	<.001
SC5. There are people I feel close to	1.366	0.114	<.001	0.760	0.023	<.001
SC6. Feel part of a group	1.478	0.117	<.001	0.604	0.028	<.001
SC7. Have a lot in common with friends.	1.275	0.105	<.001	0.585	0.031	<.001
Single-item participation factors						
SP2L by SP2	1.000	0.000	999.000	0.948	0.003	<.001
SP10L by SP10	1.000	0.000	999.000	0.948	0.002	<.001
SP12L by SP12	1.000	0.000	999.000	0.948	0.002	<.001
SP14L by SP14	1.000	0.000	999.000	0.948	0.001	<.001
Structural coefficients						
Life satisfaction						
Social connectedness	1.205	0.205	<.001	0.247	0.038	<.001
Participation	0.075	0.034	0.028	0.129	0.038	0.001
Social connectedness on						
Participation	0.014	0.007	0.053	0.120	0.041	0.003
Participation on						
SP2L. Volunteer – other	1.000	0.000	999.000	0.493	0.181	0.007
SP10L. Computer	0.197	0.292	0.500	0.171	0.234	0.464
SP12L. Baking/cooking	0.660	0.481	0.170	0.437	0.216	0.043
SP14L. Hobbies/projects	0.646	0.408	0.113	0.477	0.197	0.016
Life satisfaction on						
Age	0.022	0.006	<.001	0.121	0.034	<.001

Years education	-0.034	0.016	0.036	-0.072	0.034	0.036
Depression	-0.168	0.027	<.001	-0.238	0.038	<.001
Partnered	0.199	0.094	0.033	0.071	0.033	0.032
Self-report health	0.211	0.050	<.001	0.161	0.037	<.001
Log(wealth)	0.313	0.088	<.001	0.110	0.031	<.001
Social connectedness						
Age	0.001	0.001	0.326	0.036	0.036	0.321
Years education	0.009	0.004	0.016	0.090	0.037	0.014
Depression	-0.037	0.007	<.001	-0.255	0.043	<.001
Partnered	-0.045	0.020	0.025	-0.078	0.035	0.024
Self-report health	0.011	0.010	0.272	0.042	0.038	0.265
Log(wealth)	0.014	0.020	0.492	0.024	0.035	0.489
SP10L. Computer on						
Age	-0.073	0.008	<.001	-0.266	0.029	<.001
Years education	0.213	0.021	<.001	0.295	0.030	<.001
Log(wealth)	0.714	0.128	<.001	0.167	0.031	<.001
SP12L. Baking/cooking on						
Partnered	0.675	0.102	<.001	0.210	0.032	<.001
SP14L. Hobbies/projects on						
Years education	0.131	0.019	<.001	0.213	0.031	<.001
Factor covariances						
SP10L with SP2L	0.299	0.087	0.001	0.140	0.040	<.001
SP10L with SP12L	0.034	0.102	0.736	0.012	0.036	0.736
SP10L with SP14L	0.558	0.130	<.001	0.178	0.041	<.001
SP12L with SP2L	0.125	0.068	0.066	0.068	0.037	0.066
SP12L with SP14L	0.780	0.105	<.001	0.290	0.038	<.001
SP14L with SP2L	0.435	0.080	<.001	0.212	0.038	<.001
Residual correlations						
SC2 with SC3	0.057	0.011	<.001	0.397	0.058	<.001
SC6 with SC7	0.099	0.015	<.001	0.356	0.041	<.001
LS1 with LS2	0.466	0.089	<.001	0.364	0.051	<.001
LS4 with LS5	0.481	0.115	<.001	0.237	0.048	<.001
Factor variances						
SP2L	1.398	0.092	<.001	1.000	0.000	999.000
Residual variances						
SP2. Volunteer – other	0.156	0.000	999.000	0.101	0.006	<.001
SP10. Computer	0.492	0.000	999.000	0.102	0.003	<.001
SP12. Baking/cooking	0.282	0.000	999.000	0.100	0.003	<.001
SP14. Hobbies/projects	0.354	0.000	999.000	0.101	0.003	<.001
LS1. Life is close to ideal	1.460	0.114	<.001	0.430	0.035	<.001
LS2. Conditions of life excellent	1.122	0.107	<.001	0.318	0.031	<.001

LS3. Satisfied with life	0.737	0.077	<.001	0.241	0.026	<.001
LS4. Have important things to do in life	1.427	0.113	<.001	0.523	0.041	<.001
LS5. Would change nothing	2.883	0.175	<.001	0.669	0.037	<.001
SC1. In tune with others	0.329	0.019	<.001	0.803	0.029	<.001
SC2. Have people I can talk to	0.144	0.013	<.001	0.464	0.038	<.001
SC3. Have people I can turn to	0.141	0.016	<.001	0.469	0.043	<.001
SC4. There are people who understand me	0.158	0.011	<.001	0.448	0.033	<.001
SC5. There are people I feel close to	0.110	0.009	<.001	0.422	0.035	<.001
SC6. Feel part of a group	0.308	0.019	<.001	0.635	0.033	<.001
SC7. Have a lot in common with friends.	0.253	0.017	<.001	0.657	0.036	<.001
Life satisfaction	1.369	0.124	<.001	0.708	0.033	<.001
Social connectedness	0.072	0.011	<.001	0.884	0.024	<.001
Participation	0.000	0.000	999.000	0.000	999.000	999.000
SP10L	3.284	0.120	<.001	0.756	0.026	<.001
SP12L	2.414	0.083	<.001	0.956	0.013	<.001
SP14L	3.004	0.099	<.001	0.955	0.013	<.001

Table D.3

Estimated parameters – Model of life satisfaction as related to social connectedness and social participation – Female – Using one-factor, six-item scale measure of participation

	Unstandardized			Standardized		
	Estimate	*SE*	*p*	*Estimate*	*SE*	*p*
Factor loadings						
Life satisfaction by						
LS1. Life is close to ideal	1	0	999	0.759	0.023	0
LS2. Conditions of life excellent	1.114	0.037	0	0.829	0.019	0
LS3. Satisfied with life	1.095	0.048	0	0.874	0.015	0
LS4. Have important things to do in life	0.821	0.057	0	0.695	0.029	0
LS5. Would change nothing	0.86	0.06	0	0.58	0.032	0
Social connectedness by						
SC1. In tune with others	1	0	999	0.446	0.033	0
SC2. Have people I can talk to	1.434	0.122	0	0.735	0.026	0
SC3. Have people I can turn to	1.408	0.124	0	0.732	0.029	0
SC4. There are people who understand me	1.551	0.119	0	0.745	0.022	0
SC5. There are people I feel close to	1.367	0.114	0	0.762	0.023	0
SC6. Feel part of a group	1.477	0.117	0	0.606	0.028	0
SC7. Have a lot in common with friends.	1.275	0.105	0	0.587	0.031	0
Participation by						
SP2. Volunteer - other	1	0	999	0.389	0.038	0
SP4. Sports/social clubs	1.215	0.143	0	0.433	0.04	0
SP9. Writing	1.449	0.205	0	0.433	0.04	0
SP10. Computer	2.597	0.35	0	0.568	0.037	0
SP14. Hobbies/projects	2.095	0.253	0	0.541	0.035	0
SP15. Sports/exercise	1.988	0.231	0	0.485	0.037	0
Structural coefficients						
Life satisfaction on						
Social connectedness	1.219	0.21	0	0.248	0.04	0
Participation	0.387	0.198	0.05	0.133	0.066	0.044
Social connectedness on						
Participation on	0.087	0.047	0.064	0.147	0.076	0.052
Participation on						
Age	-0.015	0.003	0	-0.233	0.042	0
Years education	0.053	0.008	0	0.315	0.04	0
Depression	-0.022	0.01	0.03	-0.091	0.04	0.024
Partnered	-0.035	0.04	0.393	-0.035	0.04	0.376

Self-report health	0.134	0.021	0	0.293	0.04	0
Log(wealth)	0.169	0.045	0	0.17	0.043	0
Life satisfaction on						
Age	0.023	0.006	0	0.127	0.034	0
Years education	-0.045	0.019	0.017	-0.093	0.039	0.017
Depression	-0.166	0.027	0	-0.232	0.038	0
Partnered	0.232	0.094	0.014	0.081	0.033	0.013
Self-report health	0.189	0.054	0.001	0.142	0.041	0
Log(wealth)	0.272	0.093	0.003	0.094	0.032	0.003
Social connectedness on						
Age	0.002	0.001	0.193	0.047	0.036	0.188
Years education	0.006	0.004	0.148	0.062	0.042	0.142
Depression	-0.036	0.007	0	-0.249	0.043	0
Partnered	-0.038	0.02	0.059	-0.066	0.035	0.059
Self-report health	0.005	0.011	0.648	0.019	0.041	0.646
Log(wealth)	0.004	0.023	0.864	0.007	0.039	0.864
Residual correlations						
SC2 with SC3	0.056	0.011	0	0.395	0.058	0
SC6 with SC7	0.1	0.015	0	0.357	0.041	0
LS1 with LS2	0.466	0.09	0	0.364	0.051	0
LS4 with LS5	0.481	0.115	0	0.237	0.048	0
Residual variances						
SP2. Volunteer - other	1.319	0.081	0	0.848	0.029	0
SP4. Sports/social clubs	1.509	0.091	0	0.813	0.034	0
SP9. Writing	2.141	0.12	0	0.812	0.035	0
SP10. Computer	3.33	0.208	0	0.677	0.042	0
SP14. Hobbies/projects	2.505	0.14	0	0.708	0.037	0
SP15. Sports/exercise	3.024	0.148	0	0.765	0.036	0
LS1. Life is close to ideal	1.459	0.114	0	0.424	0.035	0
LS2. Conditions of life excellent	1.123	0.107	0	0.313	0.032	0
LS3. Satisfied with life	0.737	0.077	0	0.237	0.026	0
LS4. Have important things to do in life	1.427	0.113	0	0.516	0.041	0
LS5. Would change nothing	2.883	0.175	0	0.663	0.037	0
SC1. In tune with others	0.33	0.019	0	0.801	0.029	0
SC2. Have people I can talk to	0.143	0.013	0	0.459	0.038	0
SC3. Have people I can turn to	0.141	0.016	0	0.464	0.043	0
SC4. There are people who understand me	0.158	0.011	0	0.446	0.033	0
SC5. There are people I feel close to	0.111	0.009	0	0.42	0.034	0
SC6. Feel part of a group	0.309	0.019	0	0.633	0.034	0
SC7. Have a lot in common with friends.	0.253	0.017	0	0.655	0.036	0
Life satisfaction	1.378	0.126	0	0.695	0.035	0

Social connectedness	0.072	0.011	0	0.874	0.029	0
Participation	0.133	0.034	0	0.565	0.049	0