ISSUES IN EVALUATING IMPORTANCE WEIGHTING IN QUALITY OF LIFE

MEASURES

Chang-ming Hsieh

Jane Addams College of Social Work

University of Illinois at Chicago

1040 West Harrison Street (M/C 309)

Chicago, IL 60607-7134

Email: chsieh@uic.edu

ISSUES IN EVALUATING IMPORTANCE WEIGHTING IN QUALITY OF LIFE

MEASURES

Introduction

As a critical component of quality of life (QoL), life satisfaction has received wide

attention in social science research (e.g., Campbell, Converse, & Rogers, 1976; Diener, 1984;

Feist et al., 1995; George & Bearon, 1980; Myers & Diener, 1995). There have been two major

approaches to conceptualizing life satisfaction: the so-called "top-down" and "bottom-up"

approaches (e.g., Diener, 1984). The top-down approach assumes that global satisfaction is a

predispositional trait or personality which influences one's evaluation of satisfaction in various

areas of life. The bottom-up approach, on the other hand, conceptualizes global satisfaction as

being influenced by one's evaluation of satisfaction in various life domains (Diener, 1984).

These two distinct approaches parallel the reflective-indicator and formative-indicator

measurement models (Bollen & Lennox, 1991). Although many popular life satisfaction

measures (e.g., Life Satisfaction Index A, LSIA, by Neugarten et al., 1961; Satisfaction with Life

Scale, SWLS, by Diener et al., 1985) follow the reflective-indicator model, life satisfaction can

also be measured by combining satisfaction evaluations across discrete life domains (e.g.,

Campbell et al., 1976) which coincides with the bottom-up approach as well as the formative-

indicator model.

However, it has long been observed that the correlation between life satisfaction or QoL

measures using the reflective-indicator model and the formative-indicator model is far from

perfect (e.g., Campbell et al., 1976; Cummins, 1995, 1996; Cummins et al., 1994; Hsieh, 2003,

2004; Russell & Hubley, 2005; Russell et al., 2006). One factor that may contribute to the

discrepancy in correlation between QoL measures based on the two different measurement

approaches has to do with the concept of valence, or the potential inter-person differences in perceived importance of different life domains (Hsieh, in press b). For QoL measures, especially life satisfaction measures based on the formative-indicator model, it has been a common practice to use a simple sum of satisfaction scores across various life domains to represent one's global life satisfaction (e.g., Campbell et al., 1976; Beatty & Tuch, 1997; Mookherjee, 1992). This summed-domains practice does not take into account the potential inter-person differences in perceived importance of different life domains. Researchers (e.g., Campbell et al., 1976; Inglehart, 1978) have long noted the possibility of differential importance of different life domains in the overall picture of QoL. Different terms have been used to reflect the concept, such as (the most popular one) "domain importance" (Campbell et al., 1976), "value priority" (Inglehart, 1978), and "psychological centrality" (Ryff & Essex, 1992).

Although there appears to be a consensus that domain importance is not unimportant, debates over the need to incorporate domain importance, otherwise known as importance weighting, into measures of global life satisfaction remain (e.g., Campbell et al., 1976; Hsieh, 2003, 2004; Mastekaasa, 1984; Russell et al., 2006; Wu, 2008a, 2008b; Wu & Yao, 2006a, 2006b, 2007). Importance weighting is generally accomplished in one of three ways: One approach is to assign externally determined weights to various domains on a theoretical or conceptual basis. Although these weights reflect the concept of importance, the weights are applied to all respondents without taking into account any potential individual differences. A second approach is to consider regression coefficients as weights in a regression analysis when individual domain satisfaction scores are the predictor variables and a global QoL measure is the criterion variable. The results of this approach are heavily dependent on the criterion variable chosen and also do not take into consideration individual respondent differences. A third

approach is to apply different domain importance weights to different respondents. This approach of importance weighting takes into account the potential individual differences in perceived importance of various domains and is the focus of this article.

Much of the controversy surrounding importance weighting has to do with existing empirical evidence. Specifically, many studies found that domain importance as a weighting factor failed to show any detectable increase in the power to explain variations in global QoL measures, in comparison with a simple sum of domain satisfaction scores (e.g., Campbell et al., 1976; Russell et al., 2006; Wu, 2008a, 2008b). For many who chose not to incorporate importance weighting, the decision was not conceptually based but empirically oriented. Opponents of importance weighting include those who purported that importance weighting was conceptually appealing. For example, in their landmark study, Campbell et al. (1976) attempted a number of different scoring systems and techniques of analysis only to find domain importance rating as a weighting factor failed to improve the total variances explained between global QoL instruments and domain satisfaction composites. Based on their empirical findings, Campbell et al. (1976) argued against the use of importance weighting due to the lack of empirical evidence to support a significant effect of domain importance as a weighting factor. Despite the findings, the role of domain importance in QoL measures continued to be as a study topic for many researchers (e.g., Hagerty & Land, 2007; Hagerty et al., 2001; Hsieh, 2003, 2004; Rojas, 2006; Russell & Hubley, 2005; Russell et al., 2006).

For most empirical research investigating the topic of importance weighting, the prevailing approach has been to use 1) a limited choice of global QoL measures as criterion variables (often a single one) to determine the performance of importance weighting, 2) a limited option of weighting methods to develop importance weighting, and 3) a limited number of

domains to construct the (formative-indicator) measures (Hsieh, 2003; 2004; in press b; Russell et al., 2006). Examples of applying this prevailing approach to examine the issue of importance weighting in the literature are easy to find. For instance, in arguing against importance weighting, Wu and Yao (2006a) examined the performance of importance weighting, based on a single global QoL criterion variable, measured by the Satisfaction with Life Scale (Diner et al., 1985), a small number of weighting methods and the 12 domains from the World Health Organization Quality of Life questionnaire (WHOQOL-100). Other examples following this similar approach to examine the issue importance weighting are abundant in the literature (e.g., Philip et al., 2009; Wu, 2008a, 2008b; Wu, Chen, & Tsai, 2009; Wu & Yao, 2007).

However, it is important to note that there are at least three potential issues associated with the prevailing approach described above. Without properly acknowledging and addressing these issues, study findings on importance weighting may be interpreted incorrectly. Although two of these issues have been recognized previously, a third issue has not received sufficient attention in the literature. More specifically, only a small amount of studies recognized that using a limited number of domains to construct formative-indicator measures of QoL might influence the findings and our understanding of the topic of importance weighting (Hsieh, 2003, 2004, in press a, b). Unfortunately, none of these studies discussed what the impact of using a limited number of domains to construct formative-indicator measures of QoL on importance weighting might be. The purpose of this article is to first provide an overview of the two issues that have been recognized previously and then to discuss in greater detail using a limited number of domains to construct formative-indicator measures of QoL with empirical evidence. By focusing on the issue that has been largely neglected in previous studies, this article helps to bridge the gap in the literature on QoL research.

Issue One: Limited Choice of Criterion Variable

A discussion of the potential problem of using a limited choice of global QoL measures as criterion variables to determine the performance of importance weighting could be found in the study by Russell et al. (2006). Research on QoL is an area that keeps growing, and new measures of QoL continue to be developed (Evans, 1994; Hagerty et al., 2001; Land, 2000). There is no single agreed-upon "best" QoL measure. Although it is reasonable to assume that scores from different QoL measures should, to some degree, overlap, there is no reason to assume that all QoL measures capture identical constructs (Russell & Hubley, 2005; Russell et al., 2006). The degree of overlap between QoL measures is likely to vary. Therefore, the association between any given two measures of QoL is unlikely to be the same. Varying degrees of discretion in correlation between different measures of QoL based on different conceptual and measurement approaches should be expected. As Russell et al. (2006) indicated, using a limited choice of global QoL measures as criterion variables (especially in the situation only a single criterion variable) to determine the performance of importance weighting could be problematic since there is no single QoL measure that is universally agreed-upon to be the criterion of QoL. Given the performance of importance weighting is likely to be related to the criterion variables chosen, considering multiple criterion variables would be beneficial as it can help to avoid potential judgment or misjudgment of the performance of importance weighting based on an arbitrary selection of a single criterion. More specifically, caution must be taken in interpreting findings from studies using a single QoL measure as the criterion variable to evaluate the performance of importance weighting. Any positive or negative findings from these studies should only be considered as the performance of importance weighting in relation to the particular criterion variable chosen and should not be generalized to any other QoL measures.

Issue Two: Limited Choice of Weighting Methods

The potential problem of using a limited choice of weighting methods to evaluate the performance of importance weighting has been pointed out previously (e.g., Hagerty et al., 2001; Hsieh, 2003, 2004, in press a, b). The weighting methods applied so far in the literature on importance weighting have mostly centered on the assumption of a linear function of domain importance scores (Hsieh, in press b). For example, surveying a group of injection drug users, Russell et al. (2006) evaluated importance weighting using the Injection Drug User Quality of Life Scale (IDUQOL) and found that scores of IDUQOL weighted by importance rating did not perform better than unweighted scores in measuring QoL. Using a 15-item satisfaction and importance measure, Wu and Yao (2006b) found that in a sample of undergraduate students, importance rating as a weighting mechanism did not improve the correlation between composite of domain satisfaction the Satisfaction with Life Scale (Diner et al., 1985). Both Wu and Yao's (2006b) and Russell et al.'s (2006) studies assumed a linear function of domain importance rating. As Hsieh (2003, 2004) showed, the performance of importance weighting was dependent upon the weighting approaches applied. There is no reason to believe that the function of importance (weighting) must be a simple linear one. Decades ago, Campbell et al. (1976) suggested the possibilities of various types of weighting function of domain importance, such as "hierarchy of needs," "threshold," and "ceiling." The "hierarchy of needs" approach suggests certain kinds of needs are more essential than others, and one would then expect that unless these most essential domains are reasonably satisfied, it probably does not matter much in terms of overall life satisfaction what happens in the less essential domains. The "threshold" approach suggests that overall life satisfaction depends on the presence of some threshold number of satisfactions. If the number of domains with which are satisfied does not meet this threshold, it is

suggested the person would not feel satisfied with life as a whole. The "ceiling" approach

suggests there is a top limit or ceiling to the number of domain satisfactions experienced by an

individual, and satisfaction with domains beyond the top limit would not produce increased

satisfaction with life as a whole. The actual weighting function of domain importance is unclear

and is not necessarily limited to the functions discussed by Campbell et al. (1976). More

recently, Rojas (2006) detailed the limitations and implications of an assumed linear relationship

between global life satisfaction and domain satisfaction. As Rojas (2006) suggested, assuming a

linear relationship between global life satisfaction and domain satisfaction would result in a loss

of many features of the relationship and would clearly restrict our understanding of QoL.

<div align="center">Issue Three: Limited Number of Domains</div>

The impact of using a limited number of domains to construct formative-indicator

measures of QoL has not received much attention in the literature on importance weighting

(Hsieh, 2003; 2004). As indicated by Cummins (1996), there are at least 173 different domain

names that have been used in the QoL literature, and the possible number of domains is even

larger. To cover all possible domains of life in a single QoL measure is extremely unlikely, if not

impossible. The number of domains covered in different QoL measures varies, depending on the

focus of the measure. For example, the Quality of Life Index (QLI) by Ferrans and Powers

(1985) covers the domains of health and functioning, social and economic,

psychological/spiritual, and family. Mookherjee (1992), on the other hand, chose to cover the

domains of residence, networking activities, family life, friendship, health and physical

conditions, and financial situation. Since most QoL measures based on the formative-indicator

model do not, or cannot, exhaustively cover all possible domains, it seems unrealistic to expect a

perfect correlation between any given two QoL measures. In fact, given the countless potential of

the number of domains that QoL measures can cover and the fact that many domains appear to

overlap (Martin, in press), efforts to organize domains into a manageable number for efficiency

purpose were not unreasonable (Cummins, 1996). However, it is important to note that

depending on the measurement model, not covering all possible domains in a measure can have

different consequences. For measures that follow the reflective-indicator model, items or

indicators in the measure should be interchangeable. Items or indicators should also have the

same or similar content or share a common theme. Therefore, comprehensiveness of the domains

covered in the measure is often not a major concern, given dropping items or indicators should

not alter the construct measured. For measures that follow the formative-indicator model, items

or indicators in the measure do not need to be interchangeable. Items or indicators also do not

need to have the same or similar content nor share a common theme. Therefore, the extent to

which the domains covered in the measure may be a cause for major concern, since not including

or dropping indicators may alter the construct measured (e.g., Bollen & Lennox, 1991;

Diamantopoulos, 2006; Diamantopoulos et al., 2008; Jarvis et al., 2003; MacKenzie et al., 2005;

Hardin et al., 2011). Unfortunately, how the lack of comprehensiveness of domains included in

QoL measures based on a formative-indicator model may influence the findings of importance

weighting has not received much attention in literature on importance weighting in OoL

measures. It is argued here that the lack of comprehensiveness of domains included in QoL

measures using a formative-indicator model can lead to misleading results in evaluating the

performance of importance weighting. In particular, the common analytical methods of

evaluating the performance of importance weighting, correlation and moderated regression, can

produce misleading results due to the lack of comprehensiveness of domains included in QoL

measures using a formative-indicator model. To demonstrate what's argued here, an empirical

example that shows the consequences of using a limited number of domains to construct the

(formative-indicator) QoL measures in relation to the issue of importance weighting is presented

below.

An Empirical Example

Data analyzed in this study came from a survey of 112 frail older adults who were the

recipients of elderly case management services in a large city in the Midwest region of the

United States (see Hsieh [2006] for details). The case management unit provides case

management to persons age 60 or older and in need of in-home services. Clients who scored

lower than 21 on the Mini-Mental State Examination (MMSE; Folstein et al., 1975) were also

excluded to avoid any potential problems that were the result of cognitive impairment. A trained

graduate research assistant set up interview appointments, and face-to-face structured interviews

were conducted by the research assistant at the homes of the participants. Most of the study

participants were female (81%) and African American (92%). The mean age of the study

participants was 76.4 ($SD = 7.3$), ranging from 62 to 94. The mean years of schooling completed

were 9.8 ($SD = 3.0$), ranging from 2 to 16. Most of them were retired (96%) and had an annual

household income below $15,000 (90%).

Domain satisfaction and importance. Participants were asked to rate their domain life

satisfaction as: completely satisfied (coded 7), somewhat satisfied (6), slightly satisfied (5),

neither satisfied nor dissatisfied (4), slightly dissatisfied (3), somewhat dissatisfied (2), or

completely dissatisfied (1). A total of seven life domains were included: (a) health, (b) spare

time, (c) financial situation, (d) neighborhood, (e) family life, (f) friendships, and (g) religion.

Satisfaction with each domain was measured by responses to a single question. These seven

questions were derived from Campbell et al.'s (1976) measures. Domain importance was

measured by the same approach as Campbell et al. (1976). That is, participants were asked to

rate the importance of each of the seven domains as: not at all important (coded 1), not too

important (2), somewhat important (3), very important (4), or extremely important (5).

Global satisfaction criterion variable. The Life Satisfaction Index A (LSIA) by

Neugarten, Havighurst and Tobin (1961) was used to measure global life satisfaction. The LSIA

is a 20-item measure designed to assess psychological well-being among older adults. This

measure focuses on respondents' level of happiness and satisfaction with their present lives in

general in comparison with the past and with lives of others. Following the scoring method

proposed by Wood et al. (1969), two points were assigned for each positive answer, 0 for each

negative answer, and one for each "don't know" answer. Previous studies reported that the

internal consistency reliability of LSIA ranged from .66 (Serrano et al., 2004) to .80 (Wagnild,

2003). The coefficient (Cronbach's) alpha was .77 for the current sample.

In order to show the potential consequences of evaluating the performance of importance

weighting in the situation where a formative-indicator measure with non-comprehensive domains

was used, this analysis made an assumption that the seven domains included in the measure were

comprehensive. The impacts of non-comprehensive domains on importance weighting could

then be examined by constructing QoL scores with non-comprehensive domains (with six or less

domains). Although a large number of QoL scores with non-comprehensive domains could be

constructed (ranging from six domains to one domain) in the current case, results presented here

were based on six and five domains. The impacts of non-comprehensive domains on importance

weighting could be sufficiently demonstrated by the results with one-domain reduced (i.e., six

domains) and two-domain reduced (i.e., five domains) measures. Two popular approaches used

to evaluate the performance of importance weighting (Hsieh, in press a) were employed to

explore the potential consequences of a formative-indicator measure with non-comprehensive domains.

First, a correlation analysis was used. The correlation analysis has been a popular approach to the evaluation of importance weighting (e.g., Hsieh, 2003, 2004; Russell & Hubley, 2005; Russell et al., 2006). In general, the performance of importance weighting is assessed by comparing the correlation between a criterion variable and the weighted score and that between the same criterion variable and the unwieghted score, and a higher correlation is preferred. Based on the assumption made in this study, correlations between the scores with comprehensive and non-comprehensive domains would represent the impacts of measures with non-comprehensive domains. Both weighted and unweighted scores were produced for those with comprehensive and non-comprehensive domains. The unweighted scores were the simple sums of satisfaction scores across domains, while the weighted scores were the sums of the products of satisfaction scores (S) and importance scores (I) divided by the sum of importance scores ($\sum S \times I / \sum I$) across domains. The unweighted full-domain score would represent the "real" QoL score under the assumption that global QoL score was a simple sum of domain satisfaction scores and importance weighting was unnecessary. On the other hand, the weighted full-domain score would represent the "real" QoL score under the assumption that importance rating served as a weighting factor in a linear fashion linking domain satisfaction to global QoL. Two types of reduced-domain scores were produced: one-domain reduced scores and two-domain reduced scores. One-domain reduced scores were the weighted and unweighted scores with six domains: DI_{-a} to DI_{-g} and S_{-a} to S_{-g}. The subscript indicated the domain eliminated, so S_{-a} was the unweighted score with domain (a), health, eliminated, while DI_{-g} was the weighted score with domain (g), religion, eliminated. Two-domain reduced scores were the weighted and unweighted

scores with five domains: DI_{-ab} to DI_{-fg} and S_{-ab} to S_{-fg}. The subscript indicated the domains

eliminated, so S_{-ab} was the unweighted score with domains (a) and (b), health and spare time,

eliminated, while DI_{-fg} was the weighted score with domains (f) and (g), friendship and religion,

eliminated. The correlations between unweighted satisfaction score with all domains (S) and

unweighted satisfaction scores with reduced domains (S_{-a} - S_{-g} and S_{-ab} - S_{-fg}) and weighted

satisfaction score with all domains (DI) and weighted satisfaction scores with reduced domains

(DI_{-a} - DI_{-g} and DI_{-ab} - DI_{-fg}) were calculated.

Second, moderated regression analysis was also used to illustrate the impact of non-

comprehensive domains on the performance of importance weighting. The analysis followed the

three-step process outlined by Evans (1991). In general, the analysis would begin by estimating a

regression model with global QoL as dependent variable and domain satisfaction of all the

domains together as independent variables. The second step would be to add domain importance

of all domains as independent variables. The third step would be to add to the second step the

interaction/product terms of domain satisfaction by importance for all domains as independent

variables (see Hsieh, in press a for detail). The change in R^2 from step two to step three would

indicate the need for the inclusion of the interaction (i.e. domain importance weighting) terms. A

significant increase in R^2 (indicated by the by F-test) for the importance by satisfaction terms

would indicate the need for including the interaction terms, and therefore, support importance

weighting (see Evans, 1991 for detail). For the current study, the dependent variable was the

weighted score with all seven domains. This weighted full-domain score would represent the

accurate QoL score, assuming the weighting function was correct. The impacts of non-

comprehensive domains were demonstrated by excluding one or two domains of the independent

variables. Following the same approach by Hsieh (in press a), the importance variables were

constructed to reflect the weighting method (dividing the importance score of each domain by sum of the importance scores across all domains included, or $I_i/\sum I_i$).

Table 1 showed the correlations between scores with reduced domains, with full domains and LSIA scores. The correlation results for all one-domain reduced scores were presented. However, rather than presenting correlation results for all 28 two-domain reduced scores, Table 1 showed the correlations for selected results that illustrated the impacts of non-comprehensive domains in QoL measures. Based on the assumption made earlier that the seven domains were comprehensive, both weighted and unweighted full-domain scores produced with all seven domains should represent the "real" QoL scores and should be the actual standards to be used for evaluating the performance of importance weighting. As shown in the left-hand columns in Table 1, a higher correlation between a one-domain reduced unweighted score and LSIA score did not necessarily mean a higher correlation between the one-domain reduced unweighted score and the full-domain unweighted score. For example, correlation between S_{-g} and LSIA was .62 and between S_{-d} and LSIA was .63, but correlation between S_{-g} and the full-domain score was 1 and between S_{-d} and the full-domain score was .98. The same could be found in the case of two-domain reduced unweighted scores. A higher correlation between a two-domain reduced unweighted score and LSIA score did not necessarily mean a higher correlation between the two-domain reduced unweighted score and the full-domain unweighted score. For example, correlation between S_{-be} and LSIA was .57 and correlation between S_{-cd} and LSIA was .60, but correlation between S_{-be} and full-domain score was .95 and correlation between S_{-cd} and full-domain was .91.

In the right-hand columns in Table 1, the same results could be found with weighted scores. That is, a higher correlation between a one-domain reduced weighted score and LSIA

score did not necessarily mean a higher correlation between the one-domain reduced weighted

score and the full-domain weighted score. For example, correlation between DI_a and LSIA was

.40 and between DI_c and LSIA was .51, but correlation between DI_a and the full-domain score

was .98 and between DI_c and the full-domain score was .88. Similarly, a higher correlation

between a two-domain reduced weighted score and LSIA score did not necessarily mean a higher

correlation between the two-domain reduced weighted score and the full-domain weighted score.

For example, correlation between DI_{ag} and LSIA was .37 and correlation between DI_{cd} and

LSIA was .53, but correlation between DI_{ag} and the full-domain weighted score was .97 and

correlation between DI_{cd} and the full-domain weighted score was .78.

[Table 1 about here]

Next, moderated regression analysis was conducted for the weighted scores. The focus

was to determine the contribution of the importance by satisfaction interaction terms via the

change in R^2 (indicated by the by F-test). Coefficients on specific domain satisfaction, domain

importance, and satisfaction by importance of specific domains were not of interest. Similarly,

rather than presenting a total of 28 regression analysis results, Table 2 showed the R^2 and R^2

changes for selected moderated regression analysis results that illustrated the impacts of non-

comprehensive domains in QoL measures. Given that the dependent variable was the weighted

score, one would expect to see the significant increase in R^2 (indicated by the by F-test) for the

importance by satisfaction terms in order for the moderated regression analysis to produce results

that were not misleading. As shown in Table 2, the regression results for all of the one-domain

reduced composites showed significant R^2 changes from step two to step three. Therefore, with

current data, the moderated regression approach appeared to produce correct results in situations

where only one domain was missing. However, as shown in Table 2, the regression results did

not show significant R^2 changes from step two to step three for all of the two-domain reduced composites. The results indicated that when excluding financial situation (domain c) and family life (domain e), the moderated regression analysis produced an R^2 change from step two to step three that was not significant. Therefore, in situation where two domains were missing, the moderated regression approach could produce results that were misleading.

[Table 2 about here]

Summary and Discussion

In summary, our understanding of the topic of importance weighting in OoL measures has been limited by the use of 1) a limited choice of global QoL measures as criterion variables (often a single one) to determine the performance of importance weighting, 2) a limited option of weighting methods to develop importance weighting, and 3) a limited number of domains to construct the (formative-indicator) measures. Although limitations resulted from a limited choice of global QoL measures as criterion variables to determine the performance of importance weighting and a limited option of weighting methods to develop importance weighting have been recognized previously (e.g., Hsieh, 2003, 2004; Russell & Hubley, 2005; Russell et al., 2006), little attention has been paid to the impact of non-comprehensive domains in QoL measures constructed based on the formative-indicator approach.

Using empirical data, this article revealed the potential impacts of non-comprehensive domains on the evaluation of importance weighting in QoL measures. Both of the two most popular methods of evaluating the performance of importance weighting in QoL measures, correlation and moderated regression analysis, could produce misleading results in the situation when QoL measures constructed using the formative-indicator approach did not include comprehensive domains. As shown earlier, the general practice of preferring higher correlation

was not necessarily the best practice for evaluating the performance of importance weighting in

QoL measures, particularly when the measures were constructed using the formative-indicator

approach with noncomprehensive domains. The empirical example provided in this article

revealed that a higher correlation between an arbitrarily chosen QoL criterion variable and either

a weighted or unweighted non-comprehensive domain QoL score did not necessarily mean a

higher correlation between the real weighted or unweighted full-domain QoL score and a

weighted or unweighted non-comprehensive domain QoL score. In other words, using

correlation as a method to evaluate the performance of importance weighting in QoL measures

could be problematic when QoL measures constructed using the formative-indicator approach

did not include comprehensive domains. In addition, empirical data presented in this article

uncovered that results of moderated regression analysis would not necessarily capture the

significant effect of the importance by satisfaction interaction terms in the situation when QoL

measures were constructed using the formative-indicator approach with non-comprehensive

domains. That is, even when importance weighting served as the actual function of linking

domain satisfaction across various domains to global QoL, it was possible that results of

moderated regression analysis would suggest otherwise in the situation when QoL measures

were constructed using the formative-indicator approach with non-comprehensive domains. It

should be noted that findings in this article could also be interpreted based on Hagerty and

Land's (2007) work on the conditions under which the use of importance weights could improve

the accuracy of composite QoL indices. As demonstrated by Hagerty and Land (2007), a

statistical model for weighted QoL measures had two major properties: (a) an optimal composite

QoL index could be achieved by using the population or sub-population average importance

weights, and (b) in the absence of good information on domain importance, an equal-weighting

method could be used and it was a so-called minimax estimator, as it minimized extreme

disagreements on a composite QoL index. Based on the properties shown in Hagerty and Land's

(2007) study, the weighting approach (although assumed to be correct) used in this article would

be unlikely to produce an optimal composite.

Issues discussed in this article are of great importance to research in the field of QoL,

especially on the topic of importance weighting in QoL measures. As Hsieh (in press a)

indicated, much of the argument against importance weighting was based on the empirical results

of either correlation or moderated regression analysis (e.g., Wu & Yao, 2006a, 2006b).

Unfortunately, most empirical evidence used to argue against importance weighting has not fully

considered or acknowledged the potential impacts related to the issues discussed in this article,

especially the issue regarding the limit number of domains included in the QoL measures.

Therefore, it would be prudent to recognize that results from previous studies on importance

weighting in QoL measures must not be interpreted without considering the issues discussed in

this article.

References

Beatty, P. and S. A. Tuch: 1997, 'Race and life Satisfaction in the middle class', Sociological Spectrum 17, pp. 71-90.

Bollen, K. and R. Lennox: 1991, 'Conventional wisdom on measurement: A structural equation perspective', Psychological Bulletin 110, pp. 305-314.

Campbell, A., Converse, P. E., and W. L. Rogers: 1976, The Quality of American Life: Perceptions, Evaluations, and Satisfactions (Russel Sage, New York).

Cummins, R. A.: 1995, 'On the tale of gold standard for life satisfaction', Social Indicators Research 35, pp. 179-200.

Cummins, R. A.: 1996, 'The domains of life satisfaction: An attempt to order chaos', Social Indicators Research 38, pp. 303-328.

Cummins, R. A., M. P. McCabe, Y. Romeo and E. Gullone: 1994, 'The comprehensive quality of life scale: Instrument development and psychometric evaluation on tertiary staff and students', Educational and Psychological Measurement 54, pp. 372-382.

Diamantopoulos, A. :2006, 'The error term in formative measurement models: Interpretation and modeling implications', Journal of Modeling in Management 1, pp. 7-17.

Diamantopoulos, A., Riefler, P., and K. P. Roth: 2008, 'Advancing formative measurement models', .Journal of Business Research 61, pp. 1201-1302.

Diener, E.: 1984, 'Subjective well-being', Psychological Bulletin 95, pp. 542-575.

Diener, E., Emmons, R.A., Larsen, R.J. and S. Griffin: 1985, 'The satisfaction with life scale', Journal of Personality Assessment 49, pp. 71–74.

Evans, D. R.: 1994, 'Enhancing quality of life in the population at large', Social Indicators Research 33, pp. 47–88.

Evans, M.G.: 1991, 'The problem of analyzing multiplicative composites: Interactions revisited',

American Psychologist 46, pp. 6–15.

Ferrans, C., and M. Powers: 1985, 'Quality of Life Index: Development and psychometric

properties', Advances in Nursing Science 8, pp. 15-24.

Feist, G. J., Bodner, T. E., Jacobs, J. F., Miles, M., and V. Tan: 1995, 'Integrating top-down and

bottom-up structural models of subjective well-being: A longitudinal investigation',

Journal of Personality and Social Psychology 68, pp. 138-150.

Folstein, M. F., Folstein, S. E., and P. R. McHugh: 1975, 'Mini-Mental State: A practical method

for grading the cognitive state of patients for the clinician', Journal of Psychiatric

Research 12, pp. 189-198.

George, L. K., and L. B. Bearon: 1980, Quality of Life in Older Persons: Meaning and

Measurements (Human Sciences Press, New York).

Hagerty, M. R., Cummins, R. A., Ferris, A. L., Land, K. C., Michalos, A. C., Peterson, M.,

Sharpe, A., Sirgy, J., and J. Vogel: 2001, 'Quality of life indexes for national policy:

Review and agenda for research', Social Indicators Research 55, pp. 1-96.

Hagerty, M. R., and K. C. Land: 2007, 'Constructing summary indices of quality of life: A

model for the effect of heterogeneous importance weights', Sociological Methods and

Research 35, pp. 455-496.

Hardin, A. M., Chang, J. C., Fuller, M. A. and G. Torkzadeh: 2011, 'Formative measurement and

academic research: In search of measurement theory', Educational and Psychological

Measurement 71, pp. 281-305.

Hsieh, C. M.: 2003, 'Counting importance: The case of life satisfaction and relative domain

importance', Social Indicators Research 61, pp. 227-240.

Hsieh, C. M.: 2004, 'To weight or not to weight: The role of domain importance in quality of life measurement', Social Indicators Research 68, pp. 163-174.

Hsieh, C. M.: 2006, 'Using client satisfaction to improve case management services for the elderly', Research on Social Work Practice 16, pp. 605-612.

Hsieh, C. M.: In press a, 'Importance is not unimportant: The role of importance weighting in QoL measures', Social Indicators Research.

Hsieh, C. M.: In press b, 'Should we give up domain importance weighting on QoL measures?', Social Indicators Research.

Inglehart, R.: 1978, 'Value priorities, life satisfaction, and political dissatisfaction among western publics', Comparative Studies in Sociology 1, pp. 173-202.

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

Land, K.: 2000, 'Social indicators', in E. F. Borgatta and R. V. Montgomery (eds.), Encyclopedia of Sociology, Revised Edition (New York: Macmillan).

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

Martin, F.: In press, 'Perceptions of links between quality of life areas: Implications for measurement and practice', Social Indicators Research.

Mastekaasa, A.: 1984, 'Multiplicative and additive models of job and life satisfaction', Social Indicators Research 14, pp. 141-163.

Mookherjee, H. N.: 1992, 'Perceptions of well-being by metropolitan and nonmetropolitan populations in the United States', Journal of Social Psychology 132, pp. 513-524.

Myers, D. G., and E. Diener: 1995, 'Who is happy?', Psychological Science 6, pp. 11-19.

Neugarten, B., Havighurst, R., and S. Tobin: 1961, 'The measurement of life satisfaction', Journal of Gerontology 16, pp. 134-143.

Philip, E. J., Merluzzi, T. V., Peterman, A., and L. B. Cronk: 2009, 'Measurement accuracy in assessing patient's quality of life: To weight or not to weight domains of quality of life', Quality of Life Research 18, pp. 775-782.

Rojas, M.: 2006, 'Life satisfaction and satisfaction in domains of life: Is it a simple relationship?' Journal of Happiness Studies 7, pp. 467-497.

Russell, L. B. and A. M. Hubley: 2005, 'Importance ratings and weighting: Old concerns and new perspectives', International Journal of Testing 5, pp. 105-130.

Russell, L. B., A. M. Hubley, A. Palepu and B. D. Zumbo: 2006, 'Does weighting capture what's important? Revisiting subjective importance weighting with a quality of life measure', Social Indicators Research 75, pp. 146-167.

Ryff, C. D. and M. J. Essex: 1992, 'The interpretation of life experience and well-being: The sample case of relocation', Psychology and Aging 7, pp. 507-517.

Serrano, J. P., Latorre, J. M., Gatz, M., and J. Montanes: 2004, 'Life review therapy using autobiographical retrieval practice for older adults with depressive symptomatology', Psychology and Aging, 19, pp. 272-277.

Wagnild, G.: 2003, 'Resilience and successful aging: Comparison among low and high income older adults', Journal of Gerontological Nursing 29(12), pp. 42-49.

Wood, V., Wylie, M., and B. Sheafor: 1969,.'An analysis of a short self-report measure of life satisfaction: Correlation with rater judgments', Journal of Gerontology 24, pp. 456-469.

Wu, C. H.: 2008a, 'Examining the appropriateness of importance weighting on satisfaction score from range-of-affect hypothesis: Hierarchical linear modeling for within-subject data', Social Indicators Research 86, pp. 101–111.

Wu, C. H.: 2008b, 'Can we weight satisfaction score with importance ranks across life domains?', Social Indicators Research 86, pp. 468–480.

Wu, C. H., Chen, L. H. and Y. M. Tsai: 2009, 'Investigating importance weighting of satisfaction scores from a formative model with partial least squares analysis', Social Indicators Research 90, pp. 351-363.

Wu, C. H. and G. Yao: 2006a, 'Do we need to weight satisfaction scores with importance ratings in measuring quality of life?', Social Indicators Research 78, pp. 305–326.

Wu, C. H. and G. Yao: 2006b, 'Do we need to weight item satisfaction by item importance? A perspective from Locke's range-of-affect hypothesis', Social Indicators Research 79, pp. 485–502.

Wu, C. H. and G. Yao: 2007, 'Examining the relationship between global and domain measures of quality of life by three factor structure models', Social Indicators Research 84, pp. 189-202.

Table 1

Correlations Between Full-domain, Reduced-domain and LSIA Scores[a]

Full-domain Unweighted		LSIA	Full-domain Weighted		LSIA
	r	R		r	R
One-domain Reduced			One-domain Reduced		
S_{-a}	.94	.56	DI_{-a}	.98	.40
S_{-b}	.97	.60	DI_{-b}	.98	.42
S_{-c}	.94	.60	DI_{-c}	.88	.51
S_{-d}	.98	.63	DI_{-d}	.96	.46
S_{-e}	.98	.60	DI_{-e}	.96	.43
S_{-f}	1	.63	DI_{-f}	.99	.46
S_{-g}	1	.62	DI_{-g}	.99	.43
Two-domain Reduced			Two-domain Reduced		
S_{-ab}	.90	.52	DI_{-ab}	.95	.35
S_{-ag}	.94	.56	DI_{-ag}	.97	.37
S_{-be}	.95	.57	DI_{-be}	.93	.38
S_{-cd}	.91	.60	DI_{-cd}	.78	.53
S_{-ce}	.92	.58	DI_{-ce}	.78	.48

Note: [a] $\underline{n} = 110$

Table 2

Results of Moderated Regression Analysis with Reduced Domains Predicting Weighted QoL

Scores with Full Domains

	Step 1	Step 2	Step 3
	Satisfaction R^2	Importance ΔR^2	Satisfaction x Importance ΔR^2
One-domain Reduced			
DI_{-a}	.42***	.52***	.03***
DI_{-b}	.42***	.51***	.04***
DI_{-c}	.38***	.43***	.02*
DI_{-d}	.34***	.53***	.07***
DI_{-e}	.39***	.49***	.04***
DI_{-f}	.42***	.51***	.04***
DI_{-g}	.42***	.53***	.04***
Two-domain Reduced			
DI_{-ab}	.42***	.48***	.03***
DI_{-ac}	.36***	.39***	.03*
DI_{-cd}	.28***	.36***	.06**
DI_{-ce}	.35***	.32***	.03

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