Throwing the Baby Out with the Bathwater: Evaluation of Domain Importance Weighting in Quality of Life Measurements

Chang-ming Hsieh

Abstract This paper examined major inconsistencies between claims made and the evidence presented in recent studies on evaluating the role of domain importance weighting in quality of life (QoL) measures. Three inconsistencies were discussed and they were: (1) treating the failure to support a particular importance weighting function as the evidence of uncovering no importance weighting necessary at all, (2) considering domain importance weighting and multiplicative scores synonymous, and (3) extending findings with a within-domain focus to an across-domain focus. Overlooking these inconsistencies may lead to an overgeneralization of study findings, which would likely result in premature or even erroneous conclusions. Caution must be given in interpreting study results in order to avoid overgeneralization which limits our understanding of the true role that domain importance weighting may play in QoL measures.

Keywords Domain importance · Relative importance · Importance weighting · Domain weighting · Relative weighting · Multiplicative scores

1 Introduction

The topic of domain importance, or relative importance, weighting has been a research area of quality of life (QoL) studies for decades (e.g., Lix et al. 2013; Russell and Hubley 2005), especially in subjective well-being research (e.g., Hsieh 2013). Although domain importance weighting can be accomplished in different ways (Hsieh 2012a), debates have mostly focused on the need for incorporating domain importance into QoL measures at the individual level (e.g., Campbell et al. 1976; Hsieh 2003, 2004, 2012a, b, 2013; Rojas 2006; Russell and Hubley 2005; Russell et al. 2006; Wu 2008a, b; Wu and Yao 2006a, b, 2007).

C. Hsieh (✉)
Jane Addams College of Social Work, University of Illinois at Chicago,
1040 West Harrison Street (M/C 309), Chicago, IL 60607-7134, USA
e-mail: chsieh@uic.edu

Much of the debate surrounding domain importance weighting has to do with the fact that many studies observed that domain importance as a weighting factor at the individual level failed to show any detectable increase in the power to explain variations in global QoL measures, in comparison with using satisfaction scores only (e.g., Campbell et al. 1976; Russell et al. 2006; Wu 2008a, b).

At the center of the debate on the topic of domain importance weighting is the question whether or not domain importance should be incorporated along with domain satisfaction across multiple domains to form global QoL (Campbell et al. 1976; Hsieh 2012a, b; Russell and Hubley 2005). A key objective of evaluating domain importance weighting in QoL measures is to find out if domain importance plays any role in the relationship between global QoL and the composite of domain satisfactions. More specifically, the issue in question is: Should domain importance be incorporated into satisfaction scores to represent global QoL? In practice, this question turns into an investigation of whether or not using domain importance as a weighting factor can lead to any detectable increase in the power to explain variances in global QoL measures above and beyond using domain satisfaction scores alone (Campbell et al. 1976; Trauer and Mackinnon 2001). Although the investigation asks a "whether" question, there is an implicit "how" aspect to the investigation. That is, without knowing how to weight domain importance, it is unlikely that an adequate assessment of "whether" can be obtained.

Evidence used to argue for and against incorporating domain importance weighting into QoL measures certainly deserves attention. However, just like in all research, it is important to ensure that evidence is interpreted adequately without overgeneralization. Finding evidence for supporting domain importance weighting is relatively straight-forward. To provide support for domain importance weighting, study results should show that at least one type of domain importance weighting function can lead to a detectable increase in the power to explain variances in one or more global QoL measures.

Finding evidence to argue against domain importance weighting completely, on the other hand, is a task that is not so straight-forward because, in principle, the task is to demonstrate no domain weighting function can lead to a detectable increase in the power to explain variances in any global QoL measure. Considering the range of possibilities for various potential domain importance weighting functions and global QoL measures, it is very unlikely, if not impossible, for any single study to complete the task. Under the circumstances, caution is needed in interpreting study results arguing against domain importance weighting, in order to avoid overgeneralization. Unfortunately, when it comes to evaluation of domain importance weighting in QoL measures, evidence, especially in arguing against domain importance weighting, has not always been consistent with study claims. The purpose of this paper is to discuss inconsistencies between the evidence and claims presented in recent studies regarding domain importance weighting in QoL measures. Three issues pertaining to these inconsistencies were examined as follows and special attention, including empirical evidence, was given to the generalizability of evidence from a within-domain perspective to an across-domain perspective.

2 Issue One: Failure to Find a Relationship or Demonstrating No Relationship

Adequate interpretation of evidence and study results concerning evaluating domain importance weighting must consider, among others, the potential limitations associated with the analytical approaches utilized. The common analytical methods used to assess the performance of domain importance weighting have been correlation- or regression-based

(such as correlation, moderated regression, and partial least squares regression), with moderated regression analysis being most popular (e.g., Hsieh 2012a; Mastekaasa 1984; Russell et al. 2006; Wu and Yao 2006a, b). Unfortunately, there are issues related to the adequacy of this analytical approach in assessing domain importance weighting and these issues have not been fully examined (Hsieh 2013).

The use of moderated regression analysis for the evaluation of domain importance weighting, just like for all research, can be problematic if consideration of statistical power is not given. Statistical power refers to the probability that a statistical test correctly rejects the null hypothesis (e.g., Aberson 2010; Cohen 1988). For moderated regression analysis, statistical power is dependent upon, among other factors, sample size (Aberson 2010). Most of the studies on domain importance weighting had sample sizes that were small or moderate. For example, analyzing data from 130 undergraduate students, Wu and Yao (2006a) found that domain importance did not moderate the relationship between domain satisfaction and overall life satisfaction and argued against domain importance weighting. Russell et al. (2006) investigated the topic of domain importance weighting based on a sample of 241 subjects. More recently, Philip et al. (2009) argued against domain importance weighting based on data from 194 cancer patients.

Sample size affects statistical power. With a limited sample size, statistical power is likely to be limited if the effect size (Cohen 1988) to be detected is not large. In other words, unless domain importance as a weighting factor can lead to a relatively large increase in the power to explain variations in global QoL, it is unlikely for studies with limited sample sizes to find any empirical support for domain importance weighting. Evidently, the effect size of domain importance as a weighting factor is unknown. Thus, interpretation of study results that fail to support domain importance weighting should not ignore statistical power. Specifically, failing to find a statistically significant relationship could mean that the magnitude of the relationship (effect size) might be smaller than what the sample size could afford to detect. Should a conclusion of abandoning domain importance weighting be reached because the effect size is not large enough for studies with limited sample sizes to detect? Of course, failing to find a statistically significant relationship could also mean the relationships specified in the studies are inaccurate. As Hsieh (2013) indicated, the empirical evidence presented in studies on domain importance weighting has not gone beyond a linear function of domain importance. Not being able to find a linear relationship does not necessarily mean no relationship. In sum, what is important to note is that unable to find a relationship is a necessary but not sufficient condition to be regarded as the evidence of no relationship at all.

3 Issue Two: Multiplicative Scores or Domain Importance Weighting

Adequate interpretation of evidence and study results concerning evaluating domain importance weighting must also consider the fact that domain importance weighting is not equal to multiplicative scores (scores derived from multiplying satisfaction and importance ratings). Although the use of multiplicative scores as a domain importance weighting approach is not uncommon (e.g., Ferrans and Powers 1985), concerns regarding the use of multiplicative scores have been raised frequently (e.g., Hsieh 2003, 2004; Trauer and Mackinnon 2001). Major problems pertaining to multiplicative scores were eloquently discussed by Trauer and Mackinnon (2001) over a decade ago. One of the obvious problems with multiplicative scores is the lack of clarity in conceptual meaning (Hsieh 2004; Trauer and Mackinnon 2001). More specifically, different combinations of

C. Hsieh

satisfaction ratings and importance ratings can result in same scores and therefore, it is difficult to clearly define the meaning of any multiplicative score. Given the conceptual ambiguity with multiplicative scores, the use of multiplicative scores as an approach of domain importance weighting has not gone without challenges (Hsieh 2004, 2012a, b; Trauer and Mackinnon 2001).

Despite the issues associated with multiplicative scores, studies, especially those presenting empirical evidence against domain importance weighting (e.g., Russell et al. 2006; Wu 2008a, b; Wu and Yao 2006a, b, 2007), have overwhelmingly relied on using multiplicative scores to evaluate domain importance weighting. Domain importance weighting and multiplicative scores, however, are not synonymous (Hsieh 2012a, b). Multiplicative scores, a domain importance weighting approach not without controversy, should be considered only one of the many possible approaches to domain importance weighting (Hsieh 2003, 2004).

According to Hsieh (2003, 2004), the performance of domain importance weighting was dependent upon how importance was measured and how importance was weighted. The popular practice of using rating scales to measure domain importance and using multiplicative scores as weighting is just one of the many possible ways to measure domain importance weighting. As Kaplan et al. (1993)indicated, the use of rating scales was by no means the only method for assessing relative importance. Other methods, including utility assessment, such as the standard gamble, time trade-off and person trade-off methods, as well as economic measurements of choice should not be ignored (Kaplan et al. 1993). In addition to different ways of measuring domain importance, there can be many options other than simply multiplying satisfaction scores by importance scores to the construction of weighting. Campbell et al. (1976), for example, suggested various types of weighting function of domain importance, such as "hierarchy of needs," "threshold," and "ceiling." Of course, the actual weighting function of domain importance is still unclear and is not necessarily limited to the functions discussed by Campbell et al. (1976). Other literature, such as multiple-criteria decision making in operations research, offers implications on this topic as well (e.g., Luque et al. 2009; Marcenaro et al. 2010).

In sum, even if there is no evidence to support domain importance weighting based on multiplicative scores, it is unclear what the evidence would show with other types of weighting function of domain importance. It would be unwise to argue against domain importance weighting completely with evidence based only on multiplicative scores since domain importance weighting can and should mean beyond multiplicative scores.

4 Issue Three: Within Individual Domain or Across Multiple Domains

Nearly all evidence used to argue against domain importance weighting recently (e.g., Russell et al. 2006; Wu 2008a, b; Wu and Yao 2006a, b, 2007; Wu, Yang, and Huang, in press) came from analytical approaches with a within-domain focus (Hsieh 2013). For example, a number of studies used Locke's "range-of-affect" hypothesis (Locke 1969, 1976, 1984) to argue against domain importance weighting in QoL measures (Wu 2008a, b; Wu and Yao 2006a, b, 2007) by examining the relationships between global QoL and weighted and unweighted satisfaction scores of individual life domains. Similarly, a more recent study (Wu et al. in press) assessed importance weighting by comparing the predictive power of weighted (multiplicative) scores and unweighted scores of four (physical health, psychological health, social relationships, and environmental health) domains on global QoL measures. Based on the findings that weighted domain scores did not account

for more variances in the three indices for global subjective well-being than unweighted scores, the authors argued against the use of importance weighting (Wu et al. in press).

In discussing the applicability as well as generalizability of the range-of-affect hypothesis to domain importance weighting in QoL measures from both conceptual and empirical aspects, Hsieh (2012a, b, 2013) pointed out the difference in focus regarding studies that evaluated domain importance weighting. As Hsieh (2013) showed, analytical approaches and empirical results from studies that focus primarily on the relationship between global QoL, domain satisfaction and domain importance within each individual domain (within-domain focus) may not have any direct implication to studies that focus on the relationship between global satisfaction, composite of domain satisfactions and domain importance across domains (across-domain focus). The difference in study focus also means that assumptions used in a within-domain focus context can be problematic in an across-domain focus context, especially the following two assumptions.

One assumption is that if importance weighting is to be supported, then an importance-weighted score should account for more variances (measured by the R^2 value in a regression model) in a global QoL measure than an unweighted score. This assumption is problematic in an across-domain focus context. In general, the weighting function is intended to capture the effect of a particular domain on global QoL *relative to*, not independent of, other domains (hence an across-domain focus). The emphasis, therefore, should be on a weighted or unweighted *composite* (sum or average across domains) score, not scores of various domains individually. It must be noted that the variances explained in a global QoL by a (weighted or unweighted) composite score is not the sum of the variances explained by each domain score individually. In other words, it is not necessary for every (or majority of) weighted domain score to account for more variances in a global QoL measure than the unweighted score of the same domain, in order for a weighted composite score (of all domains) to account for more variances in a global QoL measure than the unweighted composite score.

The other assumption is that if importance weighting is to be supported, then importance-weighted scores should enhance the predictive pattern, in comparison with unweighted scores. This assumption cannot be applied to evaluating domain importance weighting with an across-domain focus. Examining the effect of importance weighting of a single domain (while holding other domains constant) does not provide any information regarding the effect of a particular domain on global QoL *relative to* other domains. Specifically, the purpose of domain importance weighting with an across-domain focus is not to examine the effect of any particular domain on global QoL while controlling for other domains. Instead, the purpose should be to compare the association between global QoL and the weighted composite score and the unweighted composite score. It must be noted that the predictive effect of a weighted or unweighted composite score on global QoL (using either standardized coefficient or R^2 as the standard) is not captured by the predictive pattern of weighted or unweighted score of individual domain on global QoL. In other words, it is not necessary for every (or majority of) weighted domain score to have stronger predictive effect on global QoL than the unweighted score of the same domain in order for a weighted composite score of all domains to have a stronger predictive effect on global QoL than the unweighted composite score.

In sum, evaluating importance weighting within individual domains is different from evaluating importance weighting across domains. Two critical points are worth noting. First, whether or not a weighted composite (across all domains) score accounts for more variances in global QoL than an unweighted composite score cannot be captured by comparing the difference in the variances in global QoL explained by between weighted

and unweighted individual domain scores. Second, results of comparing the predictive patterns between weighted and unweighted individual domain scores on global QoL are not indicative of the results of comparing the predictive effect of domain importance between weighted and unweighted composite scores on global QoL.

5 An Empirical Example

To illustrate the two points regarding within-domain focus versus across-domain focus made above, a re-analysis of survey data from an earlier study was conducted (see Hsieh 2003 for details).

5.1 Participants

Participants of this telephone survey were adults over the age of 50. A total of 100 telephone interviews were conducted in the greater Chicago area in the U.S. After excluding missing data, the sample size was 90. The mean age of the sample was 67 ($SD = 10.3$). The majority of the sample participants were Caucasian (48 %), female (67 %) and retired (46 %). Over one-third (36 %) were married and 34 % were widowed. Most (92 %) had at least a high school education.

5.2 Measures

The global QoL was indicated by a single-item global life satisfaction measure. The measure asked respondents to rate their satisfaction with life as a whole as: completely satisfied (7), somewhat satisfied (6), slightly satisfied (5), neither satisfied nor dissatisfied (4), slightly dissatisfied (3), somewhat dissatisfied (2), or completely dissatisfied (1). Respondents were also asked to rate their satisfaction in the same manner for each of the following eight domains: health, work, spare time, financial situation, neighborhood, family life, friendships, and religion. Respondents were also asked to rank the importance of these eight domains. Ranking was achieved by a two-step process. First, respondents were asked to rate the importance of each of the eight discrete life domains as: not at all important (1), not too important (2), somewhat important (3), very important (4), or extremely important (5). Second, respondents were asked to compare and rank among the domains with same rating scores to obtain a rank ordering of domains. It was possible for domains of equal importance to be so ranked. The most important domain received a rank score of eight (8) and the least important domain received a rank score of one (1).

5.3 Weighting

The weighting function constructed for the current study was based on the rank score. To go beyond multiplicative score, the square of the rank score was used as the weight, based on past study findings (Hsieh 2003, 2004). Therefore, a weighted domain score is the product of domain satisfaction score and the weighting function:

$$D_i R^2 = (S_i) \times \left(R_i^2 / \sum R_i^2 \right) \tag{1}$$

where S_i is the satisfaction rating in domain i and R_i is the rank score of the domain importance of domain i. A weighted composite score is:

$$DR^2 = \sum(S_i) \times \left(R_i^2 / \sum R_i^2\right) \qquad (2)$$

An unweighted domain score, on the other hand, was the satisfaction score for the individual domain, and the unweighted composite score was the sum of satisfaction scores across all eight domains.

5.4 Findings

Table 1 shows the correlation between global life satisfaction and weighted as well as unweighted domain and composite scores. Results shown in Table 1 suggest that correlations between global life satisfaction and weighted individual domain scores consistently lower than unweighted individual domain scores. However, the correlation between global life satisfaction and weighted composite score was higher than unweighted composted score.

Table 2 shows the regression results on global life satisfaction with weighted and unweighted domain and composite scores. The upper panel of Table 2 presents the multiple regression results with weighted and unweighted domain scores. The lower panel of Table 2 presents the bivariate regression results with weighted composite score and unweighted composite score. As shown in the upper panel of Table 2, standardized coefficients on weighted domain scores appeared to be larger than unweighted domain scores. However, the predictive power of weighted domain scores combined, as measured by R^2 ($R^2 = .21$), was actually lower than unweighted domain scores combined ($R^2 = .24$). Results from the lower panel of Table 2 indicated that the predicted power of weighted composite score ($R^2 = .19$) was higher than unweighted composite score ($R^2 = .15$).

6 Implications of Findings

Findings of this study illustrate the issues regarding generalizing importance weighting from a within-domain focus to an across-domain focus. As shown in Table 1, on average, a weighted domain score had a lower correlation with global life satisfaction than an unweighted score of the same domain. However, these results which had a within-domain focus could not be translated into the results with an across-domain focus. The correlation

Table 1 Correlations between global life satisfaction and weighted and unweighted domain and composite scores		Unweighted	Weighted
	Health	.31**	.10
	Work	.23*	.02
	Spare time	.28**	.12
	Financial situation	.27**	.15
	Neighborhood	.16	−.01
	Family life	.33***	.18
	Friendships	.15	.03
	Religion	.20	.07
* $p < .05$, ** $p < .01$, *** $p < .001$	Composite score	.39***	.43***

Table 2 Results of regression analysis (standardized coefficients) with domain and composite scores predicting global satisfaction

	Unweighted	Weighted
Health	.20	.27*
Work	.10	.23
Spare time	.22	.27*
Financial situation	.11	.34**
Neighborhood	−.08	.17
Family life	.32*	.38***
Friendships	−.03	.21
Religion	−.15	.37
R^2	.24	.21
Composite score	.39***	.43***
R^2	.15	.19

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

between global life satisfaction and the weighted composite score was, in fact, higher than the unweighted composite score.

Findings also show that comparing the difference in the variances in global QoL explained between weighted and unweighted individual domain scores does not tell us whether or not a weighted composite score accounts for more variances in global QoL than an unweighted composite score. As shown in Table 2, the weighted domain scores together accounted for about 21 % ($R^2 = .21$) of the variances in global life satisfaction, while the unweighted domain scores together accounted for about 24 % ($R^2 = .24$) of the variances in global life satisfaction. If these results are generalizable to an across-domain focus, one would expect that more variances in global life satisfaction would be explained by the unweighted composite score than the weighted composite score. However, the results shown in Table 2 indicated otherwise. The weighted composite score accounted for about 19 % ($R^2 = .19$) of the variances in global life satisfaction, while the unweighted composite score accounted for only about 15 % ($R^2 = .15$) of the variances in global life satisfaction. These results support the notion that findings on importance weighting with a within-domain focus cannot be generalized to importance weighting with an across-domain focus.

Furthermore, findings show that the predictive effect of a weighted composite score versus an unweighted composite score on global QoL cannot be captured by comparing the predictive patterns between weighted and unweighted individual domain scores on global QoL. Using standardized regression coefficients to compare the predictive pattern of domain importance between weighted and weighted domain scores (as in the study by Wu et al. in press) is somewhat unconventional. The comparability of standardized regression coefficients within a regression model is well established (e.g., Allison 1999), however, the comparability of standardized regression coefficients across separated regression models is unclear. Results shown in the upper panel of Table 2 suggest that in general, weighted domain scores had a stronger predictive pattern than unweighted domain scores because the coefficients were larger in magnitude and there were more significant coefficients for the weighted domain scores. However, this pattern did not translate to the overall predictive power of each model. The overall predictive power, as commonly measured by R^2, of weighted domain scores together was lower ($R^2 = .21$) than unweighted domain score together ($R^2 = .24$). The discrepancy between the overall predictive power of a model and the predictive pattern between weighted and unweighted individual domains called into the

question of the adequacy and validity of the use of standardized coefficients for evaluating domain importance weighting. A more reasonable examination would be to compare the weighted and unweighted composite scores. As shown in the lower panel of Table 2, the standardized coefficient for the weighted composite score ($\beta = .43$) was higher than the unweighted composite score ($\beta = .39$). Given that these are bi-variate regression results, the predictive power (R^2) of each model was the squared value of the standardized coefficient. Therefore, larger value of the standardized coefficient for the weighted composite score also translated to more predictive power ($R^2 = .19$) than the unweighted composite score ($R^2 = .15$). These results further support the notion that findings on importance weighting with a within-domain focus cannot be generalized to importance weighting with an across-domain focus.

7 Summary and Discussion

This paper discussed inconsistencies between the evidence presented and claims made in recent studies regarding domain importance weighting in QoL measures. In order not to overgeneralize study results, claims made for and against domain importance weighting must be supported by sufficient evidence. However, when it comes to the topic of domain importance weighting, what the evidence showed and what the studies claimed have not always been consistent, especially for claims against domain importance weighting. Specifically, there has been a tendency to (1) treat the failure to support a particular importance weighting function as the evidence of uncovering no importance weighting necessary at all, (2) consider domain importance weighting and multiplicative scores synonymous, and (3) extend findings with a within-domain focus to an across-domain focus. Overlooking these inconsistencies may lead to premature conclusion and limit our understanding of the topic of domain importance weighting.

Of these inconsistencies, the issue of generalizing results with a within-domain focus to an across-domain focus has received only limited attention (Hsieh 2012a, b). It is important to note that study results on domain importance weighting should not be interpreted without recognizing the difference in study focus. Domain importance weighting with a within-domain focus assesses the effect of importance weighting for each domain individually. The premise of studies with a within-domain focus is that importance weighting cannot be supported if weighting of each individual domain does not increase the power to explain variances in global QoL (e.g., Wu 2008a, b; Wu and Yao 2006a, b, 2007; Wu et al. in press). This premise can be problematic for two reasons. First, there is no clear conceptual justification to argue that importance weighting must significantly enhance the association between any domain-specific satisfaction/QoL and global QoL. Global QoL, either conceptually or in practice, is constructed by a composite of satisfaction scores across multiple (typically, as comprehensive as possible) domains (Hsieh 2013), not just a single domain. Expecting importance weighting to significantly enhance the association between any domain-specific satisfaction/QoL and global QoL is clearly based on the assumption of a reflective-indicator (top-down) measurement model (e.g., Bollen and Lennox 1991) between global QoL and domain-specific satisfaction/QoL. However, it is likely that global QoL and domain-specific satisfaction/QoL follow a formative-indicator (bottom-up) measurement model (e.g., Diener et al. 1985; Hsieh 2013). Therefore, the expectation should be for domain importance weighting to enhance the relationship between the *composite* of satisfaction scores and global QoL, which is different from expecting importance weighting to enhance the association between a domain-specific

score and global QoL. Second, as demonstrated earlier, findings on domain importance weighting based on the relationship between global QoL and individual domains cannot be extended to the relationship between global QoL and the composite scores across multiple domains. In fact, there is more evidence to support this point. A recent study showed that the effect of domain importance weighting is unlikely to be accurately captured by examining the relationship between global QoL and only a small number of domains (see Hsieh 2013). More specifically, in a formative-indicator measurement model, the relationship between global QoL based on comprehensive domains and global QoL based on limited domains is uncertain (Hsieh and Kenagy in press). Under the circumstances, the relationship between global QoL based on comprehensive domains and domain-specific satisfaction/QoL would be uncertain as well. The premise that importance weighting cannot be supported if weighting of each individual domain does not increase the power to explain variances in global QoL is, thus, problematic.

In conclusion, caution should be given in interpreting study findings on domain importance weighting. Evidence used to argue for and against domain importance weighting must be carefully considered in order to avoid overgeneralizing. Premature conclusions based on overgeneralization of study findings are likely to limit, not help, our understanding of domain importance weighting in QoL measures.

References

Aberson, C. L. (2010). *Applied power analysis for the behavioral sciences*. New York: Psychology Press.

Allison, P. D. (1999). *Multiple regressions: A primer*. Thousand Oaks: Pine Forge.

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

Campbell, A., Converse, P. E., & Rogers, W. L. (1976). *The quality of American life: Perceptions, evaluations, and satisfactions*. New York: Russel Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71–74.

Ferrans, C. E., & Powers, M. J. (1985). Quality of life index: Development and psychometric properties. *Advances in Nursing Science, 8*(1), 15–24.

Hsieh, C. M. (2003). Counting importance: The case of life satisfaction and relative domain importance. *Social Indicators Research, 61*, 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*, 163–174.

Hsieh, C. M. (2012a). Importance is not unimportant: The role of importance weighting in QoL measures. *Social Indicators Research, 109*, 206–278.

Hsieh, C. M. (2012b). Should we give up domain importance weighting in QoL measures? *Social Indicators Research, 108*, 99–109.

Hsieh, C. M. (2013). Issues in evaluating importance weighting in quality of life measures. *Social Indicators Research, 110*, 681–693.

Hsieh, C. M., & Kenagy G. P. Measuring quality of life: A case for re-examining the assessment of domain importance weighting. *Applied Research in Quality of Life*. doi:10.1007/s11482-013-9215-0.

Kaplan, R. M., Feeny, D., & Revicki, D. A. (1993). Methods for assessing relative importance in preference based outcome measures. *Quality of Life Research, 2*, 467–475.

Lix, L. M., Sajobi, T. T., Sawatzky, R., Liu, J., Mayo, N. E., Huang, Y., et al. (2013). Relative importance measures for reprioritization response shift. *Quality of Life Research, 22*, 695–703.

Locke, E. A. (1969). What is job satisfaction? *Organizational Behavior and Human Performance, 4*, 309–336.

Locke, E. A. (1976). 'The nature and causes of job satisfaction' in handbook of industrial and organizational psychology, edited by M. D. Dunnette (Rand McNally, Chicago) pp. 1297–1349.

Locke, E. A. (1984). 'Job satisfaction' in social psychology and organizational behavior, edited by M. Gruneberg and T. Wall (Wiley, London) pp. 93–117.

Luque, M., Miettinen, K., Eskelinen, P., Ruiz, F. (2009). 'Incorporating preference information in interactive reference point methods for multiobjective optmization', OMEGA—International Journal of Management Science 37, pp. 450–462.

Marcenaro, O. D., Luque, M., & Ruiz, F. (2010). An application of multiobjective programming to the study of workers' satisfaction in the Spanish labour market. *European Journal of Operational Research, 203*, 430–443.

Mastekaasa, A. (1984). Multiplicative and additive models of job and life satisfaction. *Social Indicators Research, 14*, 141–163.

Philip, E. J., Merluzzi, T. V., Peterman, A., & Cronk, L. B. (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*, 775–782.

Rojas, M. (2006). Life satisfaction and satisfaction in domains of life: Is it a simple relationship? *Journal of Happiness Studies, 7*, 467–497.

Russell, L. B., & Hubley, A. M. (2005). Importance ratings and weighting: Old concerns and new perspectives. *International Journal of Testing, 5*, 105–130.

Russell, L. B., Hubley, A. M., Palepu, A., & Zumbo, B. D. (2006). 'Does weighting capture what's important? Revisiting subjective importance weighting with a quality of life measure'. *Social Indicators Research, 75*, 146–167.

Trauer, T., & Mackinnon, A. (2001). 'Why are we weighting? The role of importance ratings in quality of life measurement'. *Quality of Life Research, 10*, 579–585.

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*, 101–111.

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

Wu, C. H., Yang, C. T., & Huang, L. N. On the predictive effect of multidimensional importance-weighted quality of life scores on overall subjective well-being. *Social Indicators Research*. doi:10.1007/s11205-013-0242-x.

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

Wu, C. H., & Yao, G. (2006b). Do we need to weight satisfaction scores with importance ratings in measuring quality of life? *Social Indicators Research, 78*, 305–326.

Wu, C. H., & Yao, G. (2007). Importance has been considered in satisfaction evaluation: An experimental examination of Locke's range-of-affect hypothesis. *Social Indicators Research, 81*, 521–541.