# Multiple Group Invariance analysis Across scales with Different Numerical Anchoring

Mehrdad Mazaheri

Department of Psychology, University of Sistan & Baluchestan
Zahedan, Iran
Email: mazaheri@hamoon.usb.ac.ir

*Abstract*—**Aim The current ressearch, includes two studies, was done to : (1) evaluate a hypothesized model on subjective quality of life (SQOL), (2) and to study the role of scale anchoring in satisfaction and dissatisfaction ratings. Method A sample of 456 ( for study 1) and 205 ( for study 2) volunteering students randomly assigned to two different conditions, rated their current overall life (dis)satisfaction and their (dis)satisfaction with six different domains of life. Each condition used one of two rating scale formats, differing in anchoring (-5 to +5 and 0 to 10). Findings The results of the confirmatory factor analysis (CFA) support for the conclusion that, the proposed model of SQOL fit the data well, and is able to predict SQOL. Our results also confirmed that the reproducibility coefficient of the rating scales which have been used for satisfaction and dissatisfaction ratings in this research. Moreover, the results of testing for multiple group invariance of hypothesized model indicated that a cross-validity of proposed model for measuring SQOL.**

*Keywords- Rating scales; Satisfaction; Dissatisfaction, Multiple Invariance Analysis*

## I. INTRODUCTION

The debate among researchers as to the "ideal" rating format has an extensive history. A desired effect of the rating scale method is to provide subjects with a format that allows them to make equal interval judgments thus meeting statistical assumptions of an interval scale of measurement. However, while the rating scale provides a powerful tool for investigating a wide variety of phenomenon, investigations of rating scale function reveal performance anomalies across scale formats.

Rating scales differ in the number of categories as well as number and placement of labels to aid in selection of a category. "Label" is verbal, descriptive statements placed at various locations along vectors of possible response options. Frequently, these options are numbers of increasing and/or decreasing magnitude. The respondent's task is to select the numerical response option associated with the appropriate label that he/she perceive to be the best representation of his/her attitude or belief on a latent trait. There are several characteristics of response formats that are of relevance to the quality of survey data, ranging from the labeling of response categories and the issue of administering scales with or without midpoints, to the question of whether response categories are ordered from positive to negative or the other way around.

Rating scales can be presented as a bipolar or unipolar format. There are two ways in which we may signal to respondents whether we wish them to treat a response scale as unipolar or bipolar. The usual way is by using verbal anchors which are either unipolar (eg *[no more power, much more power]*, *[not having any success, having great success])* or bipolar (eg *[much more power, much less power}*, *[much success, much failure])*. The second way is to use numeric labels which either imply a unidimensional construct (eg *[0 to 10], [0 to 5], [0 to 6}, [-5 to 0})* or bipolar construct (e.g. *[+5 to -5], [+3 to -3}, {+2 to -2})* (1).

While the numeric values are often included only for coding and response convenience, Schwartz and co-workers (2) have demonstrated that they carry more, sometimes unintended, meanings. For a particular question, "How successful have you been in life, so far?", they showed that a scale with numeric values ranging from 0 to 10 was not the same as a scale whose values ranged from -5 to +5. The verbal anchors were "not at all successful" (0 or -5) and "extremely successful" (10 or +5). They argued that when a 0 to 10 scale is used respondents infer that 0 stands for the absence of any amount; the scale becomes *unipolar*. In contrast, respondents infer that the scale is *bipolar* when the numeric values range from -5 to +5. For example, when asking people how successful they had been in their life, if a 0 to 10 scale is offered, they will assume that the low anchor (0) corresponds to *not having any success*. This contrasts with the interpretation of the lowest point on the -5 to +5 scale as being *unsuccessful* (being a failure).

Some investigators suggested that the type of response format and the scales' orientation may affect the respondent's response (3-4-2-5-6-7).

French-Lazovik and Gibson (8) demonstrated that the distribution parameters (means, skewness) of rating scale data were influenced by the labels used. They hypothesized that the negative skew found in most distributions of performance ratings is dependent upon the choice of verbal labels used to anchor rating scale points. Using evaluative rating labels, they were able to systematically change the ratings in the predicted (or hypothesized) direction. The degree of negative skew in distributions of behavioral rating measures was altered by the verbal labels used as anchors. Both means and skewness coefficients were affected by the evaluation labels used in the study. Results presented also showed that a rating scale anchored by a set of more positive evaluative labels (those having higher descriptor indices), results in a shift of the mean numerical values toward the

less positive end of the scale. On the other hand, a rating scale anchored by a set of less positive evaluative labels results in a shift of the mean numerical values toward the more positive end of the scale.

Research questions;

1- To what extent rating the six different domains of life combines to produce an overall measure of subjective quality of life?

2- To what extent does the response format affect the subjective quality of life rating?

## II. STUDY 1

### A. Participants

A sample of 456 volunteering students, , randomly assigned to two different conditions, rated their current overall life (dis)satisfaction and their (dis)satisfaction with six different domains of life (Figure 1). Each condition used one of two rating scale formats, differing in *anchoring* (*-5* to *+5* and *0* to *10*).

### B. Materials and Procedure

Overall life (dis)satisfaction and (dis)satisfaction in different domains of life were measured by a questionnaire contained a total of 14 items. Two items for assessing overall life satisfaction and overall life dissatisfaction which were followed with 12 items for assessing satisfaction (6 items) and dissatisfaction (6 items) in six different domains of life including; *Physical health*, *Psychological well-being*, *Social relations*, *Leisure*, *Financial situation* and *Student life*.

All satisfaction measures are rated on an 11-point visual analogue scale (Figure 2). Each response format presented in one of two set of anchoring (-5 to +5, and 0 to 10).

Dissatisfaction ratings were assessed in the same manner as satisfaction ratings using an 11-point visual analogue scale. For each response format, one of set of anchoring (-5 to +5, and 0 to 10) as an anchor-point were used. (see Figure 3).

### C. SEM Model for Predict Subjective Quality of Life (SQOL)

Structural Equation Modeling (SEM) is a statistical technique for building and testing models, which are often causal in nature. It is a hybrid technique that encompasses aspects of confirmatory factor analysis, path analysis and regression. Indeed all of these can be seen as special cases of SEM. Among its strengths is the ability to model constructs as latent variables which are not measured directly, but are estimated in the model from a number of manifest variables assumed to 'tap into' the construct. This allows the modeler to explicitly capture unreliability of measurement in the model, in theory allowing the structural relations between latent variables to be accurately modeled. In fact, SEM is an extension of the General Linear Model (GLM) that simultaneously estimates relationships between multiple independent and dependent variables, in the case of a structural model and/or multiple observed and latent variables, in the case of confirmatory factor analysis.



By using questions below, we would like to know how **satisfied** and **dissatisfied** you feel about your life and various aspects of your life. Please keep in mind your standards, hopes, pleasures and concerns and think about your life in the last two weeks. Be sure to answer every item.

1- *All things considered, how <u>satisfied</u> are you with your **life as a whole***?

2- *All things considered, how <u>dissatisfied</u> are you with your **life as a whole***?

3- *All things considered, how <u>satisfied</u> are you with your **physical health***?

4- *All things considered, how <u>dissatisfied</u> are you with your **physical health***?

5- *All things considered, how <u>satisfied</u> are you with your **psychological well-being***?

6- *All things considered, how <u>dissatisfied</u> are you with your **psychological well-being***?

7- *All things considered, how <u>satisfied</u> are you with your **social relations***?

8- *All things considered, how <u>dissatisfied</u> are you with your **social relations***?

9- *All things considered, how <u>satisfied</u> are you with your **leisure***?

10- *All things considered, how <u>dissatisfied</u> are you with your **leisure***?

11- *All things considered, how <u>satisfied</u> are you with your **financial situation***?

12- *All things considered, how <u>dissatisfied</u> are you with your **financial situation***?

13- *All things considered, how <u>satisfied</u> are you with your **student life***?

14- *All things considered, how <u>dissatisfied</u> are you with your **student life***?

Figure 1: Satisfaction and Dissatisfaction items



Figure 2: Two different response formats for satisfaction rating scales



Figure 3: Two different response formats for dissatisfaction rating scales

In order to find how six different domains of life combine to produce an overall measure of subjective quality of life, a *SQOL model*, was designed and the strength of this hypothesized model of SQOL was examined using structural equation modeling ( Figure 4).
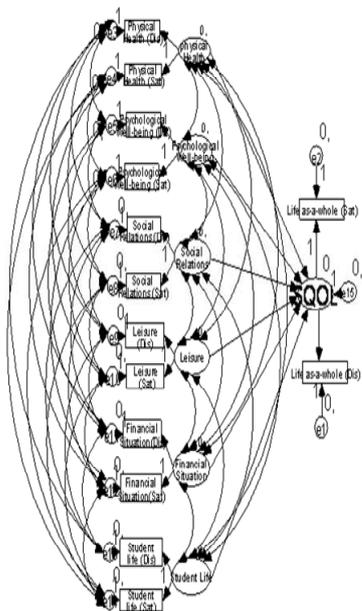
Figure 4: A Model of SQOL

In this model, as seen in Figure 4, ellipses represent latent variables, Physical Health, Psychological Well-Being, Social Relations, Leisure, Financial Situation, and Student Life, and the rectangles represent measured variables, overall life satisfaction , overall life dissatisfaction as well as (dis) satisfaction in six different domains of life, *physical health, psychological well-being, social relations, leisure, financial situation, student life*.

Using AMOS and maximum likelihood estimation, the relationships were examined between the domain of Physical Health, a latent variable with 2 indicators (*physical health satisfaction*, and *physical health dissatisfaction*, the domain of Psychological Well-being, a latent variable with 2 indicators ( *psychological well-being satisfaction* and *psychological well-being dissatisfaction*, the domain of Social Relations, a latent variable with 2 indicators (*social relations satisfaction*, and *social relations dissatisfaction*, the domain of Leisure, a latent variable with 2 indicators (*leisure satisfaction*, and *leisure dissatisfaction*, the domain of Financial Situation, a latent variable with 2 indicators (*financial situation satisfaction*, and *financial situation dissatisfaction*), and the domain of Student Life, a latent variable with 2 indicators (*student life satisfaction*, and *student life dissatisfaction*). The dependent variable was a latent variable of SQOL.

## III. RESULTS (STUDY 1)

The results of the run of the data obtained from two different response formats, Unipolar-Horizontal (0 to 10), and Unipolar-Horizontal (-5 to +5), are presented in Table 1.

Table 1: Fit Indices of the Hypothesized Model of SQOL

| Response Formats | Measures | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\chi^2$ | df | $\chi^2$/df | CFI | NNFI | NFI | RMSEA | PCLOSE | IFI | HOELTER |
| Unipolar-Horizontal (0 to 10) | 123.96 | 27 | 4.59 | .98 | .91 | .97 | .11 | .00 | .98 | 99 |
| Unipolar-Horizontal (-5 to +5) | 48.99 | 27 | 1.81 | .99 | .95 | .97 | .08 | .10 | .99 | 116 |

As seen in Table 1, the $\chi^2$ statistic obtained for the models were 123.96 ($df = 27$, $p < .001$) and 48.99 ($df = 27$, $p < .001$), which seemed to suggest an inadequate fit of the model suggest an inadequate fit of the model. Given that $\chi^2$ statistics is heavily influenced by sample size (Byrne, 2001), other goodness-of-fit indices were examined to evaluate the hypothesized model. As seen in Table 2, the remaining fit indicators for data obtained from both different response formats indicate a reasonable fit between the data and the model.

Comparative fit indexes, as seen in Table 1, were used to compare the fit of one model with the fit of another. The NFI, NNFI, and CFI indexes use an "independence model" (i.e.,"null model") as a basis of comparison by which to assess the hypothesized model. The null model is used as a "baseline" by which to evaluate the hypothesized model (and sometimes additional models, as well). The values for the NFI, NNFI, and CFI are generated by comparing the hypothesized model with the "independence model." These values range from 0.00 to 1.00. Most researchers suggest that comparative fit index values of 0.90 or better indicate a good fit (9). For current analysis, the values for the CFI, the NFI, and the NNFI obtained from two studied different response format were acceptable (higher than .90) and indicate a strong comparative fit for the model.

Table 2 provides estimated Regression Weights (**B**) and Standardized Regression Weight (**β**) coefficients of all pathways for the hypothesized model of SQOL.

Table 2: Analysis of a Model of SQOL

| Pathways | Unipolar- Horizontal (0 to 10) | | | Unipolar-Horizontal (-5 to +5) | | |
|---|---|---|---|---|---|---|
| | B | β | p | B | β | p |
| Physical Health→SQOL | .22 | .22 | .001 | .14 | .18 | .009 |
| Psychological Well-being→ SQOL | .42 | .46 | .001 | .36 | .46 | .001 |
| Social Relations→ SQOL | .11 | .12 | .018 | .25 | .31 | .001 |
| Leisure→ SQOL | .10 | .11 | .008 | .08 | .12 | .097 |
| Financial Situation→ SQOL | .10 | .12 | .004 | .08 | .12 | .066 |
| Student Life→ SQOL | .15 | .16 | .002 | .01 | .01 | .859 |
| Physical Health→ *Physical Health*(Sat.) | 1 | .86 | .001 | 1.00 | .99 | .001 |
| Physical Health→ *Physical Health*(Dis.) | -.94 | -.84 | .001 | -.98 | -.87 | .001 |
| Psychological Well-being→ *Psychological well-being*(Sat.) | 1 | .98 | .001 | 1.00 | .96 | .001 |
| Psychological Well-being→ *Psychological well-being*(Dis.) | -.97 | -.91 | .001 | -.99 | -.94 | .001 |
| Social Relations→ *Social relations* (Sat.) | 1 | 1.04 | .001 | 1.00 | .98 | .001 |
| Social Relations→ *Social relations* (Dis.) | -.89 | -.88 | .001 | -.94 | -.86 | .001 |
| Leisure→ *Leisure* (Sat.) | 1 | .99 | .001 | 1.00 | 1.03 | .001 |
| Leisure→ *Leisure* (Dis.) | -.96 | -.93 | .001 | -.84 | -.80 | .001 |
| Financial Situation→ *Financial situation* (Sat.) | 1 | .97 | .001 | 1.00 | .98 | .001 |
| Financial Situation→ *Financial situation* (Dis.) | -1.03 | -.95 | .001 | -.95 | -.95 | .001 |
| Student Life→ *Student life* (Sat.) | 1 | 1 | .001 | 1.00 | .98 | .001 |
| Student Life→ *Student life* (Dis.) | -.97 | -.93 | .001 | -1.01 | -.95 | .001 |
| SQOL→ Overall life Satisfaction | 1 | .98 | .001 | 1.00 | .94 | .001 |
| SQOL→ Overall life Dissatisfaction | -.99 | -.93 | .001 | -1.07 | -.90 | .001 |

As seen in Table 2, for scores derived from Unipolar-Horizontal (0 to 10) response format, all of the paths in the model are significant and the model of SQOL presented illustrates the hypotheses that six different domains of life,

Physical Health, Psychological Well-Being, Social Relations, Leisure, Financial Situation, and Student Life, predict overall life satisfaction and these latent variables are all interrelated. But For scores derived from Unipolar-Horizontal (-5 to +5) pathways between Subjective quality of life and following latent variables were found non significant; Leisure, Financial situation and student life. When using Unipolar-Horizontal (Not Numbered), pathways between subjective quality of life and following latent variables were shown non significant; Social Relations, Leisure, and Financial Situation.

Moreover as seen in Table 2, subjective quality of life as a latent variable was found to predict well overall life satisfaction and overall life dissatisfaction rating. Among different domains of life, Psychological Well-being was found as the strongest significant predictor of SQOL on both two different response formats.

## IV. STUDY 2

Study 2 was done to find whether or not the results of the satisfaction and dissatisfaction ratings reported in study 1, can be reproduce using the same scale with a new sample. The results of current analysis also can be considered as the evidence for the scale's *Reproducibility* in sense of the precision of the measurement under different conditions.

Reproducibility was defined as the proportion of total variance (*ie*, between-subject plus within-subject variance) explained by the between-subject variance given as a percentage. In general, the scale's *reproducibility* is the variation in a series of measurements that have been taken by different people using the same instrument to measure one characteristic or the identical characteristic on the same part. Participants:

A sample of 205, sample [1] volunteering students was, randomly assigned to two different conditions, rated their current overall life (dis)satisfaction and their (dis)satisfaction with six different domains of life (*Physical health, Psychological well-being, Social relations, Leisure, Financial situation* and *Student life*). Each subject was asked to complete a questionnaire consisting of fourteen items, and was randomly assigned to one of twelve versions of questionnaire.

Using the "Unconstrained" and "Measurement weights" model, the invariant of the components of the measurement model, SQOL model, were examined across particular groups (Samples 1 and 2). The analysis was done, separately for two pairs of parallel response formats obtained from sample 1 and sample 2 and the results are presented in Table 3.

---

[1] For current analysis, we refer to this group of participants as Sample 2 versus Sample 1 which has been participated in study 1.

Table 3: Goodness-of-Fit Statistics for Tests of Invariance across sample 1 and sample 2

| Response format | Model Description | $\chi^2$ | $df$ | $\Delta\chi^2$ | $\Delta df$ | CFI | RMSEA | $p <$ |
|---|---|---|---|---|---|---|---|---|
| Unipolar-Horizontal (0 to 10) | Unconstrained | 179.96 | 54 | - | - | .98 | .08 | - |
| | Regression weights | 190.94 | 67 | 10.98 | 13 | .98 | .07 | ns |
| Unipolar-Horizontal (-5 to +5) | Unconstrained | 110.95 | 54 | - | - | .98 | .07 | - |
| | Regression weights | 116.35 | 67 | 5.40 | 13 | .98 | .06 | ns |

As seen in Table 3, the comparison of models, Unconstrained and Measurement weights, was found no significant across sample 1 and sample 2 and this was shown for all two different response formats.

## V. DISCUSSION AND CONCLUSION

The hypothesized model of SQOL was designed to test the relationships between six domains of Physical Health, Psychological Well-being, Social Relations, Leisure, Financial Situation, and Student Life, as the latent variables with 2 indicators (satisfaction and dissatisfaction in concern domains) per each domain to predict subjective quality of life as a latent variable. Our results indicated that in spite of a significant $\chi^2$ statistic for the model, other indicators, such as CFI, *NNFI*, NFI, and *IFI* were shown all higher than .90 and proposed model to fit the data well (see Table 1).
Our results indicated that, comparing two different response formats, only for scores derived from Unipolar-Horizontal (0 to 10) response format, all of the paths in the model were found significant (Table 2). This finding suggest that an adequate predictive validity for all six different domains of life to predict subjective quality of life.

Our results of testing for multiple group invariance of model, SQOL, for each response formats, across sample 1 and sample 2 was shown non-significant across sample 1 and sample 2 (see Table 3 for different models). These results can be considered as an evidence for cross-validity of the subjective quality of life rating (SQOL model). Finally, the results of current study confirmed that the reproducibility coefficient of the rating scales which have been used for (dis)satisfaction ratings in this research.

## REFERENCES

[1] O'Muircheartaigh CA, Gaskell GD, Wright DB. Evaluating Numeric and Verbal Labels for Response Scales. Paper Presented at the 48th Annual Conference of the American Association for Public Opinion Research, May 20-23, St.Charles, Illinois, 1993.

[2] Schwartz, N., Knauper, B., Hipler, H.J., Noelle-Neumann, E., & Clark, L. (1991). Numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, *55*, 570-582.

[3] Gift, A.G. (1989) Validation of a vertical visual analogue scale as a measure of clinical dyspnea. *Rehabil Nurs*, *14*, 323-325.

[4] Friedman, L.W., & Friedman, H.H. (1994). A comparison of vertical and horizontal rating scales. *Mid-Atlantic Journal of Business*, *30*, 107-111.

[5] Mazaheri, M., & Theuns, P. (2006). A Comparison of Different Formats of the Anamnestic Comparative Self Assessment (ACSA) for the Assessment of Subjective Well-being. Pro Newletter, 36 (Spring Issue).

[6] Mazaheri, M., & Theuns, P (2009a). Effects of Varying Response Formats on Self-ratings of Life-Satisfaction. *Social Indicators Research, 90 (3)*, 381-395.

[7] Mazaheri, M., & Theuns, P (2009b). Structural Equation Modeling (SEM) for Satisfaction and Dissatisfaction Ratings; Multiple Group

Invariance Analysis Across Scales with Different Response Format. *Social Indicators Research, 91 (2)*, 203-221.

[8] French-Lazovik, U. & Gibson, C.L. (1984). Effects of verbally labeled anchor points on the distributional parameters of rating measures. *Applied Psychological Measurement, 8(1),* 49-57.

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