Nearest-Integer Response from Normally-Distributed Opinion Model for Likert Scale

Jonny B. Pornel, Vicente T. Balinas
and Giabelle A. Saldaña
University of the Philippines Visayas

This paper proposes that respondents' opinions on Likert Scale items are normally distributed around their latent ability although their observable responses will be integers in the scale nearest to those opinions. We tested the appropriateness of the model on actual data gathered by a Likert scale developed to measure attitude of teachers towards research undertaking. The soundness of common research practice of using mean and standard deviation to estimate the respondents' latent ability was tested. The results show that the NIRNDO model could be used appropriately to model responses on Likert scale. Also, the results show that using the mean response to a Likert scale, the resulting 95% confidence interval (mean $\pm$ 1.96 SEM) would be effective at least 90% of the time. This effectiveness is guaranteed for latent ability in the optimum range $[u + 0.8, v - 0.8]$ where u and v are the lowest and highest points in the scale, respectively.

Keywords: Likert Scale, NIRNDO Model, latent ability

1. Introduction

Likert scale (Likert, 1932) has become an important instrument in the fields of social sciences (Wu, 2007), education (Gay & Airasian, 2000), medicine, marketing research (Albaum, 1997) and others for measuring latent abilities such as attitude. Uebersax (2006) characterizes the Likert scale with the following main features: (1) it contains several items, (2) response levels are arranged horizontally, (3) response levels are anchored with consecutive integers, and (4) response levels are also anchored with verbal labels that are more-or-less evenly-spaced. Likert scale was intended to be a summated scale (Likert, 1932). A good Likert scale should have high internal consistency. The several items contained in a Likert scale should be designed to solicit opinions (agreement or disagreement, approval or disapproval, etc.) of the respondents on different situations related to the phenomenon being investigated. A summary measure of these opinions is intended to estimate the latent ability (e.g., attitude towards homeschooling) of the respondent.

Likert scale is a kind of polytomous test, that is, the scoring scheme is unlike the binary test where examinees have only two possible scores for each item, 1 or 0 (right or wrong). In polytomous test, the scoring scheme for each item have several score levels, such as (0, 1, 2 and 3), (1, 2, 3, 4, and 5), etc.

The popularity of Likert scale is mostly due to its ease of construction, administration, response (Albaum, 1997) and interpretation. These characteristics were important factors in the spread of its use in various fields whose practitioners cannot be rightfully assumed to be well-versed in psychometrics, and thus would gladly use tools that can be easily learned, applied and interpreted.

Other researchers are wary, however, of the common practice of using mean, standard deviation, and parametric tests on data gathered using Likert scale. Parametric statistical tools require at least interval data. The critics contend that Likert scale data are (a) of ordinal type and (b) coarse.

On the issue of the type of data from Likert scales, they contended that in Likert scale with choices such as (1) Strongly Disagree, (2) Disagree, (3) Agree, and (4) Strongly Agree, the choices can be ordered but the difference between 1 and 2 cannot be certainly claimed to be the same as the difference between 2 and 3, leading them to say that data gathered using Likert Scale are not of interval type. Thus, using parametric tools, for example, on this data are violations of sound statistical principles; they insisted that only nonparametric test should be used (Jamieson, 2004). Kuzon, Urbanchek & McCabe (1996) contended that the use of parametric analysis on ordinal data is one of seven deadly sins of statistical analysis.

On the issue of coarseness, there is a consensus that the latent variables, for which Likert Scales were devised to measure, are of an essence continuous. However, the scale forces the respondents to choose only from few choices or responses, and this results to different true scores being lumped together into the same category (Aguinis, Pierce and Culpepper, 2008). For Example, in a scale designed to measure attitude towards a topic and with possible responses of 1, 2, 3, 4, and 5, two respondents with opinions 2.7 and 3.2 (almost neutral attitude, though of varying degree and inclination) on a certain item of the scale, would most probably select response 3 as their response and will be lumped together as having the same opinion. In so doing, there is loss of information due to coarseness. This is not a trivial issue.

Russell and Bobko (1992), in their study on the effect of the response scale on the power of moderated regression, claimed that the loss of information due to Likert scale “greatly reduces the probability of detecting true interaction effects.”

Sung and Kang (2006) discussed several Item Response Theory (IRT) models for polytomous test, like Likert Scale, three of which were the rating scale model (Andrich, 1978), the partial credit model (Masters, 1982), and the generalized partial credit model (Muraki, 1992).

The rating scale model (RSM) proposes that the probability that examinee j chooses response r in a Likert scale is

$$P\left(R_{ji} = r \big/ \theta_j, \beta_i\right) = e^{\sum_{c=u}^{r} \frac{(\theta_j - \beta_i - \tau_c)}{\sum_{y=u}^{v} e^{\sum_{c=u}^{y} (\theta_j - \beta_i - \tau_c)}}}$$

where θ_j is the latent ability of respondent j, β_i is the difficulty of item i, and τ_c is the location parameter for choice *c*. Further, *u* and *v* are the least and highest value of *c* respectively.

The partial credit model (PCM) is similar to RSM with the added parameter τ_{ci} instead of being constant for each choice across items, that is the probability that examinee j chooses response r in a Likert scale is

$$P\left(R_{ji} = r \big/ \theta_j, \beta_i, \tau_{ci}\right) = e^{\sum_{c=u}^{r} \frac{(\theta_j - \beta_i - \tau_{ci})}{\sum_{y=u}^{v} e^{\sum_{c=u}^{y} (\theta_j - \beta_i - \tau_{ci})}}}$$

The graded partial credit model (GPCM) is an elaboration of PCM by adding the discrimination index α_i. That is, the probability that examinee j chooses response r in a Likert scale is

$$P\left(R_{ji} = r \big/ \theta_j, \alpha_i, \beta_i, \tau_{ci}\right) = e^{\sum_{c=u}^{r} \frac{\alpha_i (\theta_j - \beta_i - \tau_{ci})}{\sum_{y=u}^{v} e^{\sum_{c=u}^{y} \alpha_i (\theta_j - \beta_i - \tau_{ci})}}}$$

However, Item Response models such as RSM, PCM and GPCM have stringent requirements on the items of the scales being considered. Using the models when the items do not satisfy those requirements could lead to some drastic outcome. These stringent requirements, as well as the mathematical rigor involved in IRT models made those models unpopular with researchers from fields not associated with good mathematical inclination.

In contrast to IRT models, the model being proposed in this paper does not take into account the parameters of each item in the scale. It assumed that the opinion ϑ of respondents on each item of a Likert Scale (that is, a Likert scale one that was designed to measure the respondent's latent variable θ such as attitude, self-concept, motivation, etc.) is normally distributed with mean and standard deviation . Using this assumption, the paper proceeded to determine whether the mean response $\bar{r}_j = \frac{\sum_{i=1}^{n} r_{ij}}{n}$ of respondent *j* on a Likert Scale with *n* items can be a good estimate of the respondent's latent ability being considered despite of the coarseness brought about by the response format inherent to the scale.

2. The Model

This paper proposes a new mathematical model for the responses to items of a Likert Scale. Such model was then used to explore the effect of coarseness on the soundness of common practice of using mean response in the data analysis.

The idea of this paper is modeled from the tendency of respondents, when confronted with several items of a Likert scale, to signify answers of varying degree from item to item. For example, a certain respondent who was asked to answer a Likert scale with 10 items, each item of which has five choices (1 – Strongly disagree, 2 – Disagree, 3 – Undecided, 4 – Agree, 5 – Strongly Agree), may have a response vector R= (3, 4, 3, 4, 2, 3, 3, 4, 3, 3). This score vector does not mean that the latent ability of the respondent varies. Instead, it signifies that the opinions of the respondent vary from item to item but these opinions hover around the latent ability. This paper assumes that, on any item of the scale, the distance of the respondent's opinion from the latent ability is normally distributed with mean θ and standard deviation σ.

Say, a respondent of a 4-point, 30-item Likert Scale on attitude towards compulsory sex education to grade school students have a slightly positive attitude towards the issue. Assuming that on the scale of 1 to 4 with 4 as the highest, his attitude is 2.75. Then, his opinion ϑ on any item of the scale, assuming it was excellently constructed, will be around 2.75 more or less. This paper proposed that $\vartheta \sim N(2.75, \sigma)$. However, since Likert Scale forces the respondent to choose only integral answers, then the respondent's answer to a certain item would be the integer nearest to his or her opinion for the item.

Formally, let *L(k,n)* denote a k-point n-item Likert scale. That is, the scale has *n* items and each item has *k* choices. It is important that these items should be unidimensional, or at least have high internal consistency, and focus on measuring opinions for a single phenomenon. There are many standard procedures to do this.

For any item, let *u* be the choice with the lowest value and *v* be the one with the highest value. It must be noted that $v = u + k - 1$. Let $I_1, I_2, \ldots, I_n$ be the items of *L(k,n)*. Also, let *L(k,n)* be a measure of a certain phenomenon (e.g. attitude, awareness, acceptability, etc.).

A certain respondent *j* of the Likert scale *L(k,n)* will have ability (i.e. level of attitude, awareness, acceptability, etc.) θ_j with regard to the phenomenon, where $u \leq \theta_j \leq v$. When confronted with item I_i, respondent *j* will then form opinion ϑ_{ji}. The model being advance here, the Nearest-Integer Response from Normally-Distributed Opinion (NIRNDO) model, contends that $\vartheta_{ji} \sim N(\theta_j, \sigma_j)$. It has to be pointed out that this θ_j may change due to some circumstances but is assumed to be stable for some length of time which makes it worth measuring. In this model, it is the opinion ϑ_{ji} that is assumed to vary with σ_j around θ_j.

Since a Likert scale records only integral response, then respondent *j* will need to convert opinion ϑ_{ji} to response R_{ji} where R_{ji} is an integer between *u* and *v* inclusive. For example, if *u* and *v* are 1 and 5 respectively, then R_{ji} can only be 1, 2, 3, 4, or 5. This model further claims that

$$R_{ji} = r \quad \text{where } |r - \vartheta_{ji}| = \min(|x - \vartheta_{ji}|) \quad (1)$$

for all integer x from *u* to *v* inclusive

Necessarily, $u \leq r \leq v$. For example, Consider $L(4, n)$ where $u = 1$ and $v = 4$. Then

$$\overline{R}'_{ji} = \begin{cases} 4 & \vartheta_{ji} \geq 3.5 \\ 3 & 2.5 \leq \vartheta_{ji} < 3.5 \\ 2 & 1.5 \leq \vartheta_{ji} < 2.5 \\ 1 & \vartheta_{ji} < 1.5 \end{cases} \quad (2)$$

Since the model contends that $\vartheta_{ji} \sim N(\theta_j, \sigma_j)$, it follows that to solve for probability that the observed response on item I_i is 1 given that the latent ability of respondent *j* is θ_j, the following equation will be used

$$P(R_{ji} = 1/\theta_j) = \int_{\infty}^{1.5} N(\theta_j, \sigma_j) \quad (3)$$

In general, $P(R_{ji} = r/\theta_j)$ of respondent *j* on item *i* the Likert Scale *L(k, n)* is

$$P(R_{ji} = r/\theta_j) = \int_a^b N(\theta_j, \sigma_j) \quad (4)$$

Where *a* and *b* are the lower and upper boundaries of r.

It must be pointed out that a and b are functions of r. In $L(4,n)$, the ordered pair (a,b) refers to $(-\infty, 1.5)$,(1.5, 2.5), (2.5, 3.5) and $(3.5, \infty)$ when $r = 1, 2, 3$, and 4 respectively.

For notational convenience, let

$$\int_a^b N(\theta_j, \sigma_j) = \Phi_{a,b} \tag{5}$$

Then,

$$P\,(r_{ji} = r/\theta_j) = \Phi_{a,b} \tag{6}$$

From this, one may solve for the expected response $E(R_j)$ of respondent j,

$$E(R_j) = \sum_{r=u}^{v} r\Phi_{a,b} \tag{7}$$

These results will be used in determining how good is the mean response on the items of Likert scale as an estimate of the latent variable being measured by the scale. The study will use simulated data using the model as well as actual data on teachers' attitude towards research undertakings (Pornel et al., 2010) measured using a Likert scale.

This paper aimed to introduce NIRNDO model and use it to explore the soundness of the popular practice of solving for the mean response for a Likert scale to determine the respondents' latent ability. Specifically, this study aimed to do the following:

1. Given actual response vector $R_j = (r_{j1}, r_{j2},..., r_{ji}, r_{jn})$ determine whether $\bar{r}_{ji}$ significantly differs from $\bar{r}'_{ji}$ where $\bar{r}'_{ji}$ is the response vector generated using $N(\hat{\theta}_j, \hat{\sigma}_j)$ and $\hat{\theta}_j = \bar{r}_{ji}$ and $\hat{\sigma}_j = S_{rji}$.
2. Given a respondent with ability level θ_j with σ_j, determine whether $\bar{r}'_{ji} \pm 1.96$SEM is a good estimator of θ_j, where $\bar{r}'_{ji}$ is the average of the generated responses r'_{ji}.
3. Determine the RMSE of $E(R)$ across different values of σ.

Simply speaking, the first objective is to verify whether the model could generate a response vector closely related to an observed response vector using the mean and standard deviation of the observed responses. This is equivalent to a test of normality of $\bar{r}_{ji}$. In some way, it will explore whether there is basis to the practice of using parametric test on data gathered using Likert scale. The second objective is to establish whether, in the context of the model, given a theoretical latent ability θ_j with σ_j, the mean response in a Likert scale is a good estimate of θ_j. The third objective is to explore how the error caused by the coarseness of the Likert scale vary with the σ_j.

3. The Simulations

In this paper, simulations one, two and three were conducted to achieve objectives one, two and three respectively. The results of simulation 1 determined whether NIRNDO model works well with actual data. On the other hand, simulation 2 explored whether the mean response in a Likert scale is useful in estimating the latent ability of the respondent. Lastly, the third simulation studied the parameter estimation error associated with Likert scale as affected by the variance of the respondent's opinions. The algorithms for these simulations were as follows:

Simulation 1

1. Given actual response vectors $R_j = (r_{j1}, r_{j2},..., r_{jn})$ of respondents in Pornel et al. (2010), solve for $\hat{\theta}_j = \bar{r}_{ji}$ and $\hat{\sigma}_j = S_{rji}$.
2. Generate the opinion vector by generating ϑ'_{ji} from $N(\hat{\theta}_j, \hat{\sigma}_j)$ for $n = 30$ times where n is the number of items of the Likert scale.
3. Determine the generated response vector $R'_j = (r'_{j1}, r'_{j2},..., r_{jn})$ using equation 1.
4. Do steps 1 to 3 for $m = 95$ times where m is the number of respondents.
5. Determine whether $\bar{r}'_{ji}$ significantly differs from $\bar{r}_{ji}$
6. Do steps 1 to 5 for $t = 100$, where t is the number of trials.

Simulation 2

1. Let $k = 4$
   a. Let $n = 30$
   b. Let $u = 1$ and $v = k$
   c. Let $\theta_j = u$
   d. Let $\sigma_j = 0.1$
   e. Generate the opinion vector by generating ϑ_{ji} from N (θ_j, σ_j) for n times.
   f. Determine the response vector $R'_j = (r'_{j1}, r'_{j2},..., r_{jn})$.
   g. Determine $\bar{r}'_{ji}$ and $SEM = Sr'_{ji} / \sqrt{n}$
   h. Determine the interval $[x,y]$ where $x = \bar{r}'_{ji} - 1.96SEM$ and $y = \bar{r}'_{ji} + 1.96SEM$
   i. Determine whether $\theta_j \in [x,y]$.

j. Do steps e to i for *m=10,000* times.
k. Determine the proportion β_k wherein $\theta_j \in [x,y]$
l. Do steps d to k using σ_j with increasing value (increment is 0.1) until $\sigma_j = 1.0$
m. Do steps c to l using θ_j with increasing value (increment is 0.1) until *v*.

2. Do step 1 for $k = 5$ and 6.
3. Do steps 1 and 2 for $n = 25$ and 20.

<u>Simulation 3</u>

1. Let $k = 4$
2. Let $u = 1$ and $v = k$
3. Let σ = .01

a. Determine $\theta - \hat{\theta}$ using equation 7 for $\theta = u+.1, u+.2, ..., v-.1$. Where $\hat{\theta} = E(R)$

b. Solve for $[\theta - \hat{\theta}]^2$ for $\theta = u +.1, u +.2,...,v - .1$

c. Solve for the root mean square of error $\sqrt{\dfrac{\sum(\theta - \hat{\theta})^2}{(k-u)/.1}}$ for $\theta = u +.1, u +.2, ..., v -.1$

4. Do step 3 for $\sigma = 0.01, 0.02, ..., 1.50$
5. Do steps 1 to 4 for $k = 5$ and 6.
6. Determine Optimum range of σ that result to minimum RMSE.

4. Results and Discussion

First Simulation: Deviation of Generated Mean Response Using NIRNDO Model from the Observed Mean Response to a Likert Scale

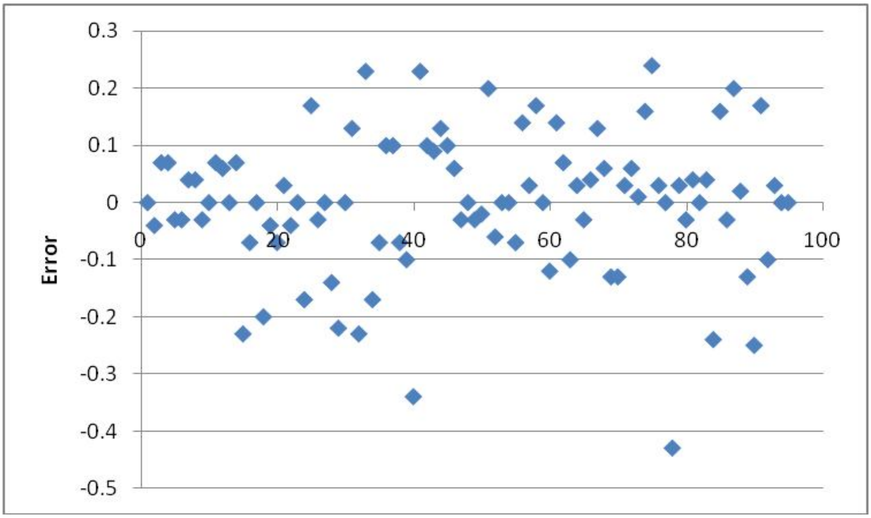
In the first simulation, the answers of 95 respondents to the Likert scale used by Pornel et al. (2010) to measure the teachers' attitude towards research undertakings were considered. The scale has originally 36 items. However, after item analysis and revision, it was reduced to 30 items. The revised instrument has a Cronbach Alpha reliability of 0.910. The least observed mean response of the respondents is $\bar{r}_{ji} = 2.17$, while the highest observed value is $\bar{r}_{ji} = 3.97$. For the standard deviation of the responses s_{rji}, the least observed value is 0.18, while the highest observed value is 1.49 as shown in Table 1.

Table 1 Summary Measures of Responses to the Likert Scale used in Pornel et al. (2010)

	Teachers' Response to Pornel et al. (2010)	
	Mean Response	SD
Maximum	3.97	1.49
Minimum	2.17	0.18
Mean	2.91	0.62

Using the mean, $\hat{\theta}_j = \bar{r}_{ji}$, and standard deviation of the responses, $\hat{\sigma}_j = s_{rji}$, the researchers generated opinions ϑ'_{ji} from $N(\hat{\theta}_j, \hat{\sigma}_j)$ for each respondent as predicted by NIRNDO model, then simulated the most probable responses, r'_{ji}, based on equation 1. Computing for the absolute error, $|\bar{r}'_{ji} - \bar{r}_{ji}|$, the researchers found that the maximum absolute error in one run of the simulation to be 0.43, the minimum is 0.00 and the average of the absolute errors is 0.09. When the errors were plotted they were found to hover randomly around zero as shown in Figure 1.

Figure 1 Scatterplot Showing Errors of Estimate for each Respondent in One Trial



Using the Wilcoxon Signed Rank test to determine the significance of the difference between the mean simulated response $\bar{r}'_{ji}$ and the mean observed response $\bar{r}_{ji}$ showed that there was no significant difference between the observed and projected mean 92% of the time. This test was performed 100 times.

Second Simulation: Efficiency of Mean Response on a Likert Scale to Estimate the Respondent's Latent Ability

In the second simulation, the researchers generated the responses of respondents given their latent ability θ_j (with the corresponding standard deviation σ_j) and then determined the average response $\bar{r}_{ji}$ and the standard error of the mean $SEM = s_{rji}/\sqrt{n}$. The researchers, then, determined how often the interval $\bar{r}_{ji} \pm 1.96SEM$ contains.

The result of the simulation shows that in $L(4,30)$**,** when σ is equal or less than 0.3, the interval $\bar{r}_{ji} \pm 1.96SEM$ is inconsistent in containing θ_j across different values of σ. That is, for some value of θ_j, the interval $\bar{r}_{ji} \pm 1.96SEM$ contains θ_j at least 90% of the time while in some values of $\theta_{j,}$ the interval contains θ_j, less than 90% of the time. However, when σ is greater than 0.3, the interval $\bar{r}_{ji} \pm 1.96SEM$ contained θ_j more than 90% of the time when the value of the ability θ_j is from 1.8 to 3.2 as shown in Table 2. The same trend was observed for $L(5,30)$ and $L(6,30)$ as indicated in Tables 3 and 4. That is, $\bar{r}_{ji} \pm 1.96SEM$ contains θ_j at least 90% of the time when θ_j is at least a distance of 0.8 from the edges and σ is greater than 0.3.

Similar patterns were found when n is either 25 or 20. The results for these simulations are found in Tables 7 to 12 of the Appendix.

Since, the interval $\bar{x} \pm 1.96SEM$ is supposed to be a 95% confidence interval for normally distributed variable, these results showed that the coarseness due to Likert scale lessen the effectiveness of the confidence interval. When θ_j is in the optimum range (0.8 from the edges of the scale), the confidence interval $\bar{r}_{ji} \pm 1.96SEM$ can be effective at least 90% of the time.

Checking the distribution of data in Pornel et al. (2010), it was found that 80% of the respondents to the Likert scale used in the study had mean responses between 1.8 and 3.2 inclusive as shown in Table 5. Also, 90% of the respondents had σ greater than 0.3 but not more than 1.0.

Table 2 Effectiveness of $\bar{r}_{ji} \pm 1.96SEM$ in containing θ_j in $L(4,30)$

θ	σ 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1.0	100.00	100.00	94.54	60.23	27.94	11.76	5.17	2.40	1.28	0.69
1.1	0.00	14.89	76.45	87.23	69.76	49.16	35.37	24.99	18.09	13.36
1.2	0.00	32.72	87.42	**95.72**	**90.02**	80.60	68.17	54.82	44.25	36.12
1.3	0.06	52.06	90.49	95.57	93.36	87.73	80.10	71.18	63.02	54.59
1.4	9.35	74.98	91.08	93.97	93.70	**91.50**	87.51	82.58	76.46	69.73
1.5	95.80	95.94	96.01	95.75	94.27	92.19	**90.31**	87.29	83.55	79.33
1.6	8.89	75.09	90.38	93.76	93.85	93.51	91.55	**90.62**	88.07	84.51
1.7	0.03	52.64	89.45	93.50	94.02	93.29	93.51	91.78	**90.63**	87.81
1.8	0.00	32.46	83.41	93.15	94.66	94.23	93.53	93.01	91.85	**90.48**
1.9	0.00	14.93	77.06	92.29	93.28	93.70	94.53	93.56	93.12	92.14
2.0	100.00	99.99	97.05	93.01	92.79	92.84	93.25	93.12	92.87	92.98
2.1	0.00	14.65	76.93	92.13	93.32	93.62	94.28	94.03	94.01	93.53
2.2	0.00	31.89	83.45	92.91	94.17	93.89	93.52	93.56	93.78	93.97
2.3	0.08	53.11	89.13	93.79	93.75	93.54	93.49	93.56	94.07	94.02
2.4	9.09	75.20	90.39	93.59	94.07	93.79	94.09	93.90	94.49	93.87
2.5	95.94	96.24	95.79	94.80	94.30	93.67	93.52	94.04	93.85	94.32
2.6	8.84	75.57	90.51	93.77	94.67	93.88	94.55	93.88	94.06	93.72
2.7	0.04	52.50	89.12	93.55	94.13	93.41	93.90	93.94	93.83	93.95
2.8	0.00	32.22	83.02	92.92	94.22	94.26	94.13	93.55	93.80	93.43
2.9	0.00	14.59	77.04	92.73	93.24	93.77	93.80	94.19	94.32	93.77
3.0	100.00	100.00	100.00	93.44	92.49	93.26	93.62	93.41	93.25	92.67
3.1	0.00	14.15	77.29	92.17	93.38	93.61	94.17	93.62	93.52	92.75
3.2	0.00	32.58	82.57	92.91	94.21	94.70	94.06	93.80	**92.33**	**91.23**
3.3	0.09	52.53	88.56	93.59	93.45	93.56	92.92	92.22	89.97	88.21
3.4	8.99	74.92	90.91	93.83	93.94	93.04	91.83	**90.41**	88.08	85.34
3.5	95.65	95.90	95.73	95.12	94.02	92.34	**90.25**	87.54	83.88	79.27
3.6	9.30	75.03	90.71	93.34	93.54	**91.61**	87.64	82.74	77.02	69.83
3.7	0.06	53.26	90.24	95.39	**92.98**	87.53	81.26	72.03	62.64	54.63
3.8	0.00	32.42	87.03	**95.26**	89.96	79.31	68.09	56.20	44.70	35.94
3.9	0.00	15.25	76.42	87.10	69.71	50.09	34.95	25.66	18.46	13.77
4.0	100.00	100.00	94.50	61.14	27.77	12.02	5.08	2.49	1.63	0.75

A low σ may signal that the respondent exhibited the blind effect, a response bias wherein the respondent would choose the answers without reading each item. This happens when a disinterested respondent would decide to have a fix answer, say 3, in $L(4,30)$ scale, and decide to answer all or most of the items with the predetermined choice without reading any of them. With response vectors from this kind of respondents, the model will fail. However, it is also possible that the respondent has a very low *s*, especially when his latent ability θ is so near an integral choice in the scale (say θ = 2.9, this is so near the integral choice 3 of a 4-point Likert scale). In this situation, the respondent will signify answers that are mostly, if not all, 3. With this kind of respondent, the mean response would be a good estimate of the respondent's latent ability.

Table 3 Effectiveness of $\bar{r}_{ji} \pm 1.96SEM$ in containing θ_j in $L(5,30)$

θ	σ 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1.0	100.00	100.00	95.19	61.05	27.35	11.24	5.23	2.71	1.27	0.68
1.1	0.00	14.62	76.56	87.16	69.41	50.22	35.63	25.42	18.30	14.46
1.2	0.00	32.58	86.91	**95.42**	**90.01**	80.74	68.52	55.16	44.73	35.38
1.3	0.03	52.69	90.27	95.63	93.51	88.24	80.37	72.08	62.42	55.18
1.4	8.78	73.86	90.96	93.95	93.94	**91.68**	87.94	82.45	75.84	69.52
1.5	96.54	95.32	95.44	95.82	94.52	92.83	**90.58**	87.08	84.00	79.10
1.6	9.23	76.03	90.51	93.65	93.81	93.12	91.95	**90.64**	87.87	84.94
1.7	0.08	53.07	89.26	93.81	93.76	93.74	93.18	91.97	**90.93**	88.12
1.8	0.00	31.83	83.94	93.06	94.33	94.35	93.50	92.99	92.18	**90.15**
1.9	0.00	14.88	77.07	92.24	94.06	93.47	94.22	94.01	93.34	91.61
2.0	100.00	100.00	96.84	92.62	93.20	93.37	93.45	93.25	93.05	92.41
2.1	0.00	14.61	77.31	92.33	93.70	94.17	93.94	94.76	93.93	92.98
2.2	0.00	31.98	83.22	93.13	94.34	94.08	93.52	94.00	94.21	93.87
2.3	0.07	52.67	89.44	93.32	94.19	93.74	93.70	93.60	93.58	94.01
2.4	8.90	75.17	90.46	93.63	94.03	93.84	94.13	93.86	94.42	93.80
2.5	95.85	95.85	95.54	95.31	95.75	94.14	93.53	94.17	93.85	94.17
2.6	9.52	74.88	90.93	93.64	94.12	93.74	93.92	94.14	94.17	94.19
2.7	0.07	53.10	88.44	93.93	93.88	93.49	94.31	93.95	94.08	93.76
2.8	0.00	33.14	83.01	93.10	94.71	93.91	94.03	93.76	94.36	93.58
2.9	0.00	14.82	77.37	92.49	94.05	94.20	94.15	94.52	93.97	94.06
3.0	100.00	100.00	96.40	92.83	93.03	93.64	93.81	93.44	93.50	94.26
3.1	0.00	15.34	76.63	92.76	93.33	94.16	94.35	94.55	93.84	94.46
3.2	0.00	32.24	83.05	92.81	94.14	94.57	93.68	94.09	93.59	93.92
3.3	0.05	52.09	89.22	93.74	93.90	94.12	94.15	93.72	94.21	93.72
3.4	8.80	74.57	91.00	93.98	94.30	93.68	97.75	93.87	93.89	94.18
3.5	95.94	95.95	96.22	94.80	93.88	93.67	93.84	93.84	93.93	93.81
3.6	9.31	74.70	90.59	93.50	94.02	93.92	94.09	93.63	93.71	94.08
3.7	0.05	52.38	89.17	93.65	94.03	94.08	93.89	93.88	93.80	93.38
3.8	0.00	31.65	83.15	93.02	93.96	94.26	93.71	93.75	93.78	93.70
3.9	0.01	14.49	77.26	92.29	93.50	93.99	94.25	94.13	94.03	93.51
4.0	100.00	99.99	97.09	93.09	92.88	93.25	93.54	93.12	93.03	92.72
4.1	0.00	15.14	77.54	92.32	93.53	93.67	93.35	93.83	92.97	92.40
4.2	0.00	32.07	82.94	93.20	94.61	94.20	94.26	92.75	**91.62**	**90.64**
4.3	0.04	52.36	89.65	93.74	93.88	93.60	93.53	91.92	89.91	87.12
4.4	8.83	75.12	90.88	93.31	93.44	92.95	91.56	**90.10**	88.04	85.22
4.5	95.55	95.93	95.56	95.12	93.78	92.29	**90.01**	87.17	84.02	79.18
4.6	9.12	75.29	90.43	93.46	93.72	**91.09**	88.00	82.58	75.93	70.28
4.7	0.11	52.43	89.94	95.26	**93.08**	87.79	79.77	71.85	63.04	55.39
4.8	0.00	32.71	87.84	**95.33**	90.40	80.56	67.27	56.05	45.54	36.71
4.9	0.00	14.37	76.37	86.81	69.78	49.36	36.04	24.88	18.20	13.93
5.0	100.00	100.00	94.70	61.47	28.08	11.65	5.47	2.75	1.40	0.81

Table 4 Effectiveness of $\bar{r}_{ji} \pm 1.96SEM$ in containing θ_j in $L(6,30)$

θ	σ									
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
1.0	100.00	100.00	94.54	60.23	27.94	11.76	5.17	2.40	1.31	0.71
1.1	0.00	14.89	76.45	87.23	69.76	49.16	35.37	24.99	18.09	13.39
1.2	0.00	32.72	87.42	**95.72**	**90.02**	80.60	68.17	54.83	44.27	36.19
1.3	0.06	52.06	90.49	95.57	93.36	87.73	80.10	71.18	63.02	54.60
1.4	90.35	74.98	91.08	93.97	93.70	**91.50**	87.51	82.58	76.49	69.80
1.5	95.8	95.94	96.01	95.75	94.27	92.19	**90.31**	87.29	83.55	79.34
1.6	8.85	75.09	90.38	93.76	93.85	93.51	91.55	**90.62**	88.06	84.47
1.7	0.03	52.69	89.45	93.50	94.02	93.29	93.51	91.78	**90.63**	87.81
1.8	0.00	32.46	83.41	93.15	94.66	94.23	93.53	93.01	91.84	**90.41**
1.9	0.00	14.93	77.06	92.29	93.28	93.70	94.53	93.56	93.09	92.07
2.0	100.00	99.99	97.05	93.01	92.79	92.84	93.25	93.13	92.88	92.88
2.1	0.00	14.65	76.93	92.13	93.32	93.62	94.28	94.00	93.96	93.40
2.2	0.00	31.89	83.45	92.91	94.17	93.89	93.52	93.56	93.73	93.84
2.3	0.08	53.11	89.13	93.79	93.75	93.54	93.49	93.57	93.93	94.16
2.4	9.09	75.20	90.39	93.59	94.07	93.80	94.07	93.95	94.41	93.80
2.5	95.94	96.24	95.79	94.80	94.30	93.67	93.60	93.97	93.69	94.27
2.6	8.84	75.57	90.51	93.77	94.67	93.91	94.59	93.92	94.13	93.82
2.7	0.04	52.50	89.12	93.55	94.13	93.41	93.94	93.93	93.80	94.00
2.8	0.00	32.22	83.02	92.92	94.21	94.38	94.18	93.70	94.18	93.95
2.9	0.00	14.59	77.04	94.73	93.25	93.85	94.00	94.22	94.85	94.24
3.0	100.00	100.00	96.71	93.44	92.52	93.31	93.60	93.65	93.89	93.90
3.1	0.00	14.15	77.29	92.17	93.47	93.76	94.43	94.07	94.27	94.34
3.2	0.00	31.48	83.13	92.96	94.33	94.58	93.73	93.83	94.03	94.08
3.3	0.04	52.57	89.20	93.70	93.80	94.09	93.70	94.08	93.80	94.10
3.4	9.07	74.89	91.15	94.12	93.84	94.14	94.10	94.09	94.03	94.41
3.5	95.8	95.78	95.22	95.13	94.16	93.74	93.84	93.60	94.26	94.10
3.6	8.85	75.34	90.88	93.75	93.87	94.03	93.90	94.02	94.02	94.29
3.7	0.07	53.91	89.16	93.63	93.87	93.8	93.64	94.05	93.77	93.66
3.8	0.00	32.31	83.87	93.03	94.40	94.28	93.86	94.11	94.36	94.14
3.9	0.00	15.31	77.50	92.59	93.38	93.96	94.85	94.42	94.16	94.19
4.0	100.00	100.00	96.77	93.26	92.97	93.32	93.48	93.60	93.45	94.12
4.1	0.00	15.69	76.65	92.06	93.57	94.04	94.05	93.90	94.45	94.21
4.2	0.00	31.88	83.28	93.18	94.13	94.22	93.60	93.81	94.08	94.33
4.3	0.04	53.73	89.06	93.90	94.04	93.89	93.99	93.47	94.40	93.75
4.4	8.65	74.68	90.64	93.75	93.82	93.93	93.74	93.72	94.28	94.17
4.5	95.73	95.93	95.74	94.93	93.53	93.38	93.67	93.76	94.24	94.18
4.6	8.80	74.57	90.70	93.75	94.41	93.85	94.43	93.58	93.90	94.34
4.7	0.04	52.58	88.2	93.23	93.48	94.33	93.73	93.62	94.17	93.68
4.8	0.00	32.72	83.76	93.06	94.86	94.32	94.14	93.96	94.24	94.01
4.9	0.00	15.26	76.88	92.75	93.53	93.52	94.45	93.93	93.95	93.76
5.0	100.00	100.00	97.1	92.94	92.71	94.02	93.29	93.37	92.86	92.49
5.1	0.00	15.46	76.42	92.90	93.1	93.59	94.40	93.98	93.30	92.34
5.2	0.00	32.31	83.02	93.32	94.47	94.04	94.12	92.56	92.33	**90.51**
5.3	0.05	53.34	89.42	93.99	93.92	93.68	93.06	92.19	**90.13**	87.84
5.4	8.90	75.50	90.37	93.34	93.89	92.75	92.19	**90.49**	87.62	85.84
5.5	95.98	96.11	95.48	95.37	94.36	92.56	**90.15**	86.90	83.61	79.47
5.6	8.56	74.88	91.09	93.63	93.80	**91.64**	88.38	82.47	76.39	69.97
5.7	0.06	52.83	90.52	95.44	93.57	88.04	80.09	72.62	63.45	55.36
5.8	0.00	32.83	87.34	**95.15**	**90.03**	79.84	68.04	55.61	45.40	35.92
5.9	0.00	14.77	77.26	86.98	69.98	50.60	36.62	29.84	18.68	13.41
6.0	100.00	100.00	95.00	60.01	27.87	11.87	4.97	2.32	1.43	0.86

Table 5 Distribution of Respondents Observed in the Study of Pornel et al. (2010)

Statistics	Range of values	% of respondents
θ	Below 1.8	0
	Between 1.8 & 3.2 inclusive	80
	Greater than 3.2	20
σ	Less than or equal to 0.3	5
	Greater than 0.3 but not more than 1.0	90
	Greater than 1.0	5

Third Simulation: Parameter Estimation Error Associated with Likert Scale as Affected by the Variance of the Respondent's Opinions

In this simulation, the researchers determine the difference between the theoretical theta θ and the estimated theta, $\hat{\theta}$, where $\hat{\theta}$ is taken to be equal to *E (R)*, given the value of standard deviation *s*. Results show that the root mean square error (RMSE) is high for s = 0.01but it decreases as s approaches 0.04. When the RMSEs were computed for different ranges of θ, it was found that in $L(4, n)$ the three ranges of θ differ in their rates of change as *s* increases as shown in Figure 2.

The RMSE of $L(4, n)$ for θ ranging from 1 to 4 is minimum at $\sigma = 0.35$, and the RMSE increases as σ increases. On the other hand, the RMSE of $L(4, n)$ for θ ranging from 1.5 to 3.5 is minimum at $\sigma = 0.45$. Lastly, the RMSE of $L(4, n)$ for θ ranging from 2 to 3 is minimum at $\sigma = 0.52$.

In $L(5, n)$, the RMSE is minimum for θ ranging from 1 to 5 at $\sigma = 0.38$, for θ ranging from 1.5 to 4.5 at $\sigma = 0.45$ and 0.46 and for θ ranging from 2 to 4 at $\sigma =$ 0.54 (see Fig. 3).

Lastly, when $L(6, n)$, was considered, the RMSE is minimum for θ ranging from 1 to 6 at $\sigma = 0.36$ and 0.37, for θ ranging from 1.5 to 5.5 at $\sigma = 0.46$ and for θ ranging from 2 to 4 at $\sigma = 0.53$ as shown Figure 4.

One may observe, that the optimum value of σ for the range *(u, v)* of θ is almost the same for $L(4, n)$, $L(5, n)$,and $L(6, n)$. The same is true for ranges $(u + .5, v - .5)$ and $(u + 1, v - 1)$**.** The relationship between σ and RMSE for $L\,(4, n)$ as presented in Figure 2 showed that the model is more accurate when the θ involved is near the middle of the spectrum than at the edges. As Figure 2 shows most of the errors are due to θ at the edges. The same trend was found for $L(5, n)$ and $L(6, n)$ as shown in Table 6 and Figures 3 and 4. Examining these results show that there is minimal difference in the RMSE among the three scales despite of the differences in the number of choices (4, 5 and 6). Thus, one may say that the error is more a function of the position of θ and the magnitude of σ and not of the number of choices or points of the Likert scale. Since *E(R)* do not vary with *n*, this simulation was not tested with different value of *n*.

Figure 2. Root Mean Square Error (RMSE) for $L(4, n)$ across Different Standard Deviation σ

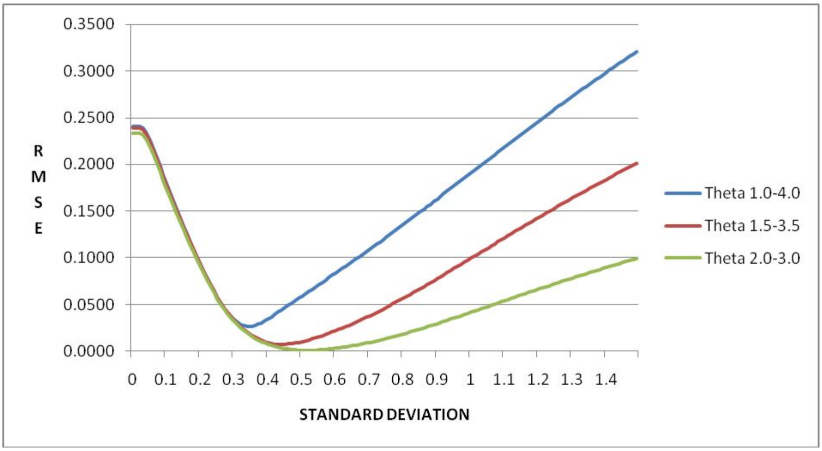


Figure 3. Root Mean Square Error (RMSE) for L (5, n) across Different Standard Deviation σ

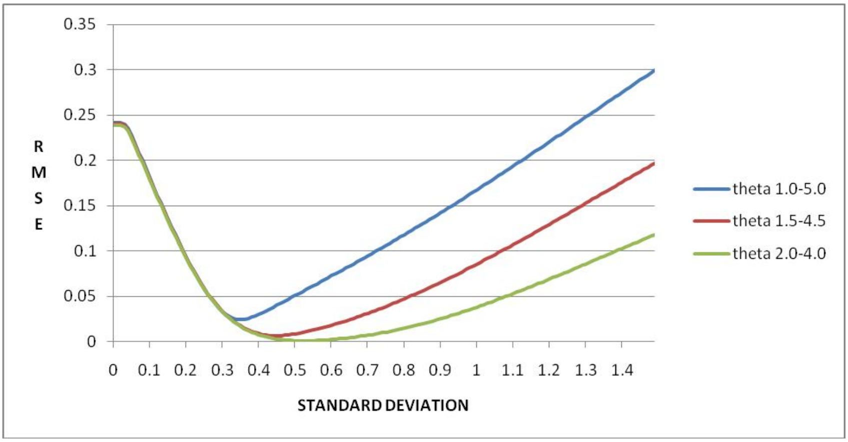


Figure 4. Root Mean Square Error (RMSE) for $L\ (5, n)$ across Different Standard Deviation σ

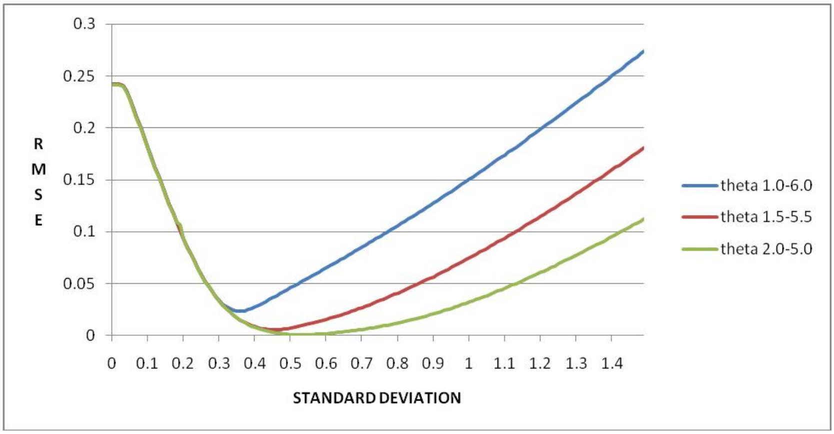


Table 6 Optimum Value of σ

Range of θ	Value of σ where RMSE is minimum		
	L (4, n)	L (5, n)	L (6, n)
(u, v)	0.35	0.38	0.36, 0.37
(u + .5, v - .5)	0.45	0.45, 0.46	0.46
(u + 1, v - 1)	0.52	0.54	0.53

5. Summary and Conclusions

The result of the first simulation showed that the model can generate a response vector with mean response not significantly different from the original response vector. Thus, for a Likert scale application that estimates the latent ability of the respondent using mean response, the NIRNDO is one good model to be considered.

The result of the second simulation showed that under the NIRNDO model, using the mean response to a Likert scale would work well in estimating the respondent's latent ability when the latent ability θ is at least 0.8 from the edges of the scale (that is, $u + 0.8 \le \theta \le v - 0.8$). Specifically, the interval $\bar{r}_{ji} \pm 1.96SEM$ contains θ at least 90% of the time when $u + 0.8 \le \theta \le v - 0.8$. This result, however, is consistent when $\sigma > 0.3$. It must be pointed out, however, that under the model, assuming $\bar{r}_{ji}$ is a normally distributed and continuous variable, the 95% confidence interval $\bar{r}_{ji} \pm 1.96SEM$ can be effective at least 90% of the time.

The result of third simulation supported the result of the second simulation. It shows that the optimum range of σ where the model has the minimum error is from 0.35 to 0.54 and that the model would not work well with θ near the edges.

The NIRNDO model could be a good model for Likert Scale, and that using mean response to estimate the latent variable being measured by the Likert Scale is feasible, although one need to be on lookout when faced with respondents who had extreme mean response, that is, mean response with 0.8 or less distance from one edge of the scale, or with very low or very high response variance. With this type of respondents, the model does not work well. Also, it is shown that Likert scale has less accuracy than a continuous scale. On the region where the model works well, an interval that would guarantee 95% accuracy for a continuous scale would guarantee 90% accuracy for Likert scale.

REFERENCES

AGUINIS, H., PIERCE, C., & CULPEPPER, S., 2008, Scale Coarseness as Methodological Artifact: Correcting Correlation Coefficients Attenuated from Using Coarse Scale. *Org. Research Methods,* 12(4):623-652.

ALBAUM, G., 1997, The Likert Scale Revisited: An Alternate Version (Product Preference Testing), *Journal of the Market Research Society,* 39(2):331-342.

ANDRICH, D., 1978, Application of a Psychometric Model to Ordered Categories which are Scored with Successive Integers, *Applied Psychological Measurement*, 2:581-594.

GAY, L. R., & AIRASIAN, P. W., 2000, *Educational Research* (6th ed.). NJ: Prentice-Hall.

JAMIESON, S., 2004, Likert Scales: How to (ab)use Them, *Medical Education, 38:*1212-1218.

KUZON, W.J., URBANCHEK,M.G., MCCABE,S., 1996, The Seven Deadly Sins of Statistical Analysis, *Annals of Plastic Surgery,* 37(3): 265-272.

LIKERT, R., 1932, A Technique for the Measurement of Attitudes, *Archives of Psychology,* 140:44-53.

MASTERS, G. N., 1982, A Rasch Model for Partial Credit Scoring. *Psychometrika,* 47:149-174.

MURAKI, E., 1992, A Generalized Partial Credit Model: Application of an EM Algorithm. *Applied Psychological Measurement*, 16: 159-176.

PORNEL, J., BAYLOCON, E., MAQUIRAN, D., CAUSING, F., & TAN, M. J., 2010, A Study on Motivations, Deterrents, Needs, and Attitudes towards Research Undertakings of Public Secondary School Teachers, *Danyag: Journal of Humanities and Social Sciences,* 15(1): 36-44.

RUSSELL & BOBKO, 1992, Moderated Regression Analyses and Likert Scales: Too Coarse for Comfort, *Journal of Applied Psychology* 77(3):336-342.

SUNG, H.J. & KANG, T., 2006, Choosing a Polytomous IRT Model using Bayesian Model Selection Methods. A paper presented at the National Council on Measurement in Education annual meeting in San Francisco, April, 2006.

UEBERSAX, J., 2006, Likert Scales: Dispelling the Confusion. Statistical Methods for Rater Agreement website. [Accessed March 14, 2008]. Available from http://ourworld. compuserve.com/homepages/jsuebersax/likert2.htm.

WU, C.H., 2007, An Empirical Study on the Transformation of Likert-Scale Data to Numerical Scores, *Applied Mathematical Sciences*, 1(58): 2851-2862.