

Using Likert Type Data in Social Science Research: Confusion, Issues and Challenges

Basu Prasad Subedi

PhD Scholar, Kathmandu University, School of Education, Hattiban, Lalitpur, Nepal
Email: bashusubedi_1@yahoo.com

Abstract

The purpose of this article is to analyze the prevailing confusions, issues and challenges regarding using of Likert type data in social science research. Review of scholarly works specifically journal articles and other related documents were the base of this academic discourse. Study results suggested that there were incompatible views about the using of the Likert type data among the community of practice. The key debates were regarding its reliability, ordinal/ interval scale dilemma, and issue of using midpoints. The study concluded that in calculating reliability of Likert type measurement Cronbach's alpha is better to use for interval nature of Likert data and ordinal alpha for ordinal nature. In relation to ordinal/interval nature of data, Likert items generate the ordinal scale data and that of Likert scale generates interval scale. Additionally, concerning midpoints issue, methodologically there is not any difference either using or not using midpoints in Likert type measurement. However, epistemologically it is better to use midpoints defining its meaning. Despite the wider use of Likert type data, it has limitations as well.

Keywords: Likert items, Likert scale, ordinal /interval dilemma, Cronbach's alpha, ordinal alpha

1. Introduction

1.1 Setting the Stage

I believe that human being is the complex in nature. It is difficult to measure the human attitude, character and personality because of its subjective nature. However, it is necessary to analyze human behavior applying various research processes. We need data to conduct any type of research. For this, appropriate tools are necessary to collect data. Over the years, numerous methods have been used to measure character and personality traits of the human (Likert, 1932). It is very difficult to measure attitude, character and personality in numerical form. The perceived need of this quantification lies in the necessity to transform an individual's subjectivity into an objective reality (Joshi, Kale, Chandel, & Pal, 2015). In qualitative research, there are the possibilities to explore human subjective feeling through follow up ideas, probe responses, and investigate motives and feelings (Bell, 1999). However, social scientists still rely on quantitative measures of attitudes, character and personality traits (Boone & Boone, 2012) which is a crucial but difficult task. In the way of searching solution, Likert (1932) introduced the summative method to measure attitudes which is now popularly known as Likert Scale and has been widely using as a tool to collect data specifically in survey research (Weng & Cheng, 2000).

One very common question is: Why Likert Scale is so popular among the researcher? In my M. Phil dissertation I also used Likert type measurement to collect data about the emotional intelligence competencies of the head teachers. Not only this, I have also been using Likert type measurement to collect data for my doctoral dissertation regarding head teacher leadership attribute and school climate. I motivated toward Likert type measurement as it has the capacity to measure the attitude of the respondents easily. Moreover, it is easy to make statements to capture the essence of a specific construct. It is also easy to understand and respondents feel easy to provide their perception through Likert type format. Supporting my view Revilla, Saris, and Krosnick (2014) state that agree – disagree scale can be used to measure a wide range of constructs. According to them visual display of the scale is easy on paper questionnaires or in web surveys as well. Administration of the questionnaire is also easier and quicker, since the scale needs only to be explained once to the respondents. For these reasons, agree – disagree scales may entail lower costs (e.g., less paper needed, less work for the interviewers, less preparation cost), which is always tempting. Furthermore, the long tradition of using agree – disagree scales in the social science research may inspire researchers to reuse established batteries of items using this response format, even if they yield lower quality data.

Despite the wider use of Likert type measurement as research tool in the social science research, so many confusions, issues, and challenges are still there. I experienced so many difficulties in using Likert type measurement in my M. Phil dissertation. I felt confusion establishing reliability of the Likert type measurement. Concerning literatures showed that in most of the cases Cronbach's alpha had been used to measure internal consistency or reliability of a psychometric instrument like Likert data. In other hand there were literatures criticizing the use of Cronbach's alpha and advocating ordinal alpha to measure internal consistency. Similarly, I also experienced the difficulty relating to the measurement scale of Likert type data. Community of practice has been used ordinal as well as interval measurement in Likert data interchangeably

in the same situation. Due to this measurement confusion, I felt difficulty in selecting appropriate statistical method to analyze the data. Moreover, I felt difficulty in the use of midpoint on Likert type measurement. The debate around this issue is the concern about whether the midpoints affect the reliability and validity of measurements.

I interested to study regarding these difficulties and wanted to find out a clear way out relating to the confusion, issues, and challenges of using Likert data in social science research. For this, in this article, I will first review the debate in using Likert data in social science research. On the basis of this review, then, I will suggest the appropriate way of resolving confusion, issues and challenges of using Likert data. I hope this paper tries to minimize the confusion, issues and challenges of using Likert data in research. In the following sections I try my best to discuss about the characteristics of Likert type data with its inception and also discuss about the stated confusion, issues and challenges. Throughout this paper I also share my own practice of using Likert data in my current PhD dissertation.

2. Discussion

Likert type data are commonly used to measure attitude providing a range of responses to a given question or statement. It is an essential tool in psychology and in social surveys, and is an ever-present method of collecting attitudinal data (Dittrich, Francis, Hatzinger, & Katzenbeisser, 2007). Typically, there are 5 categories of response: from strongly disagree to strongly agree, although there are arguments in favor of scales with 7 or with an even number of response categories (Jamieson, 2004). The original Likert scale used a series of questions with five response alternatives: strongly approve, approve, undecided, disapprove, and strongly disapprove. It can be combined the responses from the series of questions to create an attitudinal measurement scale (Boone & Boone, 2012).

Generally, Likert type data is taken as the discrete instead of continuous values, tied numbers, and restricted range (De Winter & Dodou, 2010). The reason for the frequent use of such data in social science research may be that the phenomenon to be measured most often only can be measured by nominal or ordinal scales (Jakobsson, 2004). The practice of employing statistical techniques with ordinal-level data appears to be common in educational and psychological research. Issues such as control of Type I error and subsequent interpretation problems can arise when ordinal data are employed in statistical analyses that require interval scale variables (Harwell & Gatti, 2001). Such kinds of challenges occur while using Likert type data. It is necessary to plow out the entire problem while using Likert type data and also necessary to search the common solution. For this, rigorous study of its development and uses is essential. Let me provide permission to explore its inception. How does Likert type data come into practice? The next section of the article briefly describes about its inception and development.

The Likert scale was developed and named after psychologist Rensis Likert. It was first introduced to the field in 1932 in an article entitled, "A technique for the Measurement of Attitudes" published in *Archive of Psychology* (Edmondson, 2005, p.127). Likert constructed his scale as a means of capturing an infinite number of attitudes in an ordinal scale format (Likert, 1932). According to Edmondson (2005) the ideas of Likert scale was expanded in the article titled, "A simple and Reliable Method of Scoring the Thurstone Attitude Scales" published in

1934 in the journal *Social Psychology*. It was designed to simplify the complexity of the Thurstone scaling technique. The Thurstone scaling method was the first widely used attempt to capture latent variables on a continuous scale, however, the method suffered several drawbacks including the use of multiple judges which was costly and could potentially lead to judging bias (Edwards & Kenney, 1946; Likert, Roslow, & Murphy, 1993 as cited in Rinker, 2014).

There was criticism on Likert scale. Likert tried to respond to criticism of the reliability analysis by publishing a similar study in 1934 with similar results, though there were also design faults with the analysis as well (Edmondson, 2005 as cited in Rinker, 2014). There were no articles found that utilized Likert scale prior to 1970 even though the Likert scale had been invented back in 1932 (Edmondson, 2005, p.130). Today Likert-type data have been used in many fields related to the social sciences. Since the beginning of Likert scale in 1932, there have been debates among the users about its best possible usability in terms of reliability and validity of number of points on the scale (Joshi & et al, 2015).

There is a conventional practice of calculating reliability index in Likert data. Most widely and frequently used reliability index is Cronbach's alpha. Cronbach's alpha is a statistic. It is generally used as a measure of internal consistency or reliability of a psychometric instrument like Likert data. In other words, it measures how well a set of variables or items measures a single, one-dimensional latent aspect of individuals. Though there is the great influence of Cronbach's alpha in calculating reliability index in Likert data it has repeatedly been misinterpreted and misused (Garrido, Abad, & Ponsoda, 2013). Most of the research work based on Likert data has been using Cronbach's alpha to calculate reliability.

Ordinal alpha is conceptually equivalent to Cronbach's alpha. The critical difference between the two is that ordinal alpha is based on the polychoric correlation matrix rather than the Pearson covariance matrix, and thus more accurately estimates alpha for measurements involving ordinal data (Gadermann, Guhn, & Zumbo, 2012). So, make internal consistency researchers calculate ordinal alpha which measures reliability more accurately than Cronbach's alpha for ordinal nature of items of the questionnaire (Gadermann, Guhn, & Zumbo, 2012). Ordinal reliability coefficient may differ from their non-ordinal counterparts because of their scaling assumptions. The non-ordinal coefficients focus on the reliability of the observed scores by treating the observed item responses as if they were continuous, whereas the ordinal coefficients focus on the reliability of the unobserved continuous variables underlying the observed item responses. In this way, the ordinal coefficients are nonparametric reliability coefficients in a nonlinear classical test theory sense (Lewis, 2007). In my PhD dissertation I have treated Likert data as an interval data and have used Cronbach's alpha to measure internal consistency of the scale.

The concept of ordinal *alpha* was introduced by Zumbo, Gadermann, and Zeisser (2007), and which was shown to estimate reliability more accurately than Cronbach's alpha for ordinal response scales. To calculate ordinal reliability coefficients, one needs to estimate a polychoric correlation matrix, and then calculate the reliability coefficient from the polychoric correlation matrix (R Development Core Team, 2011; Garrido, Abad, & Ponsoda, 2013). In this way, we use the different method of calculating reliability coefficient as per the nature of Likert type data. More specifically, if we develop Likert scale and take it as an interval scale we can calculate

internal consistency using Corbach's alpha. Otherwise, we use alternative method means ordinal alpha to calculate reliability of Likert data.

Before discussing about the analysis process of Likert type data I would like to discuss about the measurement scale suitable for Likert type data. Specifically, there are four measurement scales or types of data: nominal, ordinal, interval and ratio. Nominal scales are used for labeling variables, without any quantitative value. The gender, different brands of motorcycle, colors of the hair etc. are the some typical examples of nominal scale of measurement. It has two or more than two non ordered categories. Nominal scale having only two categories (e.g. male/female) is called dichotomous and that of more than two generally called simply categories (e.g. Color of the hair: brown, black, grey, others).

It is obvious that Likert scales have been developed to measure attitudes by asking people to respond to a series of statements about a topic, in terms of the extent to which they agree with them, and so tapping into the cognitive and affective components of attitudes (McLeod, 2008). According to McLeod (2008) a Likert-type data assumes that the strength/intensity of experience is linear, i.e. on a continuum from strongly agree to strongly disagree, and makes the assumption that attitudes can be measured. Respondents may be offered a choice of five to seven or even nine pre-coded responses with the neutral point being neither agree nor disagree. So, data obtained from such kind of responses have not merely categories, there is ordered categories. By this nature, Likert type data generally cannot be treated as nominal measurement as Likert data has specific ordered categories.

Considering this fact, Likert type data has been treated as ordinal level of measurement (Jaminsson, 2004). Ordinal scales arrange or rank things (Brown, 2011). According to Brown (2011) interval scales also show the order of things, but equal intervals between the points on the scale is also essential for the interval scale. In the case of Likert type data it is difficult to assume this. However, in practice it has been observed that Likert type data is being treated as interval scale. So, there is ordinal/ interval dilemma regarding Likert type data. In the case of ratio scale measurement, it differs from interval scales in that it has a zero value and points along the scale make sense as ratios. Treating Likert type data as ratio would be unusual as there is not a natural zero point.

Above discussion makes clear that there is the practice of treating Likert type data either ordinal or interval scales of measurement. In practice it has been observed that community of practice has randomly used both kind of measurement scale interchangeably while treating same sort of cases. Due to this reason, dilemma has been created regarding the analysis of Likert type data. To minimize such dilemma following section entitled "Likert Items and Likert Scale" has been developed.

Methods for ordinal variables utilize the category ordering (Agresti, 2002). But in practice Likert type data are also commonly treated as interval scales in our field (Brown, 2011). To overcome this confusion, a common consensus is necessary to establish. For this the concept of Likert items and Likert scale might play crucial role to overcome the dilemma. Likert items are single questions or statements which are mutually exclusive among each other whereas the Likert scale on the other hand is a series or battery of minimum four or more mutually inclusive

Likert items that are combined into a single composite score/variable during the data analysis process (Boone & Boone, 2012 & Joshi & et al, 2015). In my PhD dissertation I have developed Likert scale of different constructs of school climate and leadership attributes of the head teachers. I have been calculating single composite score of the each scales and treated Likert data as an interval data.

We capture the attitude and behavior of the participants through mutually exclusive items which are generally known as Likert items. Analysis of individual item is essential for this kind of Likert data. Through Likert scale, we collect the opinions or perceptions of the participants in a specific construct. Numbers of items related to particular construct are prepared as statements or questions and during the analysis the score of the entire items of a construct is combined to generate composite score.

The following example presented in Table 1 makes clear the concept of Likert – items and Likert scale.

Table 1

Example of Likert Items and Likert Scale

Items (Likert Items)	SD	D	N	A	SA
1 Lecture method of instruction is appropriate in every situation.	1	2	3	4	5
2 Strict discipline is necessary for learning.	1	2	3	4	5
3 Team work is the culture for success.	1	2	3	4	5
4 Distributive leadership is appropriate for the school improvement.	1	2	3	4	5
5 Socioeconomic factor influences student leaning behavior.	1	2	3	4	5
Items (Likert Scale)					
6 My head teacher demonstrates high moral standard.	1	2	3	4	5
7 My head teacher communicates high expectations.	1	2	3	4	5
8 My head teacher demonstrates commitment to the goals.	1	2	3	4	5
9 My head teacher displays enthusiasm and optimism.	1	2	3	4	5
10 My head teacher uses power for personal gain.	1	2	3	4	5

SD = Strongly Disagree, D = Disagree, N = Neutral, A = Agree, SA = Strongly Agree

To decide about Likert items and Likert Scale, we have to evaluate the characteristics of items such as arrangement of items in logical sequence; closely interrelated but provide some independent information; elements of coherence; and whether each item measures a distinct element of the issue (Joshi & et al, 2015). If we look out first five items they are separate and more or less mutually exclusive. Hence, these items can't combine to form a scale. It is necessary to analyze them separately. So, these items are the Likert items. While multiple questions may be used in a research instrument, there is no attempt by the researcher to combine the responses from the items into a composite scale (Boone & Boone, 2012). On the other hand, from item 6 to 10 there are items in logical sequence, closely interrelated, elements in coherence, and each item has the capacity to measure a distinct element of the issue. The distinct element of issue in this case is *transformational leadership behavior of a head teacher*. Due to these characteristics, the items may be combined to construct a composite index which measures the collective stance of the participants towards phenomenon under study and may represent as an example of Likert scale (Joshi & et al, 2015). In Likert scale, combined items are used to provide a quantitative measure of a character or personality trait (Boone & Boone, 2012).

After identification of Likert items and Likert scale, now we can trace out a distinct way to analyze the data obtained from two forms of Likert questionnaire. Now question arises about the forms of data means either Likert data is ordinal or interval. A number of scholarly research literature argued or assumed that Likert items do not form an interval scale, and should be considered ordinal scales and should be analyzed accordingly (Jakobsson, 2004; Jamieson, 2004 as cited in Brown, 2011). Empirical researches have supported the interval view and opined that it is perfectly all right to use the summed scales to conduct parametric tests (Carifio & Perla, 2007). To resolve the debate of ordinal/ interval scale Joshi and et al. (2015) noted that in the case of Likert scale measurements the composite score for all the items of a scale can be compared for another individual on an interval scale.

Likewise, Creswell (2008) suggests that for Likert data to be treated as interval data there is need to develop multiple categories within a scale, establish equality of variance between each value on the scale and normality of the data. Additionally, Pell (2005) as cited in Murray (2013) agrees that parametric tests can be conducted on the summed scores of Likert scale data provided that the assumptions are clearly stated and the data is of the appropriate size and shape. In my PhD dissertation I have been using parametric tests such as individual samples - *t* test, paired samples - *t* test and multiple regression analysis to analyze desired and school climate and head teacher leadership attribute of the community secondary school of Nepal after fulfilling assumptions needed for parametric test.

There are confusions, issues and challenges in analyzing Likert type data. Many novice researchers fall in confusion about using the appropriate statistics to analyze Likert type data. Due to the random practice of the established scholars and researcher, more confusion has been surfaced about the analysis. Reviewers of research reports frequently criticize the choice of statistical methods in the case of Likert type data (Norman, 2010). Consistency is not observed in the literature. Some researchers or community of practice uses Likert type data as ordinal where as others use it as interval. In most of the cases clear justification is not found about this ordinal/interval issue. Due to this confusion in scale the problem has transferred into the use of

statistics. There is the debate and random practice of using parametric and non parametric statistical method while analyzing the Likert type data.

Jakobsson (2004) conducted a review study about the presentation and analysis of ordinal data in three internal nursing journals in 2003. In total 166 full – length articles were reviewed during the study. Review showed that ordinal scales were used in about one third of the articles. However, only about half of the articles that used ordinal data had appropriate data presentation and only about half of the analyses of the ordinal data were performed properly. This empirical evidence clearly shows the confusion, issues and challenges in analyzing Likert type data. Incorrect presentation and analysis of the data may lead to bias and reduced ability to detect statistical differences or effects, resulting in misleading information. This highlights the importance of knowledge about data level, and underlying assumptions for the statistical tests must be considered to ensure correct presentation and analyses of data.

Choice of statistics for analysis depends upon the nature of Likert data. Issues such as control of Type I error and subsequent interpretation problems can arise when ordinal data are employed in statistical analyses that require interval scale variables. Many statistical procedures used in educational and psychological research require the assumption of a normally distributed dependent variable, which presumed at least interval-level data (Harwell & Gatti, 2001 as cited in Shaftel, 2012). Though there are numbers of research work which has been advocating and using Likert data as interval scale, the legitimacy of assuming an interval scale for Likert type categories is an important issue because the appropriate descriptive and inferential statistics differ for ordinal and interval variables and if the wrong statistical technique is used, the researcher increases the chance of coming to the wrong conclusion about the significance of his/her research (Jamieson, 2004).

I believe that much of this ordinal/interval confusion arises from the fact that many authors use *Likert scale* to refer to both the Likert items and Likert scales (sums or averages of the results on sets of Likert items) which I described in the previous section with a set of examples (Table 1). After analyzing the scholarly works and empirical evidences I perceived that it minimizes the confusion so we must think about individual Likert items and Likert scales in different ways and we should have the concept of the fundamental difference between Likert items and Likert scale (Brown, 2011). After this clear perception between Likert items and Likert scale now we can develop the concept of ordinal and interval nature of data. As per the discussion made in the previous section the data obtained from Likert items which is ordinal and that of Likert scale is interval. This is the common consensus obtained from the discussion till now.

The analyses of Likert items and Likert scale data require unique data analysis procedures. Once a researcher understands the difference between Likert items and Likert scale data, the decision on appropriate statistical procedures will be clear (Boone & Boone, 2012). The challenges of selecting appropriate statistical method depend upon the designing of Likert instruments (Joshi, et al, 2015). If we sincerely combine the items in order to generate a composite score (Likert scale) of a set of items for different participants then the assigned scale will be an interval scale otherwise it becomes ordinal.

More specifically, there are two types of statistical methods of data analysis: parametric and non parametric. To use parametric method of analysis some specific criteria such as need of large sample size, normally distributed data, continuous nature of data etc. has to be fulfilled. So, while using parametric statistical method in Likert data we have to face these challenges as well (Jamieson, 2004). Hence, a parametric test is designed for continuous, normally distributed data (with equal variance), and a nonparametric test is based on ranks.

From this academic argument and discussion it is clear that there are confusions, issues and challenges in using Likert data. So, sincere attention should be paid while using it. To analyze Likert data properly, we must understand the measurement scale represented by each. Now it is clear that numbers assigned to Likert items express a *greater than* relationship; however, how much greater is not implied. Likert items fall into the ordinal measurement scale because of these conditions. So, we use nonparametric statistical methods only to analyze the ordinal nature of data. Descriptive statistics such as mode or median for central tendency and frequencies for variability can be used to analyze Likert items or ordinal data. Additional analysis procedures appropriate for ordinal scale items include the chi-squared measure of association, Kendall Tau B, and Kendall Tau C (Jakobsson , 2004). Standard texts also advise that the appropriate inferential statistics for ordinal data are those employing non-parametric tests, such as Spearman's Rho or the Mann–Whitney U-test (Jamieson, 2004). However, there is violation of these rules in practice. Treating ordinal scales as interval scales have long been controversial. To make clear idea and search a proper tract is the challenge of today while analyzing Likert data.

To analyze Likert scale or interval data, first of all it is necessary to create Likert scale items by calculating a composite score (sum or mean) from four or more Likert items then we can use parametric statistic such as mean for central tendency and standard deviation for variance. Composite score can be calculated by transforming the data of each items of a scale using SPSS. Additionally, as per the need we can use Pearson's *r*, *t*-test, ANOVA, and regression procedures (Boone & Boone, 2012). Despite all this discussion of the Likert type data, most of the research based on Likert items and scales that I have seen in our field treats them as interval scales and analyzes them as such with descriptive statistics like means, standard deviations, etc. and inferential statistics like correlation coefficients, factor analysis, analysis of variance, etc. (Brown, 2011).

In behavioral research, while analyzing data, most of the quantitative methods correlation analysis is made. This method is the most frequently misused method in social and behavioral research specially when analyzing ordinal data as Likert items (Choi, Peters & Mueller, 2010). There are numbers of correlation analysis options. However, there is the inadequate literature about the proper use of those correlation methods. These methods range from the traditional Pearson's *r*, to more recent developments, such as Bayesian estimation of polychoric correlations (Choi, Peters & Mueller, 2010). A polychoric correlation is the appropriate statistic to measure associations between ordinal data. Polychoric correlations possess desirable properties similar to Pearson's *r* (Choi, Peter & Mueller, 2010).

To compare two independent groups, Mann–Whitney U-test can be adopted (Jamieson, 2004). An advantage with this test is that the two samples under consideration may not necessarily have the same number of observations. It is often presented as an alternative to *t* test when the data are

not normally distributed. Whereas a *t* test is a test of population means, the Mann-Whitney test is commonly regarded as a test of population medians.

Table 2 provides summary of data analysis procedures for Likert items and Likert scale data.

Table 2

Summary of Data Analysis Procedures for Likerttype and Likert Scale Data

Statistical Methods	Likert Items Data	Likert Scale Data
Internal consistency	Ordinal alpha	Cronbach's alpha
Central Tendency	Median or mode	Mean
Variability	Frequency	Standard deviation
Associations	Kendal tau B or C, Spearman's Rho, polychoric correlations	Pearson's r
Other statistics	Chi-square, Mann-Whitney U-test	ANOVA, t – test, regression

Additionally, while using Likert type data in research it is necessary to clarify about the use of midpoints as well. One concern among researchers about using midpoints on a Likert items or scale is the effects of the midpoints to the reliability and validity of measurements. Some researchers claim that the use of midpoints on Likert scale may affect research reliability and validity, but some other researchers disagree. Generally, the supporters of midpoint opinions claim that the midpoints can increase the reliability of measurement. Whereas the opponents of midpoint opinions claim that the midpoints cannot increase the reliability of measurement (Tasang, 2012). In this sense, midpoints are not necessary to benefit the internal consistence of measurements.

Some studies evaluate the impacts of midpoints on measurement validity. However, the findings are also contradictory. For instance, some studies find that the construct validity may not be influenced by the midpoints (Adelson & McCoach, 2010) but some researchers suggest the keeping out of the midpoints may weaken the validity (Johns, 2005). This discussion reveals that midpoints are not necessary to increase validity of measurements. In some cases midpoints may be viewed by the respondents as a “dumping ground” for unsure or non-applicable responses as well (Kulas, Stachowski & Haynes, 2008). The discussions till here reveals that from the methodological viewpoint both use and not use of midpoints are acceptable in Likert type measurement as the midpoints may not affect reliability and validity. However, it is necessary to pay attention concerning its epistemological issue while designing the Likert items or scale of a measurement (Tasang, 2012). It means it is necessary to clarify the using of midpoints in the Likert scale. There might be the multiple meaning of midpoints such as “neutral”, “undecided”

“don’t know” or “neither agree nor disagree” in terms of epistemological issue. In my PhD dissertation I have used midpoints giving the meaning of neither agree nor disagree, meaning in between disagree and agree in five point Likert scale considering epistemological connotation.

Despite dilemma about the using Likert data, it has set proper stage for the academic discourse. The community of practice has been advocating and using the various ideas regarding Likert type data. From its inception to till date the discussion about its implication in social science research has been continuously accelerating. The debate from different angles such as Likert items and Likert scale data, ordinal/ interval issues, use of parametric and non parametric statistical methods to analyze the data, using midpoints in Likert type data etc. really created a meaningful discussion and it has established as one of the tools of research in its more than eighty years long age. So, grounded confusions, issues, and challenges always keep it alive in the process of research. In comparison to other study tools in quantitative study the role of Likert scale lies in topmost position. Due to several controversial characteristic of Likert type data it has been getting time from the significant numbers of researcher to make it strong tool in the field of social science research.

Despite wider use of Likert type measurement in social science research, it can be observed its limitation as well. Because the Likert type data measures subjective feelings, results may change depending upon when a person takes the test. In my personal experience situation or context influences the view of respondents. Further, the Likert type data does not typically assign different weights to different statements. In practice, so many biases such as central tendency bias (e.g. respondents may avoid extreme response categories), acquiescence bias (e.g. respondents may agree with statements as presented in order to “please” the experimenter), social desirability bias (e.g. portray themselves in a more socially favorable light rather than being honest) can be experienced (Bertram, n. d.). Likewise, it is difficult to achieve internal consistency of the scale and also difficult to demonstrate validity. Additionally, in my practical experience I have observed that results are easily faked where individuals want to present a false impression of their attitudes. Equally, it is time consuming and severe job to make good attitude statements.

3. Conclusion

The purpose of this article was to analyze the prevailing confusions, issues and challenges regarding using of Likert type data in social science research. The study concluded that in calculating reliability of Likert type measurement Cronbach’s alpha is better to use for interval nature of Likert data and ordinal alpha for ordinal nature of Likert data. Regarding ordinal and interval nature of data, Likert items generate the ordinal scale data and that of Likert scale generates interval scale data. This concept of Likert items and Likert scale contributes to minimize the confusion, issues and challenges about the structure of Likert tool and guideline for choosing appropriate statistical method (either parametric or non parametric) to analyze and interprets Likert type data. Additionally, in relation to the issue of using midpoints in Likert type data methodologically there is not any difference either using or not using midpoints in Likert type measurement. However, epistemologically midpoints it is better to use defining its nature of meaning (e. g. neutral, undecided, don’t know or neither agrees nor disagrees). Despite its wider use, it has limitation as well. Context may influence the reality. Biases such as central

tendency bias, acquiescence bias, social desirability bias might be faced while using Likert type measurement. Eventually, it is hoped that this article would become a proactive manuscript to minimize the mistake/ misuse of Likert type data in social science research advancing its prior knowledge.

Reference

- Adelson, J. L., & McCoach, D. B. (2010). Measuring the mathematical attitudes of elementary students: The effects of a 4-point or 5-point Likert-type scale. *Educational and Psychological Measurement, 70*(5), 796-807. doi: 10.1177/0013164410366694
- Agresti, A. (2002). *Categorical data analysis*. New Jersey, NJ: John Wiley & Sons, Inc.
- Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress, 40*(7), 64-65.
- Ary, D., Jacobs, L. C., & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). California: Thomson Wadsworth.
- Bell, J. (1999). *Doing your research project*. Buckingham: Open University Press.
- Bertram, D. (n.d.). *Likert scales*. Retrieved from <http://poincare.matf.bg.ac.rs/~kristina/topic-dane-likert.pdf>
- Boone, H. N. & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension, 50* (2). Retrieved from <http://www.joe.org/joe/2012april/tt2p.shtml>
- Brown, J. D. (2011, March). Likert items and scales of measurement? *Shiken: Jalt Testing and Evaluation SIG Newsletter, 15*(1) 10-14.
- Burton L. J, & Mazerolle, S. M. (2011). Survey instrument validity part I: Principles of survey instrument development and validation in athletic training education research. *Athl Train Educ Journal 6*(1), 27-35.
- Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education, 42* (12), 1150-1152. doi:10.1111/j.1365-2923.2008.03172x
- Choi, J. Peters, M. & Mueller, R. O. (2010). Correlational analysis of ordinal data: from Pearson's r to Bayesian polychoric correlation. *Asia Pacific Educ. Rev. 11*, 459-466. doi: 10.1007/s12564-010-9096-
- Clason, D. L., & Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education, 35*(4), 31- 35.
- Creswell, J. (2008). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. Upper Saddle Reiver, New Jersey: Pearson Education Inc.
- De Winter, J. C. F., & Dodou, D. (2010). Five-point Likert items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation, 15*, (11). Retrieved from <http://pareonline.net/getvn.asp?v=15&n=11>

- Dittrich, R., Francis, B., Hatzinger R. & Katzenbeisser, W. (2007). A paired comparison approach for the analysis of sets of Likert-scale responses. *Statistical Modelling* 7 (1), 3 - 28. doi: 10.1177/1471082X0600700102
- Edmondson, D. R. (2005). Likert scales: A history. *CHARM*, 12, 127–133. Retrieved from <http://learn.quinnipiac.edu/charm/CHARM%20proceedings/CHARM%20article%20archive%20pdf%20format/Volume%2012%202005/127%20edmondson.pdf>
- Gadermann, A. M., Guhn, M. & Zumbo, B.D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation* 17(3). A peer-reviewed electronic journal.
- Garrido, L.E., Abad, J.J. & Ponsoda, V. (2013). A new look at Horn's parallel analysis with ordinal variables. *Psychological Methods* 18(4), 454-474.
- Harwell, M. R. & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research*, 71(1), 105-131. doi:10.3102/00346543071001105
- Jakobsson, U. (2004). Statistical presentation and analysis of ordinal data in nursing research. *Scand J Caring Sci*, 18, 437–440.
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38, 1212 -1218. doi: 10.1111/j.1365-2929.2004.02012.x
- Johns, R. (2005). One size does not fit all: Sleeting response scale for attitude items. *Journal of Election, Personal Opinion and Parties*, 15 (2), 237 – 264.
- Joshi, A. Kale, S. Chandel, S. & Pal, D. K. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 7(4), 396-403.
- Kulas, J. T., Stachowski, A. A., & Haynes, B. A. (2008). Middle response functioning in Likert-responses to personality items. *Journal of Business and Psychology*, 22(3), 251-259. doi: 10.1007/s10869-008-9064-2
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 1–55.
- McLeod, S. A. (2008). *Likert scale*. Retrieved from www.simplypsychology.org/likert-scale.html
- Murray, J. (2013). Likert data: What to use, parametric or non-parametric? *International Journal of Business and Social Science*, 4 (11), 258 -264. Retrived from http://ijbssnet.com/journals/Vol_4_No_11_September_2013/23.pdf
- Rademaker, S. G. (2010). An algorithm for converting ordinal scale measurement data to interval/ratio scale. *Educational and Psychological Measurement* 70(1), 74–90. doi: 10.1177/0013164409344532
- Revilla, M. A., Saris, W.E. & Krosnick, J. A. (2014). Choosing the number of categories in agree–disagree scales. *Sociological Methods & Research*, 43(1), 73-97. doi: 10.1177/0049124113509605

- Rinker, T. (2014). *Treatment of Likert data*. University at Buffalo. Retrieved from http://www.researchgate.net/publication/262011454_Likert
- Sauro, J. (2015 July). 5 things to know about Likert scales. *Measuring U*. Retrieved from <http://www.measuringu.com/blog/likert-scales.php>
- Scales of Measurement. (2006, August). *California State University, Fresno, Department of Psychology*. Retrieved from psych.csufresno.edu/psy144/Content/Measurement/scales.html
- Shaftel, J. (2012, April). *Effects of the number of response categories on rating scales*. Roundtable presented at the annual conference of the American Educational Research Association, Vancouver, British Columbia: University of Kansas.
- Tasang, K. K. (2012). The use of midpoint on Likert scale: The implications for educational research. *The Hong Kong Teacher's Centre Journal*, 11, 121 – 130.
- Weng, L. & Cheng, C. P. (2000). Effects of response order on Likert-type scales. *Educational and Psychological Measurement*, 60, 908 -924. doi: 10.1177/00131640021970989
- Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, 6, 21-29.