

DESIGNING A LIKERT-TYPE SCALE QUESTIONNAIRE AND ITS ANALYZING PROCEDURE

**Mohan Kumar Sharma¹, Shanti Prasad Khanal²,
Bhagwan Aryal³, Laxman Sigdel⁴, Samjhana Subedi^{5*}**

1. Graduate School of Education (GSE), Tribhuvan University, Kathmandu Nepal,
Email: attig2019@gmail.com, ORCID: <https://orcid.org/0000-0002-7600-8223>
2. Central Department of Health Education, Tribhuvan University, Kathmandu Nepal,
Email: khalshanti100@gmail.com, ORCID: <https://orcid.org/0000-0001-6001-9502>
3. Central Department of Health Education, Tribhuvan University, Kathmandu, Nepal
Email: bhagwan.aryal@cded.tu.edu.np ORCID: <https://orcid.org/0000-0002-9215-3551>

4. Department of Orthopedics, Bheri Hospital, Banke, Nepal,
Email: lax3210@gmail.com ORCID: 0009-0003-5396-791X

5. Central Department of Population Education, Tribhuvan University, Nepal,
Email: ssdrs044@gmail.com, ORCID: <https://orcid.org/0009-0006-1973-1833>

*Corresponding Author : **Samjhana Subedi,**

Email: ssdrs044@gmail.com ORCID: <https://orcid.org/0009-0006-1973-1833>

Abstract

This article aims to provide comprehensive guidance on distinguishing between Likert-type and Likert-scale survey items, as well as the appropriate statistical approaches for analyzing their intended measurements. The study involved reviewing scholarly articles published in journals and other relevant documents pertaining to Likert-related issues. It is commonly observed that odd-type Likert items use a five-point response scale, while even-type Likert items utilize a four-point response scale. Likert-type items consist of distinct and mutually exclusive questions that cannot be combined to form a scale, and therefore, need to be analyzed separately. On the other hand, Likert-scale items are comprised of questions that are logically sequenced, closely interrelated, and can be combined to create a composite index. When it comes to measuring the central tendency of Likert-type data, the mode and median are appropriate measures for ordinal scale data. For Likert-scale data, which should be in an interval scale, the standard deviation is a suitable measure, but it is important to calculate a composite score beforehand. In order to assess the internal consistency of Likert-type survey items, Cronbach's alpha is suitable for interval scale data, while ordinal alpha is appropriate for Likert survey items of an ordinal nature.

Keywords: Cronbach's alpha, design, Likert items, Likert scale, Likert-type, ordinal alpha

Introduction

A Likert scale is a rating scale that researchers use to assess participants' opinions, attitudes, and behaviors. Generally, it is used in survey research, which allows researchers to operationalize personality traits or perceptions quickly. However, it can also be used in qualitative studies. Although several kinds of rating scales have been developed, the most widely used is the Likert scale. In general, the Likert scale

has a five or seven-point scale used to allow the individual to express the extent to which they agree or disagree with a particular statement. Furthermore, it provides five or seven possible answers to a statement that allows respondents to express their positive or negative strength/marking of agreement regarding the statement.

In 1932, the Likert scale was developed by Rensis Likert (Likert, 1932). The Likert scale has 5-point, 7-point, and 9-point pre-coded responses, with a neutral point being neither agree nor disagree on an ordinal scale. This measures the respondents' perceptions towards a given statement and rates the degree to which they agree or disagree. Although the responses can be rated on an ordinal scale, the difference between each response cannot be measured. Even though numbers are assigned in responses, researchers cannot differentiate between the responses in terms of the exact distance they are different. In contrast to this, the difference between responses can be measured in interval scale data, where the rated numbers refer to each response (Sullivan & Artino, 2013).

There are various Likert questions or statements: frequency measurement, agreement measurement, and satisfaction measurement, usually with 5 to 7 or 9 possible response items in odd type and 4 to 6 or 8 possible response options in even Likert response items. Each item is given a numerical score so that the data can be analyzed quantitatively (Table 1).

Likert scales are frequently used in social and educational research. Thus, understanding, interpreting, and analyzing data derived from Likert scales is imperative for researchers working in social science, education, and academic research. The study's prominent objective is to provide basic concepts, including extensive statistics, while analyzing and interpreting Likert-type and Likert scale data for those readers who do not have a fundamental understanding of it.

Designing Likert-Type/Scale Question Statements

Several important things should be considered while developing Likert questions/statements. Firstly, each sub-theme requires ≥ 4 questions/statements that assess a single attitude when response scores are combined. Each statement measures a separate component of overall themes/subthemes. For instance, if a researcher wants to assess respondents' attitudes towards hygiene behaviors, they can design a variety of statements that measure different aspects of hygiene behaviors of the respondents. Secondly, the researcher can develop both statement and question types of Likert phrasing. It is significant to determine how to formulate questions and statements to ensure that they are easily understood and do not bias the researcher's respondents in any way. Thirdly, use both positive and negative statements in questionnaires. Formulating questions only one way, either positively or negatively, may create an information bias towards agreeing or

disagreeing with all of them. Fourthly, be sincere while developing statements because double negatives can confuse respondents and they may be unsure of what they are agreeing to on the statements. Fifthly, do not make statements that give dual meaning, which can create more confusion for the respondents. Sixthly, commonly, it has 5 to 7 or even 9-point Likert items in practice. Fewer items (5 points) mean the study captures less detail; however, the scale is more user-friendly. On the other hand, more items (7 to 9) provide deeper knowledge of the research but make it difficult for participants to respond to the questionnaires. So, before developing a questionnaire, the researcher should think carefully about the extent of depth of the study they are going to conduct. It depends on the study's nature and intensity.

This is a psychometric scale named after an American psychologist, Rensis Likert, which helps to measure opinions and perspectives towards statements relating to study issues. It is a closed-ended question that has pre-formulated answer options and the answer choices range from one extreme point to another, such as negative marking (strongly disagree) to positive marking (strongly agree) with the given statements.

There are several Likert points that are divided into two major types: Odd and Even Likert points. The first types, odd Likert scale questions, have a neutral answer option, which is designed for respondents who might not have a strong positive or negative opinion. In such cases, the neutral answer option is appropriate for them to choose from. This option is more appropriate for them to divide marks for negative and positive extremes but cannot be interpreted as either positive or negative marks. Odd Likert-type/scale questions could have 3 to 9 point Likert answer options, i.e., i) 3 points (Disagree, Neutral, Agree), ii) 5 points (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree), iii) 7 points (Strongly Disagree, Disagree, Somewhat Disagree, Neutral, Somewhat Agree, Agree, Strongly Agree), and 9 points (Strongly Disagree, Disagree, Moderately Disagree, Neutral, Slightly Agree, Moderately Agree, Agree, and Strongly Agree).

In the second type, even Likert questions have no neutral options, so basically, they are known as forced scales because respondents have to choose from either positive or negative ends. Some researchers use even Likert scales to collect extreme feedback without providing a neutral option. In the even Likert-type/scale, Likert question comprises 2 to 8 point answer options such as i) 2 points (Disagree and Agree), ii) 4 points (Strongly Disagree, Disagree, Agree, Strongly Agree), iii) 6 points (Strongly Disagree, Disagree, Somewhat Disagree, Somewhat Agree, Agree, Strongly Agree), iv) 8 points (Strongly Disagree, Disagree, Somewhat Disagree, Slightly Disagree, Slightly Agree, Somewhat Agree, Agree, Strongly Agree). This type of Likert type/scale allows researchers to include minimum options to maximum

variation without a neutral option. Table 1 show the 5-point odd Likert type/scale and 4-point even Likert type/scale answer options, which are maximally used in social science and education research.

Table 1

Sample of 5 Points Odd Response and 4 Points Even Response

S. N.	Questions	5 Points Odd Response					4 Points Even Response			
		SD	D	N	A	SA	SD	D	A	SA
1	My school's drinking water is potable. (attitude water)	1	2	3	4	5	1	2	3	4
2	My school's toilet is clean. (attitude sanitation)	1	2	3	4	5	1	2	3	4
3	My school has proper management of garbage. (attitude hygiene)	1	2	3	4	5	1	2	3	4
4	My teacher thinks that I can use the toilet properly. (subjective norm)	1	2	3	4	5	1	2	3	4
5	I believe that I can use the toilet properly. (Self-efficacy)	1	2	3	4	5	1	2	3	4
6	I intend to keep my school clean and healthy. (intention)	1	2	3	4	5	1	2	3	4

Note: SD=Strongly Disagree, D=Disagree, N=Neutral, A=Agree, and SA=Strongly Agree.

It depends on the nature and intensity of the study to determine how many-point Likert answer options should be used. If the researcher uses the minimum points Likert answer options, it makes it easier to analyze the data. On the other hand, if the researcher uses a high variance of Likert-point answer options, it provides a better reflection of the various ways in which respondents feel/respond to the statement. Additionally, as mentioned in Table 1, other point Likert-type/scale answers can be applied in research, which provide a range of options to express the extent to which respondents agree and are satisfied with the given statement. However, researchers rarely use these options as having more answer choices can overwhelm users and lead to random selection of answers. Although using more Likert option answers provides greater measurement precision, it is less commonly used because it confuses respondents with too many options in a statement.

Understanding Likert-Type or Likert Scale Data

Before analyzing Likert data, it is important to consider whether the study items are Likert-type or Likert scale. Researchers need to evaluate the characteristics of the items, such as the arrangement of items in a logical sequence, their interrelatedness, coherence, and whether each item measures a distinct element of the issue (Joshi & Pal, 2015).

If each item within a sub-theme is separate and mutually exclusive, they cannot be combined to form a scale and need to be analyzed separately. This is known as

Likert-type (Subedi, 2016). When researchers analyze each item individually, it is referred to as Likert-type and is treated as an ordinal scale. In this Likert-type, each item is ranked in an ordinal scale from higher to lower order, but it does not clearly define the exact differences between the responses. For example, referring to Table 1, we cannot be sure of the exact difference between "ever," "rarely," "sometimes," "often," and "always." In research tools, there are multiple items within the themes/sub-themes, and researchers do not attempt to combine the responses from the items into a composite scale (Bone & Bone, 2012).

On the other hand, if research items are logically sequenced, closely interrelated, coherent, and each item measures a distinct element of the issue, they can be combined to construct a composite index known as a Likert scale. Likert scale analysis measures the collective stance of the participants towards the phenomenon under study (Joshi & Pal, 2015). Authors Boone and Boone (2012) further state that combined items are used to provide a quantitative measure of a characteristic or personality trait in the Likert scale. When constructing a composite index in a Likert scale, there should be at least four or more items, and the scale is treated as an interval scale.

In summary, if the questions are individually analyzed for deeper insights into specific attributes, it is Likert-type. On the other hand, if the questions are analyzed by their composite score, it is Likert-scale data. In other words, if the researcher combines the items and generates a composite score from a set of items from different participants, the assigned scale is the Likert scale, which is considered an interval scale. If the researcher separately analyzes each item without calculating composite scores, it is Likert-type and consists of an ordinal scale.

Table 2

Sample of Likert-Type and Likert Scale Questionnaire

Likert-type items			Response				
S. N.	Questions	SD	D	N	A	SA	
1	My school's drinking water is potable. (attitude water)	1	2	3	4	5	
2	My school's toilet is clean. (attitude sanitation)	1	2	3	4	5	
3	My school has proper management of garbage. (attitude hygiene)	1	2	3	4	5	
4	My teacher thinks that I can use the toilet properly. (subjective norm)	1	2	3	4	5	
5	I believe that I can use the toilet properly. (Self-efficacy)	1	2	3	4	5	
6	I intend to keep my school clean and healthy. (intention)	1	2	3	4	5	
Likert scale items							
7	My school toilets are clean and usable.	1	2	3	4	5	
8	I feel comfortable using school toilets.	1	2	3	4	5	
9	I think my school has separate and sufficient toilets for boys, girls, and differently-abled students.	1	2	3	4	5	

10	I think human faces and excreta should be disposed of in separate places instead of leaving in the open.	1	2	3	4	5
11	I think it is essential to wash anal after using the toilet.	1	2	3	4	5
12	I think hands should be properly washed with water and soap after using toilet.	1	2	3	4	5

Note: *SD=Strongly Disagree, D=Disagree, N=Neutral, A=Agree, and SA=Strongly Agree.*

Analysis of Likert-Type or Likert Scale Data

After identifying the Likert-type and Likert scale, the researcher can determine an appropriate approach to analyze the data obtained from survey items. The selection of suitable statistical methods depends on the design of the survey items, whether they are Likert-type or Likert scale. Generally, Likert-type data is considered discrete rather than continuous values, tied numbers, and with a restricted range (Winter & Dodou, 2023). The frequent use of such data in social science research may be due to the fact that the phenomenon being measured can often only be measured using nominal or ordinal scales (Jakobsson, 2004). Likert-type data shows a relationship between responses, such as higher or lower, but does not imply the extent of the difference. Therefore, Likert-type data falls into the ordinal measurement scale. Researchers can use nonparametric statistical methods to analyze the ordinal nature of the data. Additionally, descriptive statistics such as mode or median for central tendency and frequencies for variability can be used to analyze Likert-type items, which are considered ordinal data (Subedi, 2016). To assess the association between independent and dependent variables, additional analyses such as Chi-2, Kendall Tau B, and Kendall Tau C are considered appropriate for Likert-type or ordinal data (Jakobsson, 2004). Standard texts also recommend using appropriate inferential statistics for ordinal data, such as Spearman's Rho or the Mann-Whitney U-test (Jamieson, 2020).

Secondly, to analyze Likert scale (interval) data, it is necessary to calculate a composite score (sum or mean) from the Likert items. If there are more than 4 survey items, researchers can use parametric statistics such as mean for central tendency and standard deviation for variance (Subedi, 2016). In addition, Pearson's *r* can be used to show the extent to which independent and dependent variables are associated with each other. Pearson's *r*-test, *t*-test, ANOVA, and regression procedures can be used depending on the nature and intensity of the study (Boone & Boone, 2012). ANOVA statistics can provide information on how the data is spread out and how it compares to the average. It also measures the degree to which levels or groups of an independent variable differ from each other.

Despite these facts, several researchers treat both Likert items and scale data as interval scales and analyze them as such with descriptive statistics like means and standard deviations, and inferential statistics like correlation coefficients, factor analysis, and analysis of variance (Brown, 2011; Zaman, Bibi & Raziq, 2020; Subedi,

2016). Both descriptive and inferential statistics can be used to analyze Likert data. Descriptive statistics are used to present the data in a simple numerical or visual form, while inferential statistics are used to test hypotheses such as correlations between different variables in the entire dataset. The important thing to note is that ordinal data can be transformed into numbers known as interval data (Carifio & Perla, 2008). Researchers can then calculate means, standard deviations, chi-square tests, Mann-Whitney U-tests, Pearson's r, ANOVA, t-tests, and regression if the data are normally distributed (Sullivan & Artino, 2013). Table 3 shows the statistical methods that can be used to calculate the Likert items and Likert scale datasets.

Table 3

Measurement and Analysis Type of Likert Data

Statistical methods to measure	Likert-type questions	Likert-scale questions
Central Tendency	Median/Mode	Mean
Variability	Frequency	Standard Deviation
Associations	Spearman's r, Kendall Tav B/C	Pearson's r
Other Statistics	Chi-square, Mann-Whitney U-test	ANOVA, t-test, and regression
Internal Consistency	Ordinal alpha	Cronbach's alpha

Results and Discussion

Likert items are generally used to measure people's attitudes, providing a range of responses to a given statement for collecting attitudinal data (Dittrich et al., 2007). Typically, there are 5 categories of response: strongly disagree (lowest score) to strongly agree (highest score) with a neutral option, although there are arguments in favor of scales with 7 or 9 odd Likert options (Jamieson, 2020). In a similar vein Bone & Bone (2012) stated five response alternatives such as strongly approve, approve, undecided, disapprove, and strongly disapprove in a series of original Likert scales. In addition, there are 4 categories of responses without a neutral option in even Likert options, and some scholars stated 6 to 8 response options, which are rare in practice. The more Likert options, whether they are odd or even, the more options there are to express the perceptions, feelings, satisfaction, and agreement of respondents. However, it is more likely to create confusion for respondents with too many options in a statement and might not be applicable to those who are less familiar with the subject matter. Additionally, the general public might not be able to select the option that they intended to select or respond to.

Furthermore, a large sample size is required to show the difference between and among variables, in which exact differences can be found. The descriptive statistics of Likert scale data, such as the mean and standard deviations, cannot provide clear meaning. For instance, considering Table 1, the average scores of never, rarely, sometimes, often, and always do not present exact clear meanings. So, the median or mode can be used to measure the central tendency for Likert scale data

(Jamieson, 2004). Similarly, frequency can be applied to measure variability, and the Chi-square test, Spearman r , and Mann-Whitney U test can be applied to show the association between variables instead of the parametric test, which requires interval data (Carifio & Perla, 2008). On the other hand, parametric tests can be used with ordinal data that are more robust than nonparametric tests. The parametric tests tend to provide the right answer even when data are not normally distributed, even to an extreme degree (Norman, 2010). [14]. So, parametric tests provide robust and unbiased findings that are sufficiently close to the truth when analyzing the Likert data.

Likert-type data are commonly known as discrete instead of continuous values (Winter & Dodou, 2023). These types of data are used to measure phenomena that can be measured through nominal or ordinal scales. During the analysis of Likert data, the researcher can use both parametric and non-parametric statistics. Formerly named parametric tests, the results obtained from the sample can be assumed throughout the total population in which data is normally distributed. In contrast, the researcher cannot generalize the findings obtained from the sample throughout the total population in non-parametric tests. So, nonparametric tests are less likely powerful compared to parametric tests. There exists disagreement among scholars about whether Likert data should be analyzed with parametric statistics such as the t -test or nonparametric statistics like the Mann-Whitney U test (Carifio & Perla, 2008; Jamieson, 2020). In this concern Winter and Dodou (2023) articulated that scholars Carifio and Perla reviewed Likert-items related articles and found 13% used a nonparametric test and 34% used a parametric test out of 95 in 1994. In a similar study Gregoire and Driver (1987) did not state a clear preference towards either the parametric t -test or nonparametric counterparts, however, Rasmussen (1989) stressed that parametric tests are more powerful for a lower Type II error rate, except when the sample pairs are taken from the most non-normal combination of distributions from a uniform distribution and a mixed-normal one. The author further pointed out no large differences between parametric and non-parametric tests relating to the occurrence of false positives, such as the Type I error rate. Authors Nanna and Sawilowsky (1998) found greater power for Mann-Whitney U (MWW) test in almost all investigated cases. The difference between the results of Rasmussen and those of Nanna and Sawilowsky is just the types of analysis distribution they applied; the former study investigated relatively normal distributions, while the latter used distributions of real data that were considerably skewed (Winter & Dodou, 2023).

Despite the Likert scale being invented back in 1932, it had not been practiced before 1970 [18]. Since the invention of the Likert scale, there have been debates among users about its reliability and validity in terms of the several points on the scale [3]. To respond to this issue, Likert published a similar study in 1934 with

similar results, though there were also design faults with the analysis (Edmondson, 2005). It is common practice to calculate the reliability index in Likert data frequently using Cronbach's alpha to measure the internal consistency of research tools. Cronbach's alpha measures how well a set of items measures a single, one-dimensional latent aspect of individuals (Subedi, 2016). Authors Garrido, Abad and Ponsoda (2013) also agreed on the great influence of Cronbach's alpha in calculating the reliability index in Likert data and further highlighted its misinterpretation and misuse. The Ordinal alpha is fairly equivalent to Cronbach's alpha; the significant difference between these two is that the former is based on the polychoric correlation matrix rather than the Pearson covariance matrix, which estimates alpha more accurately using ordinal data (Gadermann, Guhn & Zumbo, 2012). Hence, it can be recommended that, to ensure internal consistency of the used tools, ordinal alpha is appropriate, as it measures reliability more accurately than Cronbach's alpha for the ordinal nature of items in the survey questionnaire. Due to the scaling assumptions, ordinal and non-ordinal reliability coefficients may differ. The former (ordinal) coefficients focus on the reliability of the unobserved continuous variables underlying the observed item response, whereas the latter (non-ordinal) focuses on the reliability of the observed scores by treating the item responses as if they are continuous (Subedi, 2016).

Conclusion

The article aims to clarify Likert-type and Likert scale data for researchers and survey item developers, guiding the type of statistical methods to use for analyzing Likert items and scales, as well as their purpose. If the researcher or survey developer has serious or unique Likert items, then the items can be analyzed separately: median, mode, and frequency are considered appropriate statistical analyses. However, if survey developers have designed a series of survey items and require combined (score) measures in a particular theme, it can be analyzed as a Likertscale: other parametric tests such as mean, standard deviation, Pearson r, t-test, ANOVA, and regression can be used, but the data should be in interval scale. So, the researcher should consider how to analyze the data before its formulation.

A Likert item consists of ordinal type data, while a Likert scale consists of interval scale data. Likert-type data shows a relationship between responses; however, it cannot differentiate between the responses. For Likert-type data, mode or median for central tendency, frequency for variability, and Chi-2, Kendall Tau B, and Kendall Tau C for the association are appropriate. It is necessary to calculate a composite score from the Likert items to analyze the Likert scale data. The mean for central tendency and standard deviation for variance are appropriate for Likert scale data. Pearson's r can be used to show the extent to which independent and dependent variables are associated with each other. Pearson's r-test, t-test, ANOVA,

and regression procedures can be used depending on the nature and intensity of the study.

Ethics Statement

This research does not involve human participants and is prepared through reviewing published articles, requiring no ethical approval from the Institutional Review Board.

Conflict of Interest

The authors declare no potential conflict of interest concerning the research, authorship, and/or publication of this article.

Funding

The authors disclose no receipt of any funding or financial support for the research.

Consent to publication

Not applicable.

References

- Boone, H. N., & Boone, D. A. (2012). Analyzing Likert data. *The Journal of Extension*, 50(2), (2012).48. <https://doi.org/10.34068/joe.50.02.48>
- Brown, J. D. (2011). Likert items and scales of measurement 15 (2011) 20–24. https://hosted.jalt.org/test/bro_34.htm
- Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical education*, 42(12), (2008), 1150–1152. <https://doi.org/10.1111/j.1365-2923.2008.03172.x>
- Dittrich, R., Francis, B., Hatzinger, R., & Katzenbeisser, W. A. (2007). Paired comparison approach for the analysis of sets of Likert-scale responses. *Statistical Modelling*, 7(1) (2007),3-28. <https://doi.org/10.1177/1471082X0600700102>
- Edmondson, D. R. (2005). Likert scales: A history. *CHARM*, 12, (2005),127–133. https://orion2020.org/archivo/investigacion/Likert_History.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, and Evaluation* 17(1) (2012) <https://doi.org/10.7275/n560-j767>
- Garrido, L. E., Abad, F. J., & Ponsoda, V. (2013). A new look at Horn's parallel analysis with ordinal variables. *Psychological Methods*, 18(4), (2013). 454–474. <https://doi.org/10.1037/a0030005>

- Gregoire, T. G., & Driver, B. L. (1987). Analysis of ordinal data to detect population differences. *Psychological Bulletin*, 101(1), (1987), 159–165. <https://psycnet.apa.org/doi/10.1037/0033-2909.101.1.159>
- Jakobsson, U. (2004). Statistical presentation and analysis of ordinal data in nursing research. *Scandinavian journal of caring sciences*, 18(4), (2004), 437–440. <https://doi.org/10.1111/j.1471-6712.2004.00305.x>
- Jamieson, S. (2004). Likert scales: how to (ab)use them. *Medical Education*, 38(12), (2004). 1217– 1218. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Jamieson, S. (2020). Likert Scales : How to (ab)Use Them Title : Likert Scales : How To (Ab)Use Them ? Descriptive title : Commentary on the appropriate use of data derived from Likert-type rating scales Word count 749 words (excluding references & pull-out quotes). (2020). <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Joshi, A., & Pal, D. K. (2015). Likert Scale : Explored and Explained Likert Scale : Explored and Explained, (2015) <https://doi.org/10.9734/BJAST/2015/14975>
- Likert, R. (1932). A technique for measurement of attitudes. *Archives of Psychology*, 22(140), 5-55. <https://www.scirp.org/reference/referencespapers?referenceid=534541>
- Nanna, M. J., & Sawilowsky, S. S. (1998). Analysis of Likert scale data in disability and medical rehabilitation research. *Psychological Methods*, 3(1), (1998), 55-67. <https://psycnet.apa.org/doi/10.1037/1082-989X.3.1.55>
- Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education: Theory and Practice*, 15(5), 625-632. <https://doi.org/10.1007/s10459-010-9222-y>
- Rasmussen, J. L. (1989). Analysis of Likert-scale data: A reinterpretation of gregoire and driver. *Psychological Bulletin*, 105, 67–170. <https://www.semanticscholar.org/paper/Analysis-of-Likert-scale-data-revisited.-Gregoire/60083024c9a908e34d2f2511240e2c89b923b08a>
- Subedi, B. P. (2016). Using Likert Type Data in Social Science Research : Confusion, Issues and Challenges. 3(2), 36–49. <https://ijcar.net/assets/pdf/Vol3-No2-February2016/02.pdf>
- Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, 5(4), 541–542. <https://doi.org/10.4300/JGME-5-4-18>
- Winter, J. D., & Dodou, D. (2023). Five-Point Likert Items : t test versus Mann-Whitney-Wilcoxon. ISSN 1531-7714 Volume 15(11).

[https://www.researchgate.net/publication/266212127_Five-Point Likert Items t Test Versus Mann-Whitney-Wilcoxon](https://www.researchgate.net/publication/266212127_Five-Point_Likert_Items_t_Test_Versus_Mann-Whitney-Wilcoxon)

Zaman, N. U., Bibi, Z., & Raziq, A. (2020). Manualizing Factor Analysis of Likert Scale. *Journal of Management Sciences*, 7(2), 56–67. <http://dx.doi.org/10.20547/jms.2014.2007204>