

Likert Scale: Explored and Explained**ABSTRACT:**

Likert scale is applied as one of the most fundamental and frequent psychometric tools in educational and social sciences research. Simultaneously it is also subjected to a lot of debates and controversies in regards with the analysis and inclusion of points on the scale. With this context; through reviewing the available literature and then clubbing the received information with coherent scientific thinking; this paper attempts to gradually build a construct around likert scale. This analytical review begins with the necessity of psychometric tools like likert scale and its' variants and the focuses on some convoluted issues like validity, reliability and analysis of the scale.

KEY-WORDS: psychometrics, likert scale, points on scale, analysis, education.

INTRODUCTION:

Nothing is more frightening than a fear you cannot name.

— Cornelia Funke, Inkheart

Since the inception of human race; we have been mesmerized to capture the intangible and concealed attributes of human behaviour and performance; although it is always a challenge to quantify a thing which cannot be measured through conventional measurement techniques. The need of quantification lies in the necessity to transform an individuals' subjectivity into an objective reality. Attitude, perceptions and opinions fit into above statements which capture the degree of accomplishment of any interventions from stakeholders/participants perspective. Qualitative research techniques do capture the complexity of human thoughts, feelings and outlooks through several social science techniques; still the quantification of these traits remains a predicament.

PSYCHOMETRICS AND LIKERT SCALE

Psychometrics techniques are being developed, instituted and refined in order to meet the quantification of traits like ability, perceptions, qualities and outlooks...a requirement of social sciences and educational researches [1, 2].Psychometrics operates through formulating of approaches (theoretical construct) for measurements followed by development of measuring instruments and their validation. Stanford Binet test (measures human intelligence), Rasch measurement model (use for

estimation of ability), Minnesota Multiphasic Personality Inventory (measures human personality), Likert scale (measures human attitude) are some of the instruments in Psychometrics used widely in the social science & educational research [3, 4, 5].

Likert scale was devised in order to measure 'attitude' in a scientifically accepted (validated) manner in 1932 [6, 7]. An attitude is relatively enduring organisation of belief and ideas around an object, a subject or a concept acquired through social interactions which directs the individual to behave & act in a preferential manner on confrontation [8]. This is clear from the discourse that thinking (cognition), feeling (affective) and action (psychomotor) all together in various combination/permutation constitute delivery of attitude in a specified condition. The issue is how to quantify these subjective preferential thinking, feeling and action in a validated and reliable manner a help is offered by Likert scale [9, 10].

The original Likert scale is a set of statements for the situations (conditions- real or hypothetical) under study and demands participants' to show their level of agreement (from strongly disagree to strongly agree) with a given statement (preferences) on a certain metric scale. Further, the composite score of all the statements is supposed to measure a single dimension (single trait) of attitude for that person and comparisons with score of other persons [11].

LIKERT SCALE AND ITS VARIATION

This position of neutrality exactly in between two extremes of strongly disagree (SD) to strongly agree (SA) provides independence to choose any response in a balanced and symmetric way in either directions (Symmetric scale); as opposed to a asymmetric Likert scale in which less choices are given on one side of neutrality (average) as compared to other side. Asymmetric scale in some case also indicates ipsative (forced) choices where there is no perceived value of indifference/neutrality of researcher [12, 13,14].

7 or 10 point scale is a variation of 5 point scale in which more continuums of choices are granted as an attempt to compare the trait in a more intense manner [15]. These variations are discussed in more details (in reference with validity and reliability) further in this paper.

Another issue of concern at this juncture is to understand the difference between the Likert and Likert type scale. When scores of the all items of the questionnaire are combined (sum) to generate a composite score, which in turn measures a uni-dimensional trait (as also described in beginning); this instrument is known as Likert scale. Sometimes the primary interest of the researcher is not to synthesize the stance of the participants per se but to grab hold on several inter-related feelings, actions and pragmatic opinion of the participants around an event/phenomenon. This demands the individual analysis of item in terms of degree of agreement collectively for participants. This 'direction of enquiry' must be decided during the designing the questionnaire and not at the time of analysis. The scale used so can be labelled as Likert type and not Likert scale [16].

64

65

66 **IS 7 POINT LIKERT SCALE BETTER THAN 5 POINT LIKERT SCALE? - A**

67 **PERSPECTIVE CONTROVERSY OR ESTABLISHED WITH A CONSENSUS?**

68 Since the advent of Likert scale in 1932, there have been debates among the users about its best possible usability in term of reliability and validity of number of points on the scale [17-20].

70 Considering reliability of the responses from participants in a survey, chances are that the 7 point scale may perform better compared to 5point scale. The 7 point scale provides more verities of options increasing the probability of meeting the objective reality of people. As a 7-point scale reveals more description about the motif and thus appeals practically to the "faculty of reason" of the participants [19, 20].

75 A respondents' absolute agreement with the motif of topic may lie between the two descriptive options provided on a 5 point scale. On repeated administration he/she may differ in choosing one of the options e.g. 3 instead of 4 when the person thinks in between the two of the response option on 5 point scale. A 7 point scale may eliminate this problem up to an extent, by eliciting retrieval beyond the utmost level of agreement provided by a 5 point scale. Hence, the dilemma of choosing between the two undesirable points on 5 point Likert scale may convert into the slight subjective predicament towards a response option or even an absolute agreement as a result of provision of degree of measurement beyond the extent of 5 point scale [21-23]. The provision of number of scale points, 5 point or 7 point, would be more engaging to the minds of respondents when the items on the scale

carry the statement of ideas near the truth of the universe for both the participants and the surveyor. It may create the 'curves of reliability' around the 'zenith of validity'. The dilemma of choice and explicit greater extent of measurement by 7 point scale is very much in the territory of the reason of response without which consideration of reliability is of no weight [19].

Validity of Likert scale is driven by the applicability of the topic concerned, in context of respondents' understanding and judged by creator of the response item. We can appreciate it by an example: "How efficacious is a therapeutic modality in treating a particular disease?" This question when asked to a group of individuals, indifferent with the disease or the modality, the response pattern may remain similar, independent of the number of point on the scale. The responses may cluster around center or to the extreme ends. On the contrary, when the topic concerned is relevant to the respondents' context provision of more option may add to the content & construct validity of the scale. Providing options more close to the original view of the respondent reduce the role of ambiguity in the responses.^[22, 12] Furthermore, comprehension of all items and points on a scale needs a judgment time and a memory span different for different means and also depends on communication mode. While listening to the responses of a long scale may discern the various options on the scale with lesser time to judge compared to a written scale. Written scale thus will add to validity even with more points on the Likert scale. Also research concerning span of immediate memory support this notion of accuracy of response categories around seven as human mind has span of absolute judgment that can distinguish 7 categories at a time [24].

ANALYSIS OF THE RESPONSE ITEM

Before proceeding to real analysis, a very fundamental question, rather controversy, needs to be addressed - which type of scale Likert is?

There are two schools of thoughts embedded on equivalence and equidistance. Researchers and statistician often equate Liket scale with ordinal scale in which choices or responses are arranged in some instrumental order. However, this scale doesn't show the relative magnitude between two responses. On the other hand another interpretation offered for this dilemma is that when the aim of the researcher is to 'combine' all the items in order to generate a 'composite' score for an individual

then these individualistic summative score do reflect a sensible realistic distance among individuals hence can be labelled as 'interval estimates' [25, 16].

To understand this concept; let's assume a scenario in which the aim of the researcher is to measure the attitude towards classroom lectures and to make out relative preferences (library reading and small group teaching) compared with lecture. He design the following survey instrument on 5 point likert scale for the above said purpose-

Items	SD	D	N	A	SA
1. Lectures are usually dominated with monologue without revealing the application of what is being taught.	1	2	3	4	5
2. The information provided in lecture is very much textual and can be retrieve easily though text book.	1	2	3	4	5
3. Lectures usually fail to generate curiosity and further motivation to know more about the topic	1	2	3	4	5
4. I would like to spend time in self directed reading owing to the opportunity of control on pace and choices of learning	1	2	3	4	5
5. I would like to attend Small Group Teaching (SGT) session owing to the opportunity of more interaction and liveliness of the environment.	1	2	3	4	5

Fig.1. Survey Instrument for Measuring Attitude towards Classroom Lectures

The first question of importance is-'Can the items be clubbed in order to generate a composite index to measure the attitude?' As item-1, 2 and 3 are constructed in (a) the logical sequence (b) closely interrelated but provides some independent information as well (c) coherence /expectedness in the next responses based upon previous one and (d) ability of the each item to measure a distinct element of the issue; so they may collectively construct an index to measure the stance of a student towards lecture.

132 On the other hand Item-4 and item-5 offers separate and sovereign preferences regarding self
133 directed reading and small group teaching they should be analyzed independently from item 1,2 and 3
134 and even from each other.

135 The next question arises- On what scale can item1, 2 and 3 can be treated and what is the
136 appropriate measurement scale for item4 and item-5?

137 The answer for the above lies in another question asked by Stevens in his famous paper-‘what are
138 the rules (if any) under which numerals are assigned?’ Here we see (a) the minimum score one can
139 secure for three item is 3 (and not a absolute zero)-as in psychometrics attitude is measured in
140 positive degree and being the ‘strongly disagree ‘cannot be equated with ‘absolute
141 disagreement’....there is always something below than strongly disagree; and zero also gives the
142 notion of neutrality rather disagreement (the attitude is zero...means one is apathic to issue) (b) each
143 numerals conveys the same meaning in all three items (i.e. 3 denotes the neutral in all three items)
144 (c) as mentioned above, all three items can be clubbed while satisfying the content and criterion
145 validity (d) considering a, b and c; one can safely assume the equality of interval among response
146 variable and thus the obtained data can be treated on a interval scale.

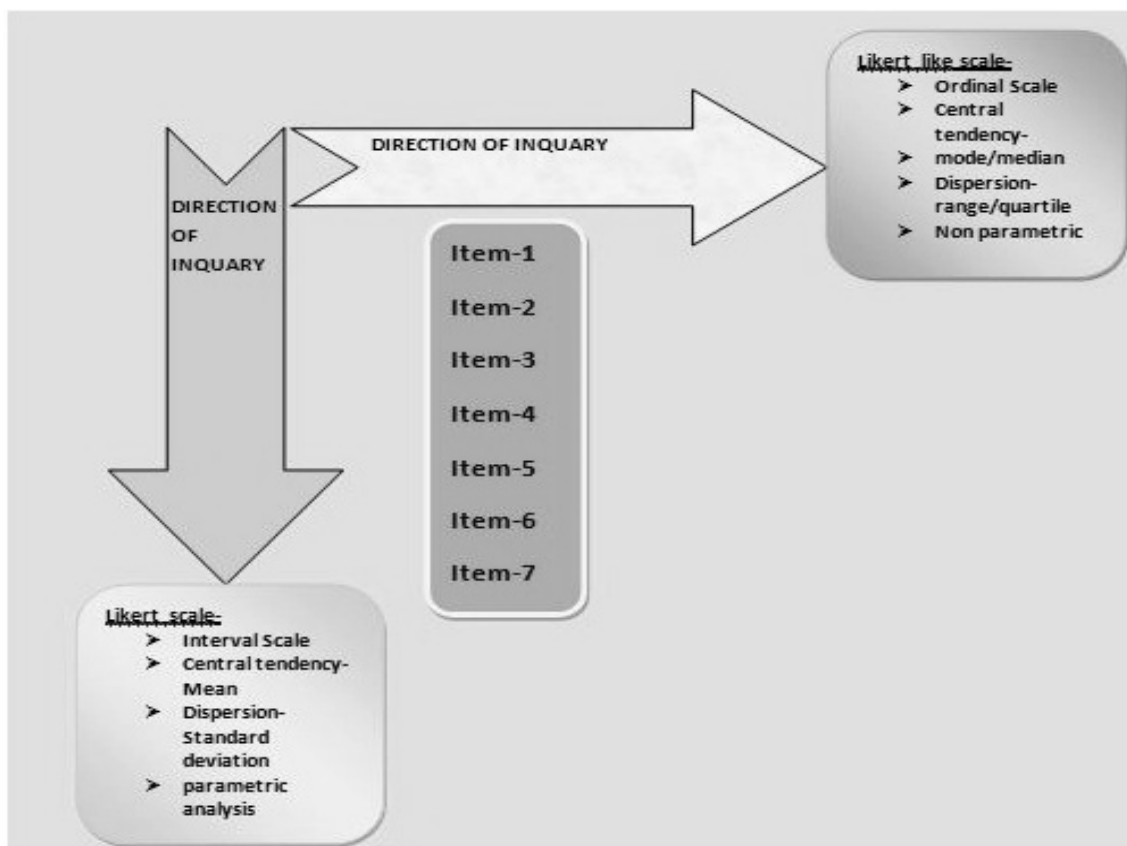
147

148 Conversely, Item -4 and 5 being an autonomous observation from each other and item1, 2 and 3;
149 should be treated differently. As it is clear from response variables /set, if one applies any order
150 preserving transformation (like square, multiplication, square root etc.) to the response variable; the
151 rank order will be unaffected so a ordinal scale’s assumptions and treatment is applicable on this
152 subset.

153 The crux that can be extracted from the above discussion is that the methods adopted for analysis
154 depends on the aim of researcher decided beforehand during the construction of the instrument.

155 If one wished to combine the items in order to generate a composite score (Likert scale) then the
156 scale will be an interval scale (fig.2). The measure for central tendency and dispersion for this type of
157 scale will be mean and standard deviation. Further pearsons’ correlation coefficient (r), ANOVA &
158 Regression analysis can be performed for such data set.

159 As opposed to, if researcher wishes to analyse separate item (without combining them for an
160 individual-Likert type scale) then this pattern of analysis is for ‘ordinal data’ (fig.2). Needless to say;



the recommended measure of central tendency and dispersion are the median (or the mode) & frequency (or range) for such pattern. In addition to this, this pattern may be further statistically tested through Non-parametric techniques such as chi-square test, Kendall tau B or C test.

Fig.2. Choice of Analysis of Likert Items: Aim and Construct of Research

REFERENCES:

1. Cook DA, Beckman TJ. Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine*. 2006; 119(2): doi:10.1016/j.amjmed.2005.10.036
2. Shea JA, Fortna GS. Psychometric Methods. In: Norman GR, Van der Vleuten CPM, Newble DI, editors. *International Handbook of Research in Medical Education*. Dordrecht: Kluwer Academic Publisher; 2002. p. 97-126.
3. Youngstrom EA, Glutiing JJ, Watkins MW. Stanford-Binet Intelligence Scale: Fourth Edition (SB4): Evaluating the Empirical Bases for Interpretations. In: Reynolds CR, Kamphaus R W,

- 175 editors. *Handbook of psychological and educational assessment: Intelligence, aptitude, and*
176 *achievement*. New York: Guilford press; 2003.p. 217-242.
- 177 4. Wu M, Adams R. *Applying the Rasch model to psycho-social measurement: A practical*
178 *approach*. Melbourne: Educational Measurement Solutions; 2007.
- 179 5. Croasmun JT, Ostrom L. Using Likert-Type Scales in the Social Sciences. *Journal of Adult*
180 *Education*.2011; 40: 19-22.
- 181 6. Edmondson DR. Likert Scales: A History. Proceedings of the 12th Conference on Historical
182 *Analysis and Research in Marketing (CHARM)*.April 28-May1, California, USA; 2005.
- 183 7. McLeod S. Likert Scale. *Simplypsychology.org*. Retrieved January 27, 2014, from
184 <http://www.simplypsychology.org/likert-scale.html/pdf>.
- 185 8. Park K. Social Sciences and Health. In Park K, editor. Park Text Book of Preventive and
186 Social Medicine. Jabalpur: Bhanot Publishers; 2013.p. 54x-54x.
- 187 9. Culbertson JM. What is an attitude? Journal of cooperative extension. 1963: 79-84.
- 188 10. Schwarz N, Bohner G. The Construction of Attitudes. In: Tesser A & Schwarz N, Editors.
189 *Blackwell Handbook of Social Psychology*. Oxford: Blackwell; 2001. p. 436-457.
- 190 11. Singh YK. *Fundamental of Research Methodology and Statistics*. New Delhi: Newage
191 International (P) Ltd. Publisher; 2006.
- 192 12. Lobsy J, Wetmore A. CDC Coffee Break: Using Likert Scales in Evaluation Survey Work
193 [Internet].Cdc.gov. Retrieved January 12 , 2014, from
194 www.cdc.gov/dhds/pubs/docs/CB_February_14_2012.pdf.
- 195 13. Tsang KK. The use of midpoint on Likert Scale: The implications for educational research.
196 *Hong Kong Teachers Centre Journal*. 2012; 11:121-130.
- 197 14. Malhotra NK. Questionnaire Design and Scale Development. In: Grover R, Vriens M, editors.
198 *The Handbook of Marketing Research*. California: Sage Publications, Inc; 2006.
- 199 15. Dawes J. Do data characteristics change according to the number of scale points used? An
200 experiment using 5-point, 7-point and10-point scales. *International Journal of Market*
201 *Research*. 2008; 50(1):61-77.
- 202 16. Boone HN, Boone DA. Analyzing Likert Data. *Journal of Extension*. 2012; 50(2). Available
203 from <http://www.joe.org/joe/2012april/tt2.php>.

- 204 17. Colman AM, Norris CE, Preston CC. Comparing rating scales of different lengths:
205 Equivalence of scores from 5-point and 7-point scales. *Psychological Reports*. 1997; 80: 355-
206 362.
- 207 18. Preston CC, Colman AM. Optimal number of response categories in rating scales: reliability,
208 validity, discriminating power, and respondent Preferences. *Acta Psychologica*. 2000. 104; 1-
209 15.
- 210 19. Chang L. A psychometric evaluation of 4-point and 6-point likert-type scale in relation to
211 reliability and validity. *Applied Psychological Measurement*. 1994; 18:205-215.
- 212 20. Cox EP. The optimal number of response alternate for a scale: a review. *Journal of Marketing*
213 *Research*. 1980; 17(4):407-422.
- 214 21. Pearse N. Deciding on the scale granularity of response categories of likert type scales: The
215 Case of a 21-Point Scale. *Electronic Journal of Business Research Methods*. 2011;
216 9(2):2011.
- 217 22. Finstad K. Response interpolation and scale sensitivity: evidence against 5-point scales.
218 *Journal of Usability Studies*. 2010; 5(3):104-110.
- 219 23. Komorita SS and Graham WK. Number of Scale Points and the Reliability of Scales.
220 *Educational and Psychological Measurement*. 1965; 25(4):987-995.
- 221 24. Miller GA. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for
222 Processing Information. *The Psychological Review*. 1956. 63:81-97.
- 223 25. Carifio J and Perla RJ. Ten common misunderstandings, misconceptions, persistent myths
224 and urban legends about likert scales and likert response formats and their antidotes. *Journal*
225 *of Social Sciences*. 2007; 3 (3):106-116.

226