https://doi.org/10.48009/2_iis_2005_96-102

RANKING ORDINAL SCALES USING THE CONSENSUS MEASURE

William J. Tastle, Ithaca College, tastle@ithaca.edu
Mark J. Wierman, Creighton University, mwierman@creighton.edu
U. Rex Dumdum, Marywood University, rexdum2@marywood.edu

ABSTRACT

A new measure is applied to the analysis of ordinal scales, specifically the Likert scale. Using this measure, sets of Likert scales can be hierarchically compared, and the measure is mathematically simple to apply. A set of rules for consensus is identified and an example of its use as a measure of the strength of consensus is illustrated. Using an existing Likert scale of responses, the strength of consensus measure is applied and the results are shown to provide greater insights into the importance of the various rows of data.

Keywords: consensus measure, strength of consensus, Likert scale, decision making

INTRODUCTION

It is typical for information systems professionals to solicit project-specific information from management as well as teams of individuals directly involved with the system in question when addressing the requirements determination phase of systems analysis. Distinguishing the cause of the perceived problems under investigation is a problem that continues to provide incentive for research, especially as to a method by which symptoms can be separated from cause. Each stakeholder views their understanding of the system, and its corresponding problem, as true, even though their perceptions might differ from others. Further, author experience continues to show that it is not atypical for the view held by the highest-ranking individual(s) around the conference table to override the views of subordinates.

While there is much in the literature concerning motivations and methods for motivating team involvement with respect to establishing an agreement on business problem identification, there is little with respect to mathematical measures that can be effectively used to guide a team. We introduce a measure, based on the Likert scale, which permits an almost continuing monitoring of group discussion as a search for problem identification consensus ensues. Further, the measure is extended to show that it can be a powerful tool in bringing hierarchal meaning to Likert scale propositions.

MEASURING GROUP CONSENSUS

The use of teams in decision-making has dramatically increased in organizations [2, 3]. Teams, when appropriately deployed and properly enabled, tend to execute better, learn faster, and deliver better results [1, 2, 3].

Agreement cannot be presumed in the analysis of complex, ill-structured situations, so tools are necessary for identifying agreement during the divergent and convergent processes in problem analysis. The divergent process includes the identification, investigation, and evaluation of issues

while the convergent process includes consensus formation regarding issues that need to be addressed, their priority level and the most efficacious means to address them.

This paper presents a mathematical measure of agreement as an indication of agreement (consensus) and dissension. Although consensus building is a typical method used in team decision-making, few measures exist which allow for the easy determination of the degree to which a team is nearing the point of agreement. The measure presented in this paper is intuitive and easy to apply on ordinal scales.

AGREEMENT AND DISSENSION

We consider agreement and dissension to be diametrical concepts. An agreement is a concurrence of opinion, a compatibility of observations reached by a team of individuals acting as a whole; it may also considered consensus. Dissension is defined as a difference of opinion such that strife is caused within the team undertaking to make a decision. We define consensus as complementary to dissension. However, the purpose of this section centers on the understanding and measurement of the concept of agreement/consensus. We consider either term to capture the intent of team acceptance.

Consensus is a function of shared team feelings towards an issue. This "feeling" can be captured through an ordinal scale, specifically the Likert scale, which measures the extent to which a person agrees or disagrees with the question. For example,

> Question: "The order entry process needs to be fixed."
>
> 1 Strongly disagree
> 2 Somewhat disagree
> 3 Undecided
> 4 Somewhat agree
> 5 Strongly agree

Other number-assignments can be made, such as: -2 = strongly disagree, -1 = disagree, 0 = not sure, 1 = agree, and 2 = strongly agree, or 0.0 = strongly disagree, 0.25 = disagree, 0.50 = neutral, etc. We have shown elsewhere [5] all these scales produce identical values for the Consensus measure. Likert scales can also be from two to nine categories in width. The issues of scale, symmetry, selection of clusters, and ordinal vs. interval data are not addressed here, but Munshi [4] has produced a very nice paper that describes these aspects in straightforward terms. The paper by Munshi [4] also contains an excellent bibliography.

RULES FOR A MEASURE OF CONSENSUS

We establish a set of rules that must be satisfied before any measure can be considered a viable solution to the Likert scale agreement/consensus problem.

1. For a given (even) number of individuals participating in a discussion on some question of interest, if an equal number of individuals, n/2, separate themselves into two disjoint teams, each centered on the strongly disagree and strongly agree categories, the team is considered to have no agreement/consensus.

2. If all the participants classify themselves in the same category of the Likert scale, regardless of the category, then the agreement/consensus of the team is considered to be complete at 100%.

3. If the mix of participants is such that n/2 + 1 participants assigns themselves to any one category, the degree of agreement/consensus must be greater than 0, for the balance in the team is no longer equal.

4. As the number of categories to which each participant classifies himself/herself diminishes, the agreement/consensus must increase, eventually approaching 1 on the unit interval. Thus, when all participants place themselves in a single category, consensus has been maximized and it considered to be perfect, and that is given a value of 1.

5. The dispersion of the categorical values must be captured by the consensus to provide an indication of the variance of the data.

Hence, a complete lack of consensus must generate a value of 0, and a complete consensus of opinion must yield a value of 1. Every other combination of Likert scale categories must result in a value within the unit interval. The issue of classification of individuals into Likert categories and the makeup of questions such that a Likert scale can be properly applied, is not addressed here.

A MEASURE OF CONSENSUS

The Consensus Measure (Cns) is defined as:

$$Cns(X) = 1 + \sum_{i=1}^{n} p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X}\right) \qquad [1]$$

where X is any finite discrete random variable with probability distribution $p(x)$. As usual $E(X) = \sum_{i=1}^{n} p_i X_i = \mu_X$ is the mean of X and d_X is the width of X, $d_X = X_{\max} - X_{\min}$. μ_X is the mean, or it may also be an arbitrary value other than the mean if the measure is used as a strength of consensus. This measure adequately fulfills the above rules as evidenced by the following illustrations [5].

For the situation represented by table 1, the Cns for each of the first four rows is zero. Thus, the number of individuals participating in the team does not have any impact on the value of consensus and rule #1 is satisfied. Note that the last row of table 1 illustrates a very modest shift of one person from Strongly Agree to Strongly Disagree. That change causes the balance to shift slightly towards the SD side of the Likert Scale, the result being a very slight increase in the degree of consensus.

Table 1. Data from Table 1 with the Consensus Measure Associated with each Row

SD	D	N	A	SA	Cns
5	0	0	0	5	0
50	0	0	0	50	0
500	0	0	0	500	0
5000	0	0	0	5000	0
51	0	0	0	49	0.0003

Rule 4 states that as the number of participants in the team shifts their judgment such that the categories begin to gravitate towards a central value, the degree of consensus must also correspondingly increase to reflect agreement. Hence, the degree of proximity increases as the numbers of individuals in the team adjust their perceptions about the question or issue under discussion and move towards agreement. Table 2 shows a movement in proximity from complete opposition to complete agreement. Note that the final row represents the total of all team members.

Table 2. The Movement of Values Towards Coalescence on a Single Category.
Note that the Consensus Ranges from 0 to 1.

SD	D	N	A	SA	Cns
5	0	0	0	5	0
5	0	0	5	0	.3219
5	0	5	0	0	.5850
5	5	0	0	0	.8074
10	0	0	0	0	1

Finally, as the number of participants increases in size, the consensus measure should not be affected (Rule #4). Regardless of the number of participants, the proportion of the team in each category is constant and hence, the measure of consensus remains unchanged (see Table 3).

Table 3. As the Number of Members in the Team Increases, the Consensus Remains Constant as Long as the Category Percentages Remain Constant.

SD	D	N	A	SA	Cns
0	1	0	3	0	.6860
0	10	0	30	0	.6860
0	20	0	60	0	.6860
0	30	0	90	0	.6860
0	300	0	900	0	.6860

USING THE CONSENSUS MEASURE TO DETERMINE HIERARCHY

The consensus measure yields a number in the unit interval regardless of the categories involved. Hence, Table 4 shows three representations, all of which have the same Cns value.

Table 4. Three Representations of Consensus Values

SA	A	N	D	SD	Cns
5	4	3	0	0	0.7233
0	5	4	3	0	0.7233
0	0	5	4	3	0.7233

If we have phrased the question such that the desired response is SA, then the current measure does not capture the strength of association with the "strongly agree" category. Hence, we force the measure to reflect the SA category (or any other category we choose) to be the central value, and calculate the Cns as the *strength of affirmative consensus* (**sCns**), or *strength of accord*, or *consensus strength*. The proper nomenclature on this aspect of the measure is currently under internal debate.

Forcing the mean to equal 1, that is, placing the focus of the measure on the SA category and comparing the distribution to be evaluated with respect to SA, generates the values in Table 5.

Table 5. Consensus Values after Forcing the Mean to 1

SA	A	N	D	SD	sCns
5	4	3	0	0	0.723
0	5	4	3	0	0.327
0	0	5	4	3	0.334

The further the individual values migrates from the forced central value (SA in this case), the less the consensus strength. At present, the endpoints have some interesting properties that cause the sCns to be non-monotonic (note the increase in value from 0.327 to 0.334), but these properties are currently being investigated. There are at least two novel solutions to this peripheral problem.

This method permits the comparison of Likert scales based on the proximity of the values to a specific value. Certainly, the measure can be applied to virtually any situation in which the Likert scale is used to collect data, and this is especially relevant in the IS domain where agreement on complex issues can be a challenge. To illustrate the effectiveness of this strength of consensus measure, we use a survey published on the Internet. Though the example does not deal with IS topics, the power of the measure can be ascertained.

STRENGTH OF CONSENSUS EXAMPLE

Recently a survey based on the Likert scale was published [6] by the Episcopal Church of New Hampshire that addressed a number of important issues. The Church presented both the

questions and Likert scale values (as percentages) for view by the public. Applying the strength of consensus measure to these numbers by placing the weight on the "strongly agree" category yielded some interesting results.

The questions are contained within Table 6 below (a subset of the questions is shown to satisfy publishing requirements) followed by the responses by surveyed group, i.e., the ordained

Table 6. Strength and Ranking of Consensus Responses

	sCns	Imp		SA	A	N	D	SD
1			It is important that the next Bishop of New Hampshire help us define and implement our vision for ministry.					
	0.7490	**1**	**Ordained**	51	43	3	2	0
	0.7273	**3**	**Lay**	42	53	3	1	0
	0.7382	**2**	**Total**	93	96	6	3	0
2			It is important that the next Bishop of New Hampshire make ministry with youth and young adults a priority in terms of time and energy.					
	0.5259	**24**	Ordained	29	54	9	8	0
	0.6891	**5**	Lay	45	46	6	3	0
	0.6075	**15**	**Total**	74	100	15	11	0
3			It is important that the next Bishop of New Hampshire continue to recruit and deploy women clergy.					
	0.6031	**16**	Male	36	46	11	5	3
	0.6842	**6**	Female	44	44	7	3	1
	0.6800	**7**	Ordained	40	45	11	1	2
	0.6532	**10**	Lay	41	45	8	4	2
	0.6666	**9**	**Total**	81	90	19	5	4
4			It is important that the next Bishop of New Hampshire emphasize the ministry of the laity.					
	0.7245	**4**	Ordined	41	55	5	0	0
	0.6259	**14**	Lay	33	54	11	2	0
	0.6754	**8**	**Total**	74	109	16	2	0
5			It is important that the next Bishop of New Hampshire maintain the financial integrity and health of the diocese through stewardship education.					
	0.6451	**11**	Ordained	32	59	7	2	0
	0.6327	**13**	Lay	29	62	7	2	0
	0.6389	**12**	**Total**	61	121	14	4	0
	sCns	**Imp**		**SA**	A	**N**	**D**	**SD**

clergy, lay people, rural, suburban, male, female, etc. The first column shows the calculated strength of consensus, and the second column the ranking by decreasing sCns order. Sometimes, the difference is insignificant, as in the first question concerning help in defining and implementing their "vision for ministry." Both the ordained and lay classified it is similarly important, which the "total" row expectedly captures.

However, on the issue of the importance of making a ministry with youth and young adults (item 2), the matter ranked very high with the lay people [5] but considerably lower for the clergy [24]. Conversely, item 4 is very important to the clergy [4], but of substantially less importance to the laity [14]. Thus, a set of Likert scales can be confidently used to logically order outcomes. The application of this method to business in general, and team/group decision-making, is obvious.

Utilizing this method, we undertook an analysis of the faculty in one AASCB school of business. The results are both interesting and compelling, and are released in another paper.

CONCLUSIONS AND FURTHER WORK

A new measure of consensus is introduced as a means by which ordinal scale measures can be compared. The measure used is generally the Likert scale, and given a set of questions and responses, an ordering is possible. This method provides a means by which group and team discussions can be objectively managed and, when it is known that the team is tending towards agreement, the facilitator can better manage the desired outcome. The measure as only a consensus of ordinal scale values works very well, but some unexpected deviations at the extremes require additional research.

REFERENCES

1. Avolio, B.J., Kahai, S., Dumdum, U.R. & Sivasubramaniam, N. (2001). Virtual teams: Implications for e-Leadership and team development. In M. London (Ed.) *How People Evaluate Others in Organizations*. NJ: Lawrence Erlbaum Associates.
2. Katzenbach, J.R. & Smith, D.K. (2003). *The Wisdom of Teams: Creating the High Performance Organization.* HarperBusiness Essentials.
3. Mohrman, S.A., Mohrman, S.G. & Mohrman, A.M. (1995). *Designing Team-Based Organizations: New Forms of Knowledge Work.* Jossey-Bass.
4. Munshi, J. (1990). A Method for Constructing Likert Scales. http://www.munshi.4t.com, downloaded June 2004.
5. Wierman, M. J. & Tastle, W. J. (2005). *Consensus and Dissension: Theory and Properties*. North American Fuzzy Information Processing Systems Conference, Ann Arbor, MI.
6. Diocese of New Hampshire, Survey Report, downloaded 3/23/2005, www.nhepiscopal.org/BishopSearch/leader_questions.htm.