IGI PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.igi-global.com

ITJ4382

This paper appears in the publication, International Journal of E-Business Research, Volume 4, Issue 4
edited by **In Lee © 2008, IGI Global**

E-Business Decision Making by Agreement

William J. Tastle, Ithaca College, USA

Mark J. Wierman, Creighton University, USA

ABSTRACT

Gathering customer data over the Internet is largely limited to collecting the responses to a set of easily answerable questions, such as Yes/No questions and Likert scale questions. These data are then analyzed to identify customer trends or other items of interest to management. The data can be useful, but key to their usage is the application of suitable mathematical tools. Traditionally little more than standard statistics has been used in the analysis of ordinal, or category, data. This can be inaccurate and in some cases, misleading. This paper introduces measures of agreement and dissent to the field of eBusiness analysis and shows how ordinal data can be analyzed in more meaningful ways.

Keywords: *agreement; consensus; dissent; Likert scale; ordinal data*

INTRODUCTION

Gathering data from customers is a common activity and much research has gone into design and planning (Parsons, 2007; Solomon, 2001), improving response rates (Cook, Heath, & Thompson, 2000; Kaplowitz, Hadlock, & Levine, 2004; Schmidt, Calantone, Griffin, & Montoya-Weiss, 2005), the study of privacy and ethics (Couper, 2000), mode of questionnaire delivery (Denscombe, 2006), the effect of subject lines of survey responses (Porter & Whitcomb, 2005), and the analysis of Web usage using traditional statistics (Korgaonkar & Wolin, 1999; Stanton, 1998), but little has been written about the evolution of ordinal scale survey results, typical of Likert or Likert-like scale surveys. Acknowledging that getting respondents to answer surveys, either paper or digital, can be a challenge, and once the data is collected the effort to squeeze as much information from the data as possible begins.

Traditionally, data analysis is well founded in statistics, even though the same underpinnings of statistics recognize that there are limits to this branch of mathematics. Statistics are at home when dealing with ratio or interval data

(Tastle & Wierman, 2006a), but once the scale shifts to ordered categories the use of statistics is circumspect, for what does it mean to say the average of "warm" and "hot" is reported as "warm-and-a-half" (Jamieson, 2004). Ordinal scales of measurement typically consist of ordered category hierarchies such as: strongly agree (SA), agree (A), neither agree nor disagree (N), disagree (D), and strongly disagree (SD); very cold, cold, cool, tepid, warm, hot, and very hot. The instrument typically used to collect this kind of data is called the Likert scale, though there are variations of this scale such as Likert-like, Likert-type, and ordered response scales. Researchers utilize this kind of instrument to collect data that cannot be ascertained using traditional measures, for the data being collected are feelings, perceptions, sensations, emotions, impressions, sentiments, opinions, passions, or the like. Unfortunately, the application of standard statistics to these data can be improper (Cohen, Manion, & Morrison, 2000; Jamieson, 2004; Pell, 2005). This article looks at the different kinds of scales and presents a new measure for analyzing ordinal scale data.

The identification of consensus in a group environment was the motivation for the original research into ways of assessing ordinal data. The authors sought to identify some mathematical way by which a discussion leader could be guided towards getting a group of discussants to arrive at consensus as quickly as possible. The consensus measure can be easily applied to situations whereby a quick survey of perceptions of discussants to one statement is taken. Given the statement "The group has arrived at consensus" the discussants would check either SA, A, N, D, or SD. The resulting calculation of consensus could guide the leader in the direction of conversation or to determine if there is sufficient agreement to move forward. The authors have expanded on this idea to identify the group agreement with a targeted category, such as SA, on a data collection instrument. It would be nice to know if, in response to some survey statement on a matter of critical importance to the organization, the overall percentage of agreement for each Likert category, not just the mode category. Notice we do not use the mean, for the meaning of the average of two ordered categories is not clear, that is, the average of *acceptable* and *unacceptable* is *acceptable-and-a-half*, or so the interval and ration scale mathematics tells us. Also, standard deviation is based on the presence of at least an interval scale, so its use on ordinal scales is suspect at least, and invalid at most. The dissent measure gives a result that is much easier to interpret and carries more intuitive meaning. In this article we focus on the agreement measure and how it can be used to foster a group agreement assessment that is especially important when a business is largely limited to Internet activities and must rely on survey-type data for assessments that might typically be ascertained through an in-person sales force.

BACKGROUND

We begin with a discussion of the meaning of *consensus*, for it plays a critical role in the analysis and interpretation of **ordinal** data that is collected using Internet-based survey forms, and then conclude this section with a discussion of other works.

It is common for a group of well-intentioned individuals, engaged in purposeful dialogue, to utilize the concept of consensus in making decisions, especially when it is important to maintain some sort of collegiality. In America there exists a set of rules used by most boards and organizations as the arbiter of the structure for group discussions and it is called Robert's Rules of Order. While Robert's Rules are effective, it usually results in someone or some group *losing* in the resulting decision if the leader or chair calls for a vote having *sensed* that most are in agreement. Such feelings may be incorrect. Although consensus building is a typical method used in decision making, few measures exist that allow for the easy determination of the degree to which a group is nearing the point of agreement. When dealing with Internet-based surveys, the ordinal data collected must be analyzed to determine the level of consensus or agreement of the respondents with respect to the questions or issues raised. The purpose of this article is to show a mathematical measure (Tastle & Wierman, 2005, 2006a, 2006b, in press-a, in press-b, in press-c; Wierman & Tastle, 2005) that is intuitive, satisfies the requirements of a measure of consensus, and is easy to apply to the analysis of ordinal surveys.

The survey analysis requires finding some means by which the consensus of the respondents to an ordinal survey can be identified, understood, and compared. As a number of business and political analysts have pointed out in the past, there are problems associated with determining consensus in a group or by survey; the problems are similar. If a too-strict requirement of consensus is asserted, it is possible for a minority group to hold a veto power over decisions. Conversely, a too-loose requirement permits the domination of the minority by the majority, an equally undesirable outcome. It is entirely possible for a decision by consensus to take an extremely long time to occur, and thus may be inappropriate for urgent matters such as decisions involving strategic policy or competitive advantage. Sometimes, consensus decision making encourages groupthink, a situation in which people modify their opinions to reflect what they believe others want them to think. This can lead to a situation in which a group makes a decision that none of the members individually support and may lead to a few dominant individuals making all decisions. Fortunately, survey respondents are not impacted by this problem. Finally, consensus decision making may fail when there simply is no agreement possible, a problem that is theoretically possible when half of the survey respondents select *strongly agree* and the other half select *strongly disagree*. Does such a possibility justify reporting a *neutral* category being the average category chosen? Even if including a standard deviation in the resulting report (for a five category Likert scale, such a standard deviation is 2), there is an expectation and visual image by most readers of values scattered around a mean value of neutral, clearly an erroneous expectation.

Consensus (n.d.) has two common meanings. One is a general agreement among the members of a given group or community; the other is as a theory and practice of getting such agreements. Many discussions focus on whether agreement needs to be unanimous and even dictionary definitions of consensus vary. These discussions miss the point of consensus, which is not a voting system but a taking seriously of everyone's input, and a trust in each person's discretion in follow-up action. In consensus, people who wish to take up some action want to hear those who oppose it because they do not wish to impose, and they trust that the ensuing conversation will benefit everyone. Action despite opposition will be rare and done with attention to minimize damage to relationships. In a sense, consensus simply refers to how any group of people who value liberty might work together.

To capture how someone *feels* towards an issue under discussion, some mechanism must be used by which that person may express his/her opinions, but in a manner such that the data can be quantified. The Likert scale easily fulfills this requirement. Unfortunately, the Likert scale has no interval property. To *solve* this problem some have advocated placing numbers next to the linguistic labels, that is, strongly agree = 1, in an effort to force an interval. This does not work; the presence of an interval means that a respondent has carefully reviewed the available data (or searched his/her mind for a proper feeling) and has evidence that 2.1 is too high and 1.9 is too low, so the choice of 2 is checked. Forcing the presence of numbers does not change an ordinal scale to an interval scale. It remains simply a set of ordered categories and the use of ratio and interval scale mathematics is not conceptually sound when analyzing ordered categories, though the results are accepted as accurate. We propose another way of analyzing ordinal data, and it has great potential in e-business as we attempt to gather as much information as possible out of available data.

There is substantial work on the ranking of discrete data (Chamberlin, Cohen, & Coombs, 1984; Murphy & Martin, 2003) and every good statistics text has a section devoted to Kendall, Spearman, and Cayley rankings (Murphy & Martin, 2003), and sometimes the Hamming and Euclidian distances. Ranking is a means by which items in a collection can be evaluated such that any two items can be compared to see which should be placed higher in the ranking. Hence it is easy to see that presidential candidates can be ranked, as can the top golfers, the National Football League, or World Soccer teams, the flavors of ice cream, and attributes of a product. Unfortunately, we sometimes confuse ordinal ranking with ordinal measures. An ordinal ranking is the assignment of a unique ordinal number to all items in a collection. An ordinal measure is the assignment of a degree of acceptability, desirability, favor, discernment, and so forth to each single attribute. To ask a subset of Internet customers to rank the products in order of desirability is quite different from asking them to assess their agreement that property X is an important quality of product Y. In the latter ranking the customers merely need the list of products and a space next to each item into which their number value can be placed. In the former example, some ordinal scale is provided to which the customer will check a response. For example, in response to the statement "it is important for product Y to have property X" the ordinal scale might be *strongly agree*, *agree*, *neither agree nor disagree*, *disagree*, or *strongly disagree*. It is obviously not useful to take a set of responses to these Likert attributes and attempt to forcibly rank order them, and that is the purpose of this article: to show a new method by which such data can be evaluated from a perspective of group agreement. Using ranks, there is a winner and a looser! Using this novel method of assessing group agreement, each Likert category has a degree of agreement.

Davies (2005) investigated the combination of a fully anchored Likert scale with a numerical rating scale and found that by providing visual cues yielded a more discriminating result in which respondents more consistently applied their ratings. Applying our method to his data yielded identical results without the need for any other visual cues. The method presented here is computationally easy to apply and gives consistent results. We do, however, acknowledge that this work is still in-process, and much more must be done before the measure becomes main stream (see the Conclusion). We hope that readers will build upon our efforts.

METHOD

Let us assume that a data set of Likert scale responses has been collected by means of an Internet survey. The data are represented by a listing of numbers, each one from 1 to 5 representing the standard Likert categories from strongly agree to strongly disagree. We can apply standard statistics to this listing, but a more conceptually accurate method is offered on some reflections on the properties needed to analyze these data.

We postulate that the following set of rules must be satisfied before any measure can be considered a viable solution to the Likert-scale problem:

1. For a given (even) number of **n** individuals participating in a survey on some matter of interest, if **n/2** select the strongly disagree category and the other **n/2** select the strongly agree category, the group is considered to have **no** consensus. This is called the **diametric opposition** quality.
2. If all the participants classify themselves in the same category, regardless of the label given that category, then the group is considered to be in consensus.
3. If the mix of participants is such that **n/2 + 1** survey respondents assign themselves to any one category, the degree of consensus must be greater than 0, for the balance in the group is no longer equal.

Conversely, dissention requires the following set of rules be satisfied:

1. For a given (even) number of **n** individuals participating in a survey on some matter of interest, if n/2 select strongly disagree and the remaining **n/2** respondents select strongly agree, the group is considered to have **maximum dissention**.
2. If all the respondents classify themselves in the same category, regardless of the label given that category, then the dissention is considered to be **zero**.
3. If the mix of respondents is such that n/2 + 1 respondents assign themselves to any one category, the degree of dissention must be less than maximal.

Consensus and dissention are inverse functions of shared group feelings towards an issue. This *feeling* can be captured through a Likert scale that measures the extent to which a person agrees or disagrees with the statement under investigation. The most common scale is 1 to 5. Often the scale will be 1 = strongly agree, 2 = agree, 3 = not sure, 4 = disagree, and 5 = strongly disagree. Other number assignments can be made, such as: -2 = strongly agree, -1 = agree, 0 = not sure, 1 = disagree, and 2 = strongly disagree, or 0.0 = strongly agree, 0.25 = agree, 0.50 = neutral, and so forth. 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 (1990) has produced a very nice article that describes these aspects in straightforward terms. A rather complete bibliography can also be found there.

THE CONSENSUS AND DISSENTION MEASURES

The properties of a consensus measure is defined (Tastle & Wierman, 2005) 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 the random variable X represents the Likert scale values, X_i is the particular Likert attribute value, p_i is the probability associated with each X_i, d_X is the width of X, and $E(X) = \sum_{i=1}^{n} p_i X_i = \mu_X$ is the mean of X. This measure adequately fulfills the previous rules as evidenced by the following *illustrations*.

The mirror image of consensus is dissention and has the following form:

$$Dnt(\mathbf{X}) = -\sum_{i=1}^{n} p_i \log_2 \left(1 - \frac{|X_i - \mu|}{d_X} \right) \qquad (2)$$

In other words, *Cns = 1 – Dnt* and *Dnt = 1 – Cns*. One of the interpretations of the dissent measure is that of dispersion. If the frequency distribution is balanced on the extreme categories of the Likert scale, for example at *strongly agree* and *strongly disagree*, the dispersion is maximized at 1 (and the consensus is zero). As the frequency distribution approaches the assignment of all probability to a single category, the dispersion approaches 0 (and the consensus approaches one). This is the essence of the consensus measure: the more the respondent assignments are tightly clustered around one category, the higher the consensus and the less the dissent. This dispersion is always a value in the unit interval, [0..1].

Let us assume that we have a five-attribute Likert scale: strongly agree (SA), agree (A), neutral (N), disagree (D), and strongly disagree (SD). Let us further assign a numerical scale of SA = 1, A = 2, N = 3, D = 4, and SD = 5. Then X = {1, 2, 3, 4, 5}, X_1 = 1, and so forth, d_x = 5 - 1 = 4. Using an arbitrary number of random integer values to populate the scale, the following table denotes the required properties.

Table 1 contains data on eight aspects: the first column is simply an index of the rows, columns SA through SD denote the frequencies assigned to the Likert scale attributes (for comparison purposes all frequencies sum to 12), the expected

Table 1. Illustration of ten sets of values ranging from the most extreme (row 1) to the most concentrated (row 10). Calculations of the mean, standard deviation, consensus, consensus as a percent, dissent, and dissent as a percent, are shown.

	SA	A	N	D	SD	Mean	St Dev	Cns	Cns%	Dnt	Dnt%
1	6	0	0	0	6	3.0	2	0	0%	1	100%
2	6	0	0	1	5	2.917	1.93	0.049	5%	0.951	95%
3	6	0	0	2	4	2.833	1.86	0.097	10%	0.903	90%
4	5	1	0	2	4	2.917	1.80	0.146	15%	0.854	85%
5	5	1	2	4	0	2.417	1.32	0.425	43%	0.575	58%
6	1	5	2	4	0	2.750	1.01	0.605	61%	0.395	40%
7	1	5	4	2	0	2.583	0.86	0.675	68%	0.325	33%
8	0	6	5	1	0	2.583	0.64	0.758	76%	0.242	24%
9	0	9	3	0	0	2.250	.043	0.896	90%	0.104	11%
10	0	12	0	0	0	2.0	0	1	100%	0	0%

mean for the attribute values, the standard deviation for the attribute values, Cns and Cns% are the consensus values in decimal and rounded percent, and Dnt and Dnt% are the dissension values in decimal and rounded percent (Cns = 1 - Dnt). Row 1 shows a maximum amount of dissent in consensus since n/2 observations are reflected in each of the extreme attributes. As a point of interest, had the n/2 values been associated with agree and disagree, the consensus would have been 3.0 and the standard deviation 1.0, since these attributes are closer to each other. Rows 2 through 9 show a convergence of opinion moving towards agree. An examination of the mean column shows a modest fluctuation of the values but, in general, a movement of value from neutral (3) to agree (2). This is supported by the StdDev column as the values continue to converge towards 0 as the values surrounding the attributes merge. The consensus shows continuous movement towards 1; it is arguably easier to associate the consensus as a percent to easily visualize the movement towards a consensus. Conversely, one can monitor the dissent from total presence (row 1) to total absence (row 10). Finally, row 10 shows the attribute values firmly in one category. The mean is trivially at 2, the StdDev is now zero, consensus is complete at 100%, and dissent does not exist.

The proof that this measure of ordinal data satisfies the rules listed previously is found in Wierman and Tastle (2005).

THE AGREEMENT MEASURE

Consensus (Equation 1) can become agreement (Equation 3) when the mean μ_X is replaced with some target value, τ, and we divide by twice the width, $2d_x$, in the denominator. The target, τ, is usually some desired value identified by the manager. For our purposes let us assume that the desired response is *strongly agree*. Since that is the first category in our Likert item, it is assigned a numerical value of 1. Hence, in response to the declarative statement "Customer service is exceptionally good," we desire for our survey respondents to strongly agree with this statement, that is, the target is $\tau=1$. This measure is called *agreement* to distinguish it from measures that use an unspecified target such as the mean, median, or mode. Equation 3 shows τ in place of μ and an expanded width. Doubling the width prevents the equation from exploding when extreme values are reflected in the frequency distribution. We have found the agreement function to work especially well in practice and, for this current work, have limited ourselves to the $2d_x$ denominator. For the most part, either consensus or agreement will work very well, but it is necessary to be consistent in their use. It should also be mentioned that consensus, dissent, and agreement are invariant with respect to linear transformations of the random variable *X*.

$$Agr(X|\tau) = 1 + \sum_{i=1}^{n} p_i \log_2\left(1 - \frac{|X_i - \tau|}{2d_X}\right) \quad (3)$$

While targeting provides a novel way of measuring distance from a desired goal, it assumes that all elements of the assessment are equally important.

Table 2 shows the Table 1 Likert data and the mean, and adds the agreement measures for each of the Likert categories. The usefulness of the measure becomes evident with practice. For example, the first row shows a higher agreement for target 3 (neutral) than for the other values, and the extreme values, 1 (SA) and 5 (SD), have the smallest measure of agreement. At first, this seems counterintuitive, but the agreement measure is actually mid-way between consensus and consent for SA and SD. Table 1 shows a consensus for the entire first row having the value of 0. With respect to the entire distribution of categories, there is no *consensus* whatsoever. However, with respect to the neutral category there is a 50% *agreement*. When the set of surveys support the extreme categories, there is a de facto agreement on the middle category (neutral in this case) but the level of agreement is certainly not 0, nor is it 100%. It is logical that some middle value is appropriate, like 50%. Looking down to the values in row 7 we note that there are more respondents who have selected A than any other category. The mean is 2.58, which indicates that the average is almost between agree and neutral, perhaps agree-and-a-half. There is a 68% consensus on the part of the respondents with a dispersion of about 33%. The respondent values are becoming clustered, but what is the data telling us? The agreement with a target

Table 2. Illustration of the same ten sets of values ranging from the most extreme (row 1) to the most concentrated (row 10). Calculations of the mean and agreement values for each category, that is, Agt(1) is read as "agreement with respect to strongly agree as the target."

	SA	A	N	D	SD	Mean	Agt(1)	Agt(2)	Agt(3)	Agt(4)	Agt(5)
1	6	0	0	0	6	3.0	0.500	0.565	0.585	0.565	0.500
2	6	0	0	1	5	2.917	0.527	0.587	0.603	0.581	0.484
3	6	0	0	2	4	2.833	0.554	0.608	0.622	0.597	0.468
4	5	1	0	2	4	2.917	0.538	0.625	0.641	0.619	0.495
5	5	1	2	4	0	2.417	0.689	0.749	0.747	0.651	0.393
6	1	5	2	4	0	2.750	0.625	0.813	0.821	0.738	0.501
7	1	5	4	2	0	2.583	0.668	0.851	0.853	0.706	0.464
8	0	6	5	1	0	2.583	0.674	0.885	0.888	0.712	0.472
9	0	9	3	0	0	2.250	0.752	0.952	0.856	0.641	0.388
10	0	12	0	0	0	2.0	0.807	1.000	0.807	0.585	0.322

of neural has the greatest value, 0.853, which is interpreted as an 85.3% agreement of the respondents for the neutral category. The four who selected neutral are not the deciders of the agreement but rather, the five plus 2 that surround it. Agreement takes into account all the data in the distribution.

Finally, examination of row 10 shows complete consensus for the overall distribution as would be expected with all respondents having selected the same category, and the agreement with respect to agree is also 1.0. It is also evident that there is some agreement with the contiguous categories, like 80% agreement with both strongly agree and neutral, and even a modest level of agreement with strongly disagree of 32%. The absence of data from one or more categories does not mean an absence of agreement. All agreement values are shown in Figure 1.

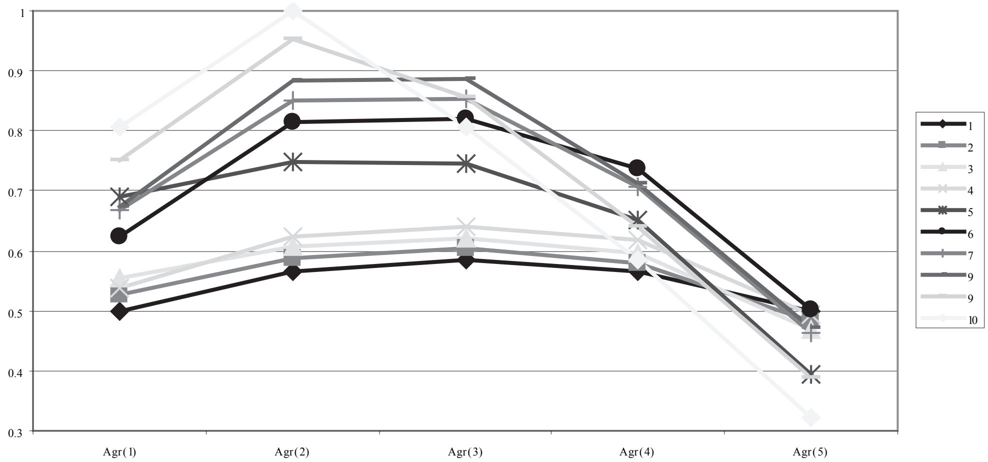
CONCLUSION

Data collected through the Internet can be analyzed in many statistically proper ways, but the fundamental premise of the presence of an interval or ratio scale is absent from ordinal data. This method is a new way of examining ordered ordinal data, is very intuitive, and requires little effort in calculating. The authors have used a spreadsheet to perform the calculations. The category that is most targeted by the survey respondents can be identified, and the degree of overall consensus with respect to the frequency distribution is interpreted as a percentage of the whole. Also, the measure of dissent is an indicator of dispersion. A visual representation of the proximity of categories can show confusion (multiple categories with close agreement measures, or a mandate for a particular category (steep drop-off in agreement measure for categories on either side). Combined, these measures permit the examination of data from any entirely new perspective.

Other work in the identification of measures to assist group decision making and analysis of non-interval or ratio scale data offers considerable potential for the interested researcher. What follows is an incomplete listing of potential opportunity:

- The authors think that a statistical change in Likert categorical importance, assuming a normal distribution of potential responses, can be approximated by the normal measure of significance. However, it has not yet been proven.

Figure 1. The measures of agreement for all 10 rows of data shown in Table 2



- The Likert scale demands the selection of only one category, but what about the possibility of degrees of categorical selection, as would be found in a typical fuzzy set.
- Our preliminary investigations suggest that the presence or absence of an interval does not change the result of the measure. Thus, we can be very confident in the resulting choice of the selected category, but we have not yet undertaken this research.
- We seek to identify a measure of consistency; preliminary attempts using covariance do not satisfy our intuition as to the properties we require of such a measure.
- Lastly, nominal measures are reputed to have only one valid statistical measure, that of the mode. We have found, only in the most preliminarily way, that a measure of dispersion can be validly calculated for nominal measures, but much work remains before an article can be written.

Using these measures and the family of measures that could develop from them, analysis of nominal and ordinal data might be able to move past the traditional statistical approach

ACKNOWLEDGMENT

The authors are grateful to the insightful comments from the anonymous reviewers. We wish we could thank them in person, but that would invalidate the peer review process so we thank them publicly.

REFERENCES

Chamberlin, J. R., Cohen, J. L., & Coombs, C. H. (1984). Social choice observed: Five presidential elections of the American Psychological Association. *The Journal of Politics, 46*(2), 479-502.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: Routledge Falmer.

Consensus. (n.d.). Retrieved from http://essential-facts.com/primary/ethics/Consensus.html

Cook, C., Heath, F., & Thompson, R. (2000). A meta-analysis of response rates in Web- or Internet-based surveys. *Education and Psychological Measurement, 60*(6), 821-836.

Couper, M. (2000). Web surveys: A review of issues and approaches. *The Public Opinion Quarterly, 64*(4), 464-494.

Denscombe, M. (2006). Web-based questionnaires and the mode effect. *Social Science Computer Review, 24*(2), 246-254.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38,* 1217-1218.

Kaplowitz, M., Hadlock, T., & Levine, R. (2004). A comparison of Web and mail survey response rates. *Public Opinion Quarterly, 68*(1), 94-101.

Korgaonkar, P., & Wolin, L. (1999). A multivariate analysis of Web usage. *Journal of Advertising Research, 39*(2), 53-68.

Munshi, J. (1990). A method for constructing Likert scales. Retrieved June 2004, from http://www.munshi.4t.com

Murphy, T. B., & Martin, D. (2003). Mixtures of distance-based models for ranking data. *Computational Statistics & Data Analysis, 41,* 645-655.

Parsons, C. (2007). Web-based surveys: Best practices based on the research literature. *Visitor Studies, 10*(1), 1064-5578. Retrieved July 15, 2007, from http://www.informaworld.com/10.1080/10645570701263404

Pell, G. (2005). Use and misuse of Likert scales. *Medical Education, 39,* 970.

Porter, S., & Whitcomb, M. (2005). E-mail subject lines and their effect on Web survey viewing and response. *Social Science Computer Review, 23*(3), 380-387.

Schmidt, J., Calantone, R., Griffin, A., & Montoya-Weiss, M. (2005). Do certified mail third-wave follow-ups really boost response rates and quality? *Marketing Letters, 16*(2), 129-141.

Solomon, D. (2001). Conducting Web-based surveys. *Practical Assessment Research and Evaluation, 7*(19).

Stanton, J. (1998). An empirical assessment of data collection using the Internet. *Personal Psychology, 51*(3), 709-725.

Tastle, W. J., & Wierman, M. J. (2005). Consensus and dissention: A new measure of agreement. *North American Fuzzy Information Processing Society (NAFIPS) Conference,* Ann Arbor, MI.

Tastle, W. J., & Wierman, M. J. (2006a). An information theoretic measure for the evaluation of ordinal scale data. *Behavioral Research Methods, 3*, 487-494.

Tastle, W. J., & Wierman, M. J. (2006b). Consensus and dissension: A new measure of agreement. In *NAFIPS 2006,* Montreal, Canada.

Tastle, W. J., & Wierman, M. J. (in press-a). Consensus: A new measure of ordinal dispersion measure. *International Journal of Approximate Reasoning.*

Tastle, W. J., & Wierman, M. J. (in press-b). The development of agreement measures: From general to targeted. *International Journal of Approximate Reasoning.*

Tastle, W. J., & Wierman, M. J. (in press-c). Determining risk assessment using the weighted ordinal agreement measure. *Journal of Homeland Security.*

Wierman, M. J., & Tastle, W. J. (2005). Consensus and dissention: Theory and properties. *North American Fuzzy Information Processing Society (NAFIPS) Conference,* Ann Arbor, MI.

William J. Tastle received his PhD in advanced technology with specialization in systems science from the Thomas J. Watson School of Engineering and Applied Science of the State University of New York, University Center at Binghamton in 1994 and an MBA in MIS from the School of Management at Binghamton. He is a professor of information systems in the business school at Ithaca College and is active in the IS community having served as the president of the Association for Information Technology Professionals, Education Special Interest Group (EDSIG), and on many other committees. He is the managing editor of the International Journal of General Systems. Currently he is organizing the 2008 conference of the North American Fuzzy Information Processing Society and serves on its board of directors as treasurer. He is a consultant for numerous organizations including Los Alamos National Laboratory. Dr. Tastle's current areas of interest involve measures of consensus, dissention, agreement and disagreement, requirements analysis elicitation, and outsourcing.

Mark J. Wierman is an assistant professor of computer science at Creighton University in Omaha, Nebraska. He received his PhD in Systems Science from Binghamton University. His current research interests are the application of the mathematics of uncertainty to political science and generalized information theory. He has published over 30 papers and his second book, Applying Fuzzy Mathematics to Formal Models in Comparative Politics *(with TD Clark, JM Larson, JN Mordeson, and JD Potter) will be published in spring 2008.*