Practical Assessment, Research, and Evaluation

Volume 17 *Volume 17, 2012* Article 5

2012

Terminating Sequential Delphi Survey Data Collection

Sema Kalaian

Rafa M. Kasim

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 17, Number 5, January 2012 ISSN 1531-7714

Terminating Sequential Delphi Survey Data Collection

Sema A. Kalaian, *Eastern Michigan University*
Rafa M. Kasim, *Statistical and Research Consultant*

The Delphi survey technique is an iterative mail or electronic (e-mail or web-based) survey method used to obtain agreement or consensus among a group of experts in a specific field on a particular issue through a well-designed and systematic multiple sequential rounds of survey administrations. Each of the multiple rounds of the Delphi survey administration is augmented with continuous summary feedback of aggregated responses from the same group of experts. Statistical methods to analyze data from the Delphi surveys to make decisions for terminating subsequent Delphi data collection are needed to ensure that (a) stability of the responses of the panel of experts is reached; and (b) termination of the rounds of the Delphi survey administration is based on sound statistical results. The present study presents an overview of the parametric and nonparametric statistical methods that can be used to analyze the structured Delphi survey data to make decisions about terminating the sequential Delphi survey data collection.

The *Delphi technique* is an iterative and sequential mail or electronic (e-mail or web-based) survey method for forecasting and decision-making purposes to obtain informed anonymous agreement and consensus among a panel of experts and leaders in the field on a particular issue or problem. The process of the Delphi survey can be accomplished through a well-designed multiple sequential administrations of either mail or electronic survey questionnaires augmented with continuous summary feedback of aggregated responses of a panel of experts. The Delphi methodology was developed in the early 1950's by the Rand Corporation to structure interactive and informed communication among a group of experts in the military to solve complex security problems such as the possibility of a military or nuclear missile attack (Yang, 2003). Since then, the use of the Delphi survey methodology has increased in many diverse academic disciplines and fields of study such as business, technology, management, education, medicine, nursing, health, and family therapy. In a typical Delphi study, the opinions of a carefully selected group or panel of experts are sought concerning various different complex issues or problems for informed decision-making purposes. Experts are defined as being the qualified and experienced professionals and scientists who have the relevant knowledge and expertise about a particular issue or problem. In other words, the Delphi technique is an organized systematic way of arriving at an informed and consensus-based decision or opinion regarding an important organizational issue or problem such as policy-oriented issue, managerial problem or decision, certification issue, organizational needs assessment, curriculum improvement, financial planning or forecasting, or any other problem or issue (Blair & Norman, 1993; Kalaian & Shah, 2006; Linstone & Turoff, 1975; Shah & Kalaian, 2009; and Yang, 2003). Therefore, the Delphi survey method is structured to maximize on the merits of anonymous informed expert group decisions and judgments while minimizing the liabilities of the individual expert's decisions (Dunham, 1996; Yang, 2003).

Many practical and logistical advantages of the Delphi survey methodology have been identified in the literature. One major advantage is that it provides in-depth anonymous information about the problem or issue under consideration. The second advantage is that it is economical because it makes the collection of opinions from the geographically dispersed experts possible via paper-and-pencil or electronic mail surveys (Delbecq & Van de Ven, & Gustafson, 1975). The third advantage, it is an anonymous and confidential process that enhances the consensus-based decision-making process and judgments. The fourth advantage, it is time-efficient methodology to reach an agreement between experts and leaders in a specific field of study to solve complex problems or policy issues that require informed decisions and judgments. The fifth and finally, it produces more objective and accurate informed expert solutions, judgments, and policies for complex problems than the traditional face-to-face meetings and discussions. It is important to note that the Delphi survey methods also have disadvantages. One of the major disadvantages is that during the course of multiple sequential rounds of collecting Delphi data some members of the experts may not return one or more of the survey questionnaires. This is considered a significant problem in a Delphi survey methodology because those experts who don't return some of the mail or electronic questionnaires are excluded from the panel of experts for further Delphi data collection. This exclusion of experts consequently leads to the elimination of these experts' records from the final Delphi data set to be analyzed to make decisions about terminating the subsequent administration of the Delphi survey.

Given the iterative nature of the Delphi survey methods, which are based on many sequential rounds of survey administration to reach a consensus among a panel of experts, there is a need to use appropriate statistical methods to make sound decisions about terminating the rounds of the Delphi survey administration and data collection. Traditionally, the Delphi researchers suggested comparing the averages or percentages of responses for each question from any two consecutive rounds of the Delphi survey administrations. To date, this method is the typical technique that is used for analyzing the collected Delphi survey data. Based on these simple mean and percentage comparisons, the Delphi researcher concludes that no additional round for administering the Delphi survey is needed because of little change in the averages (or percentages) of the responses of the experts between the pair of consecutive Delphi survey rounds (Yousuf, 2007; Hsu & Sandford, 2007). Instead of these simple methods that are based on the means and percentages, more appropriate and statistically sound parametric and nonparametric analytical methods are needed to analyze the collected data from the Delphi survey administration rounds. These analytical methods ensure that (a) all expert opinions are represented in the data set based on sound statistical procedures; and (b) terminating the Delphi rounds is based on sound statistical results and not just the comparisons of the simple summary statistics such as the averages and percentages. The results of the Delphi survey that are based on appropriate parametric and nonparametric statistical analytical methods, help the Delphi survey researchers to make sound and valid decision for terminating the Delphi data collection. Therefore, the main objective of this study is to present an overview of the parametric and nonparametric analytical methods to be used for terminating the sequential rounds of the Delphi data collection. In other words, these analytical methods can be used to analyze data, which are collected from administering sequential Delphi surveys to measure the stability and consensus of the responses of the panel of experts and make informed decisions about the termination of the Delphi survey data collection rounds based on sound and appropriate statistical methods.

STEPS OF THE DELPHI SURVEY METHODOLOGY

A typical research process of the Delphi method consists of a preparation step followed by a series of rounds of survey administrations to a panel of experts in the field of study. The planning of the Delphi survey study which includes developing operational definitions of the specific problems or issues, developing the survey questions, administering and collecting the Delphi survey responses, preparing summary feedback reports, and

DOI: https://doi.org/10.7275/g48q-je05

executing the systematic steps of the Delphi survey methodology are usually conducted by the primary researcher or a team of researchers. The following is a brief introduction to the steps necessary for conducting a research study using the Delphi survey methodology to reach an informed agreement and consensus among the experts on a specific issue or problem.

1. At the first round of the Delphi survey administration, the experts in the panel anonymously answer a small number of open-ended survey questions focused on the issue or problem under consideration. These open-ended questions give each expert in the panel a chance to suggest other alternatives as possible considerations or solutions to the issue or the problem under investigation. These alternatives become part of the Delphi-survey in the second round. It is important to note that, instead of the open questions, it is possible that the survey questions can be structured and that close-ended questions are used at the first round of the Delphi survey methodology.
2. The open-ended survey responses from the first round are reviewed and categorized by the Delphi survey research team to create a valid and reliable list of structured and Likert-type closed-ended questionnaire items to be used for the second round of the Delphi survey administration.
3. At the second round of the Delphi survey administration, the same panel of experts is provided with the closed-ended survey questions developed from the responses from the first round to get their expert opinions about the newly developed closed-ended questions. Round two Delphi survey administration allows the panel of experts to recommend changes and suggest additions and/or deletions to the survey questions. At this step, the Delphi survey is also accompanied by an anonymous summary of the expert responses from the first round in categorized or aggregated form such as frequency distributions, summary statistics, and graphical representations. Expert panelists are invited to confirm or modify their first round responses based on their reactions to the collective responses of the panel members of experts that are documented and presented in the summary reports.
4. The survey responses from round two are reviewed and analyzed by the research team to provide a comprehensive description of the experts' consensus and agreement on the issue or problem. This consensus and agreement report should be based on the parametric and nonparametric statistical methods that are presented in this article (e.g., Coefficient of Variation, F-ratio, McNemar Change Test). If agreement or consensus among the expert panelists is not reached, then a third round of survey administration is needed.
5. At the third round, a revised closed-ended questionnaire is accompanied by a summary of the findings from the second round (e.g., frequency distributions and graphs of the responses for each item in the survey) are sent to the same experts for their informed input and opinion regarding the issue or problem under consideration. It is important to note that at each additional Delphi structured survey administration, a revised closed-ended questionnaire should be developed and administered to the same panel of experts. The revised questionnaire should include only the items that consensus is not reached based on the results of the statistical methods that are described below (e.g., Coefficient of Variation, F-ratio, McNemar Change Test) from the two previous Delphi survey administrations. The items that reached consensus should be dropped from the Delphi survey for further survey administration purposes.

Finally, the Delphi rounds of questionnaire administration should continue until a predetermined level of consensus is reached or no new information is gained from further rounds of administering the Delphi survey (Linstone & Turoff, 1975). In most Delphi survey applications it is found that three iterations of the rounds of the Delphi survey data collection are enough to reach a consensus among the panel of experts (Kalaian & Shah, 2006; Yang, 2003).

The next sections cover a list of possible statistical methods (Parametric and non-parametric) that one can use to make decisions about terminating the Delphi survey administration rounds. It is important to realize that at this point we are not promoting a choice of one method over the others. Rather, we think that a good practice is to collect evidences from as many statistical methods as possible based on valid assumptions in support of the timing for the termination of sequential rounds of administering the Delphi-survey.

PARAMETRIC METHODS FOR ANALYZING DELPHI DATA

The following is an overview and presentation of four parametric statistical methods to be used in Delphi studies for setting stopping criteria for further rounds of survey administration and data collection (Shah & Kalaian, 2009; Yang, 2003). These *parametric* statistical methods require that the Delphi survey study to have at least 30 experts. *Nonparametric* statistical methods need to be used if the number of experts in the panel is less than 30 (Kalaian & Shah, 2006; Shah & Kalaian, 2009; Yang, 2003).

1. Coefficient of Variation (CV) difference for an item from two consecutive rounds.

Coefficient of Variation (CV) is the ratio of the standard deviation of the responses of the panel of experts on a specific item to its corresponding mean (average). Therefore, the responses of the experts for each of the items in the survey from each of the rounds of the Delphi survey data collection will yield one coefficient of variation. For example, a Delphi survey with 25 items will yield 25 coefficients of variation for each of the rounds of the Delphi survey administration. The Coefficient of Variation (CV) is calculated using:

$$CV = \frac{Standard\ Deviation}{Mean}$$

A large value of the coefficient of variation (CV) for an item in the Delphi survey (larger than 1) indicates that the responses of the experts are scattered compared to the mean of the responses for the item. In other words, the amount of the variation among the opinions (responses) of the expert panelists compared to the mean (average) of the item is large. In contrast, a small value of CV indicates that the amount of variation among the responses of the expert panelists is small compared to the mean of the responses to the item.

Next, to measure the stability of the responses for an item an absolute CV difference for each item needs to be calculated by subtracting CV obtained for each item in the Delphi survey from any two consecutive rounds (Dajani, Sincoff, & Talley, 1979; Kalaian & Shah, 2006; Shah & Kalaian, 2009; Yang, 2003). For example, a Delphi survey study with two rounds (x and y) will have the following CV difference between the coefficients of variation for each of the items in the structured Delphi survey

$$CV\ Difference\ (Rounds\ x\ \&\ y) = CV\ (Round\ y) - CV\ (Round\ x).$$

A small value and close to zero of absolute CV difference for each item in the Delphi survey indicates that the stability of responses and consensus among the experts on a specific item from the two rounds is reached and there is no need for further rounds of survey administration and data collection. In contrast, a large value of absolute CV difference indicates that a consensus or agreement among the expert panelists for a specific item is not reached and there is a need to further rounds of data collection from the same experts after dropping the items that have an absolute CV difference value of zero or close to zero.

2. F-ratio for comparing the variances of an item from two consecutive rounds.

An *F-ratio* is performed to determine the ratio of two variances of an item (question) for any two consecutive rounds of a Delphi survey administration. In other words, an F-ratio is used to examine the equality of the variances of an item from two consecutive rounds of Delphi survey administration. For example, in a Delphi study with two rounds, there will be one F-ratio representing the ratio of the variances of an item from Round y (e.g., Round 3) and Round x (e.g., Round 2). The F-ratio is represented as

$$F - ratio\ (Rounds\ x\ \&\ y) = \frac{Variance\ of\ the\ item\ (Round\ y)}{Variance\ of\ the\ item\ (Round\ x)}$$

An F-ratio value of one indicates that the variances of the same item from the two rounds of the Delphi survey administration are equal and a perfect stability of consensus is reached. On the contrary, any further distance and deviation from a ratio of one is an indication of a lack in stability of consensus as well as the need for another round of Delphi survey administration. Hence, when the value of F-ratio of an item from the two rounds is one or close to one for an item, then the item should be dropped from further rounds of Delphi survey administration because it indicates that a stability of consensus is reached regarding this specific item. It is important to note that we do not recommend this statistical method to be used for terminating further rounds of Delphi data collection because the assumption of independence of the two variances for an item, which is required for conducting the F-ratio, is violated when using data collected from the same panel of experts in any two consecutive Delphi rounds of data collection (Yang, 2003).

3. Pearson correlation coefficient for the experts' responses on an item from two consecutive rounds.

A *correlation coefficient* (*r*) is also known as the Pearson product-moment correlation coefficient or Pearson r and is a measure of the strength of the association between two quantitative variables. In this case, it is the relationship between the responses of the panel of experts for each item from the two consecutive rounds of the Delphi survey administrations. For example, if the Delphi survey consists of ten questions, then we will have ten Pearson r correlation coefficients from each of the two consecutive rounds (one correlation coefficient for each item).

A high and positive correlation coefficient (*r*) of one ($r = 1$) or close to one indicates that the ratings (responses) of the panel of experts on an item are similar across the two rounds of the Delphi survey administrations. This implies that stability of the consensus is reached for the specific item and there is no need to include the item in the survey for subsequent rounds of the Delphi survey administrations. A correlation coefficient (r) of zero or close to zero indicates no relationships between the ratings from the two rounds of the Delphi survey administrations. This implies that a consensus is not reached for the specific item and the item needs to be included in the subsequent rounds of the Delphi survey administration.

The reader should note that the Pearson correlation relies on the assumptions of (a) normality of the responses from the two rounds of the survey, and (b) a linear relationship between the responses from the two consecutive rounds. The correlation is not robust to these assumptions and violating such assumptions can lead to an artificial deflated correlation value. One possibility for not meeting such assumptions is the small number (less than 30) of experts. The other possibility is the skewness of the distributions of the responses. In such cases, a non-parametric correlation (e.g., Spearman's rank correlation, presented below) should provide an accurate measure of association between the responses from the two consecutive rounds of the survey.

4. Paired t-test for experts' responses on an item from two consecutive rounds.

The paired t-test evaluates whether or not the mean of the difference in responses to an item of the Delphi survey from two consecutive rounds is equal to zero. A zero mean difference in responses indicates no change in the responses of an item between the two rounds of the survey administration. The test provides the average of the response in each round with a t-test statistics and its associated p-value. Table 1 shows a hypothetical data representing the ratings of 11 experts on a

Delphi survey item from two consecutive rounds of administering the survey.

Paired t-test statistical procedure is used to analyze this data using the SPSS software. The t-value of the test equals -1.44 with a p = 0.18 (see Table 1) suggest that the difference in the ratings of the responses to the item from the two rounds of the Delphi survey is not significantly different from zero. This result indicates that there is little change in the responses from the two consecutive rounds (Round 1 and Round 2) and this specific item should be removed from the Delphi survey in the third round of the survey administration.

Table 1: Ratings of the panel members on a specific item from the two consecutive rounds of the Delphi survey administrations

Experts	Round 1 (R1)	Round 2 (R2)
Expert 1	3	2
Expert 2	5	4
Expert 3	3	1
Expert 4	4	4
Expert 5	2	5
Expert 6	2	1
Expert 7	3	5
Expert 8	3	5
Expert 9	1	4
Expert 10	2	5
Expert 11	2	3
Item Mean	2.73	3.55
Item S.D.	1.10	1.57
R1-R2 Mean	-0.82	
R1-R2 S.E.	0.57	
t-value	-1.44	
p-value	0.18	

NONPARAMETRIC METHODS FOR ANALYZING DELPHI DATA

The following are the nonparametric statistical methods that are suggested in the literature (Yang, 2003; Kalaian & Shah, 2006) that can be used to analyze categorical and ordinal Delphi survey data. These methods are recommended to be used instead of the parametric methods when the number of experts in the panel is less than 30 and/or the distribution of the responses for each of the items are skewed (non-normal distribution).

1. *McNemar Change Test of the responses of the experts on an item from two consecutive rounds.*

McNemar Change Test is used in the Delphi survey studies when the responses to the Delphi questions are dichotomous (e.g., yes-no or agree-disagree responses). A hypothetical example showing the structure of the dichotomous Delphi question is presented in Table 2. The A and D diagonal cells are those indicating a change in response from the previous round to the current round (in this example, a change in response from Round 2 to Round 3). The B and C diagonal cells are those indicating no change between the responses of the panel of experts from the two consecutive rounds of the Delphi survey administrations (Yang, 2003).

Table 2. The structure of dichotomous data from two rounds of a Delphi study

Delphi Survey Rounds		Round 2	
		Disagree	Agree
Round 3	Agree	A	B
	Disagree	C	D

In the above example, the interest is only in the cells that reflect a change of opinion about the question, which are cells A and D. The null hypothesis in this case is: There are an equal number of changes in responses of the panel of experts in both directions (in this example, changes from Agree to Disagree and changes from Disagree to Agree). Hence, we are testing that the expected frequencies in cell A will be equal to the expected frequencies in cell D. Using McNemar Chi-Square test statistic that is computed as

$$McNemar = \frac{(A-D)^2}{(A+D)}$$

A continuity correction is applied to the above equation to make the approximate sampling distribution of Chi-square calculated from the above equation more precise (Siegel & Castellan, 1988; Yang, 2003), which can be represented as:

$$McNemar_{adj} = \frac{(A-D-1)^2}{(A+D)}$$

In this example, there is one degree of freedom associated with McNemar change statistic because there are two rows and two columns in the above table. Degrees of freedom equal (number of rows – 1) x (number of columns – 1) and in this example equal (2-1) x (2-1) = 1. Looking up for the critical value of a Chi-square distribution with one degree of freedom at α = 0.05 (Chi-square Distribution Table can be found in any statistics textbook or online), we get 3.841. If the computed McNemar Chi-square value is smaller than the critical Chi-square value, which is 3.841, the null hypothesis is not rejected. Hence, for this example, it can be concluded that the experts' opinions about the specific issue have not significantly changed from one direction to the other from the two consecutive Delphi survey administrations (from Round 2 to Round 3). Therefore, this specific item should be dropped from the Delphi survey and not be included in the subsequent round of the Delphi survey administration (in this example, Round 4).

2. Spearman's Rank Correlation Coefficient between the ratings of the experts on an item from two consecutive rounds.

Spearman's rank correlation, r_s represents the correlation between the ranks of the ratings or responses of the experts for each of the items in the Delphi survey. It is calculated by using

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

where, d_i is the difference between the ranks of the responses or the ratings on the i^{th} item of the Delphi survey, for example, from rounds 1 and 2. The number of experts in a panel is represented by *n*.

The value of r_s always falls between -1 and +1, with +1 indicating perfect positive correlation between the responses on an item from two consecutive rounds. The closer r_s falls to +1, the greater the correlation between the ranks from the two consecutive rounds. The closer r_s is to 0, the less the relationship indicating no correlation between rankings from the two consecutive rounds. The closer r is to -1 the greater the correlation between the responses (e.g., ratings) of the item in an opposite direction indicating a disagreement in the responses of the experts on the item from the two consecutive rounds.

For example, in a hypothetical Delphi study two rounds of a survey questionnaire are administered. The six panel members rate the items on a Likert-scale of 1 to 5 as shown in Table 3.

Table 3. Ratings of the panel members on a specific item from the two rounds of Delphi survey administrations

Experts (*i*)	Round 1	Round 2	Difference between Ratings (d_i)	d_i^2
Expert 1	1	2	-1	1
Expert 2	4	5	-1	1
Expert 3	3	4	-1	1
Expert 4	4	5	-1	1
Expert 5	2	2	0	0
Expert 6	5	3	2	4
$\sum di^2$				8

Calculating r_s for the above data using the above formula for the Spearman Rank Order Correlation we get

$$r_s = 1 - \frac{6(8)}{6(6^2-1)} = 0.772.$$

The value of 0.772 shows a positive correlation between the ratings of the same experts on the specific item from the two consecutive rounds of the Delphi survey administration (in this example, Round 1 and Round 2). Looking at the table for the critical values of the Spearman's rank correlation coefficient, at α = 0.05, we get the value 0.829. Since the calculated r_s = 0.772 is smaller than the critical value of 0.829, we conclude that the association between the ratings of the panel members on this particular item from the two rounds of the Delphi survey administration is not significantly different from zero. Since the test result of the relationship between the two ratings on an item from the two consecutive rounds is not significantly strong, this specific item should be included in the next round of the Delphi survey administration (in this example, Round 3).

3. Wilcoxon Paired Signed-Ranks T Test.

Wilcoxon paired signed-ranks T test assesses whether or not the ranks of the difference in responses to an item on the survey from two rounds is equal to a zero. In other words, there is no difference between ranks of the responses of the experts from the two rounds (Privitera, 2012). The test is used as the nonparametric alternative to the parametric paired t-test (see the parametric section of this article). The test provides the sum of each of the positive and negative ranks of the differences between any consecutive rounds of Delphi survey responses (e.g., ratings) with a Z statistic and its asymptotic p-value. A hypothetical data of the ratings of 8 experts on items of a Delphi survey is listed in Table 4.

We used SPSS to perform the Wilcoxon paired signed ranks T test to analyze the data. The SPSS results show (see Table 4) that 3 ranks of the difference between Round 3 and Round 2 are negative with a mean sum negative ranks of 4.83. The results also show (see Table 4) that 4 ranks of the difference between Round 3 and Round 2 are positive with a mean sum positive of 3.38. The Z-value of the test is -0.087 with a p-value of 0.93. These results suggest that the differences in ranks of the responses to the item from the two rounds of the Delphi survey (Round 2 and Round 3) are not significantly different from zero. This result indicates that there is little change in the responses from the two consecutive rounds (in this case, Rounds 2 and Round 3) and the item should be removed from the survey.

Table 4. Wilcoxon Paired Signed-Rank T Test for the ratings of the panelists on a specific item from two rounds of Delphi survey administrations

Experts	Round 2 (R2)	Round 3 (R3)
Expert 1	2	3
Expert 2	4	5
Expert 3	3	1
Expert 4	3	3
Expert 5	4	2
Expert 6	1	3
Expert 7	5	4
Expert 8	3	4
Mean Sum Negative Ranks of R3-R2 (#)		4.83 (3)
Mean Sum Positive Ranks of R3-R2 (#)		3.38 (4)
# Tied Ranks of R3-R2		1
Z-value		-0.087
p-value		0.93

CONCLUSIONS

The Delphi survey is a sequential and iterative mail or electronic (e-mail or web-based) survey data collection method to investigate and gain informed anonymous consensus among a panel of experts on a particular complex issue or problem in a specific field of study. It is used as an alternative method to conventional face-to-face meetings to anonymously gain consensus for differing and alternative ratings of positions and opinions between the expert panelists regarding an issue or problem. The ultimate goal is to make informed decision about a specific issue or problem.

In this paper we presented four parametric and three nonparametric analytical methods that can be used to analyze a collected Delphi survey data to make decisions about terminating further Delphi data collection rounds. The use of these statistical methods are needed to ensure that (a) all expert opinions are represented in the data set; and (b) terminating the Delphi survey rounds is based on sound statistical results, which helps the Delphi survey researchers to conclude that there is no significant variation existing among the opinions of the experts. These parametric and non-parametric statistical methods are easy to calculate and the calculations can be done with the help of any available spreadsheet (e.g., EXCEL), statistical software (e.g., SPSS), or a calculator. It is important to note that the parametric statistical methods require that the Delphi survey study involves 30 or more experts and/or the distribution of the responses for each item is normal. Otherwise, nonparametric statistical methods such as the McNemar Change Test should be used. In other words, the choice of any of the four parametric and the three nonparametric methods presented in this article should be based on (a) number of the expert panelists; and (b) meeting the assumptions of the statistical method. For example, it is recommended that nonparametric methods should be used if the number of the experts in the Delphi study is less than 30 or the distributions of the responses to the Delphi survey questions are skewed (not normal). Shah and Kalaian (2009) compared the three parametric method presented above and found that the three methods yielded similar conclusions for terminating the Delphi data collection process. They recommended using the Coefficient of Variation (CV) method for termination decisions.

To avoid the problems associated with using the traditional simple means and percentages of the responses on each item in the Delphi survey and comparing these means and percentages across the rounds of the Delphi survey, which are suggested in the literature, Delphi survey researchers need to use one of the four parametric or one of the three nonparametric statistical methods presented in this paper to make valid, objective, and informed conclusions about the results of the Delphi study. Therefore, analyzing the collected data from a Delphi study using the appropriate statistical methods is a necessary step and should be taken into consideration when planning and designing a Delphi study to have successful findings and conclusions for policy making purposes.

REFERENCES

Blair, S., & Uhl, N.P. (1993). Using the Delphi method to improve the curriculum. *The Canadian Journal of Higher Education, 23*(3).

Dajani, J. S., Sincoff, M. Z., & Talley, W. K. (1979). Stability and Agreement Criteria for the Termination of Delphi Studies. *Technological Forecasting and Social Change*, 13, 83-90.

Delbecq, A.L., Van de Ven, A.H., & Gustafson, D.H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes.* Scott Foresman and Company: Glenview, IL.

Hsu, Chia-Chien & Sandford, Brian A. (2007). The Delphi Technique: Making Sense of Consensus. *Practical Assessment Research & Evaluation*, 12(10). [Available online: http://pareonline.net/getvn.asp?v=12&n=10].

Kalaian, S. A., & Shah, H. (2006, October). *Overview of Parametric and Non-parametric Statistical Methods for Analyzing Delphi Data*. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association. Columbus, Ohio.

Linstone, H. A., & Turoff, M. (1975). *The Delphi method: Techniques and applications.* Addison-Wesley: MA.

Privitera, G. J. (2012). Statistics for the Behavioral Sciences. California: Sage Publications, Inc.

Shah, H., & Kalaian, S. A. (2009). Which Parametric Statistical Method to Use For Analyzing Delphi Data? *Journal of Modern Applied Statistical Method (JMASM).* 8(1), 226-232.

Siegel, S. & Castellan, N. J. (1988). Nonparametric Statistics for the Behavioral Sciences. New York: McGraw Hill.

Yang, Y. N. (2003). *Testing the Stability of Experts' Opinion between Successive Rounds of Delphi Studies.* Paper presented at the Annual Meeting of the American Educational Research Association. Chicago, Illinois.

Yousuf, Muhammad Imran (2007). Using Experts' Opinions through Delphi Technique. *Practical Assessment Research & Evaluation*, 12(4). Available online: http://pareonline.net/getvn.asp?v=12&n=4.

Citation:

Kalaian, Sema & Rafa M. Kasim, (2012). Terminating Sequential Delphi Survey Data Collection. *Practical Assessment, Research & Evaluation*, 17(5). Available online: http://pareonline.net/getvn.asp?v=17&n=5

Corresponding Author:

Sema Kalaian, Ph.D.
Professor of Statistics and Research Design
College of Technology
School of Technology Studies
161 Sill Hall
Eastern Michigan University
Ypsilanti, MI 48197
SKalaian@emich.edu