ISAHP 2005, Honolulu, Hawaii, July 8-10, 2003

QUESTIONNAIRE DESIGN FOR SURVEY RESEARCH: EMPLOYING WEIGHTING METHOD

Yuji Sato
Graduate School of Policy Science, Mie Chukyo University
1846, Kubo, Matsusaka
Mie, 515-8511 Japan
ysatoh@mie-chukyo-u.ac.jp

Keywords: questionnaire design, Feeling Thermometer, weighting method

Summary: *In this paper, we propose an application of the Analytic Hierarchy Process to survey research to overcome the deficiency of traditional questionnaire methods. The procedure of the AHP results in not only the identification of the most important alternative but also the preference of all alternatives for each respondent. Therefore, by applying the AHP to survey questionnaire, eliciting respondent's opinion more precisely than by traditional methods may be expected. This study showed that the weights derived from the AHP improved R^2 in regression analysis, which had Feeling Scores as independent variables and respondent's political slant as dependent variable.*

1. Introduction

The design of questionnaires for survey research (e.g., public opinion polls) represents one of the most controversial issues among survey researchers in terms of accuracy in measuring respondents' perceptions. The use of multiple-choice questions, one of the more traditional methods, is highly suitable because respondents find them easy to answer and they allow researchers to easily identify the main concerns of the respondents. This method has two types of formats: one is Simple- multiple-choice; the other is Modified-multiple-choice. In the Simple-multiple-choice format, respondents must choose one from among the given alternatives. Simple-multiple-choice identifies only the most important alternative for each respondent, thus preventing the respondent from expressing his or her preference concerning a selected alternative over the others. Furthermore, no information regarding the relationship among the non-selected alternatives is derived. In the Modified-multiple-choice format, respondents are given the option of indicating their top-two (or more) alternatives. Since respondents are allowed to express their preferred alternatives, Modified-multiple-choice can be expected to be an effective way to make up for the lack of information incurred by Simple-multiple-choice. However, the difference in the degree of importance among the selected alternatives is not clarified, nor is the information concerning non-selected alternatives reflected in the results.

Let us consider the case asking respondents why they are non-partisan, for example. If the question is designed in the Simple-multiple-choice format, respondents must express their opinion by choosing one from among the given reasons. Respondents with a definite reason regarding the issue could choose one alternative without confusion if they found that it exactly represented their perception. This format could be expected to function quite well for these respondents. On the other hand, it might be that some respondents are non-partisan for no particular reason, while others are non-partisan for complex reasons. Multiple-choice would not be an effective format for these respondents.

To overcome the deficiencies of traditional methods, many ways of asking questions have been proposed and much discussion has been generated. Ronald Inglehart and Paul Abramson (1993) used ranking scale questions in the World Values Survey. This method asks respondents to rank all given alternatives in a question, from the most preferred to the least, thus allowing researchers to identify a respondent's preference order for all alternatives. However, the more alternatives a questionnaire offers, the more

difficult it is for the respondent to answer. Indeed, one respondent in a survey needed 25 minutes to complete the ranking of over 40 alternatives.

In contrast, a group of researchers at the University of Michigan developed rating scale questions called the Feeling Thermometer (FT), which has a scale from 0, the coldest feeling toward alternatives, to 100, the hottest, with 50 being neutral. In surveys, respondents express their perception by indicating their "temperature" to all alternatives given in a question. This method helps respondents precisely clarify their preference; however, consistency among responses to the alternatives is not always satisfactory.

One possible option for such survey research is to use the Analytic Hierarchy Process (AHP). One of the most popular methods for decision-making, the AHP was proposed by T. Saaty (1977, 1980). Since then, the AHP has been used by many individuals and groups in various fields because of its user-friendly interface for multi-criteria decision-making (Vargas, 1990). In the AHP process, data on decision-makers' judgments, called pair-wise comparisons, are aggregated, and the degree of importance of each alternative is quantified for each decision-maker. This procedure identifies not only the most important alternative but also the preference for all alternatives for each decision-maker. Therefore, by applying the AHP to survey research questionnaires, respondents' perceptions can be clarified more precisely than by traditional methods (Sato, 2003, 2004).

In this study, we compared the answers to two types of questions on a particular issue, each formatted in a different way, one using the FT and the other using the AHP. First, we focused on the difference of the aggregated ranking of alternatives across all respondents between the FT and the AHP. Since both rankings reflect the entire trend concerning each alternative of a population, they are likely to produce similar results. Second, we evaluate the effectiveness of each method in terms of its ability as independent variables in regression analyses. In the above-mentioned comparisons, a data set obtained from a 2004 survey on public opinion was employed. Details of these data sets are shown in Section 3.

2. Outline of the AHP

Saaty developed the AHP as one of the supporting systems for multi-criteria decision-making and as a tool for analyzing the decision-making process. The AHP has the subjective judgment of each decision-maker as input and the quantified weight of each alternative as output. Therefore, not only can objective issues be easily quantified but also the more subjective issues that do not have theoretical values. It has thus been widely used for decision-making (e.g., economic problems, policy evaluation and urban planning (Lootsma, 1989; Saaty, 1994)).

The AHP has three steps to the final output: stratifying the framework of decision-making, pair-wise comparison, and calculating the weight of each alternative. Let a_{ij} (i, j=1,…, n) denote the relative weight of alternative i (i for short) to j, and $a_{ji}=1/a_{ij}$, then the results of all pair-wise comparisons can be summarized as an n by n reciprocal matrix (a_{ij}) called a pair-wise comparison matrix, where $a_{ii}=1$ for all i=1,…, n.

In cases where a pair-wise comparison matrix satisfies transitivity for all pair-wise comparisons, i.e., for arbitrary i, j and k, $a_{ik} \cdot a_{kj}= a_{ij}$ (i, j, k =1,…, n) holds, the rank of such pair-wise comparison matrix equals 1, its Frobenius root is always equal to n, and the remaining eigenvalues are all 0 for any a_{ij} (Nikaido, 1968). Thus, the eigenvector corresponding to the Frobenius root is always non-negative, and each element of the eigenvector standardized by l_1-norm can be interpreted as the degree of importance of each alternative.

Decision-makers, however, sometimes make judgments inconsistently, in which case the aforementioned transitivity is not always satisfied. Then the Frobenius root of such an inconsistent pair-wise comparison matrix is always greater than n, and the difference between the root and n is equal to the sum of the remaining eigenvalues (Aupetit and Genest, 1993). Therefore, the smaller the difference is, the more consistent the judgment of the decision-maker would be (Murphy, 1993). In the AHP, the quotient of this difference over n-1 is defined as the Consistency Index (C.I. for short), which is the criterion for the

consistency of judgments across all pair-wise comparisons (Lootsma, 1991).

3. Questionnaire design

In this study, a data set was employed to compare the answers to two types of questions on a particular issue, each formatted in a different way: one is the FT and the other is the AHP. The data set was obtained from a 2004 survey on public opinion (the Survey, for short). The data set obtained from the survey may reflect each respondent's subjective judgment and thus served as an ideal actual sample in the comparison of the FT and the AHP.

As noted in the previous section, respondents need to respond to a series of redundant pair-wise comparisons in the AHP; therefore, we need to take into account the inconsistency of a pair-wise comparison matrix in analyzing elicited weights. In this study, however, we employed all samples, because "inconsistent respondent" is also an electorate being able to exercise the right to vote when an election would be carried out.

The Survey was carried out in March 2004, one month after Japan Self-Defense Forces was dispatched to Iraq. Respondents were lay citizens of Matsusaka city, Japan. The purpose of the survey was to identify "the political attitude of citizens when the Self-Defense Force was dispatched to a country in the state of warfare." The sample size was 30; each respondent's political attitude was elicited by interviewing him/her one by one.

The Survey included 33 questions. In Qs.9 and 17, respondents were asked their intention to vote for a part in the next election. The parties were: Liberal Democratic Party of Japan (LDP); Democratic Party of Japan (DPJ); New Komeito (NK); Japanese Communist Party (JCP); and Social Democratic Party (SDP). Q.9 was formatted in the FT and Q.17 in the AHP. The FT gave Feeling Score for each party, and the AHP gave the weight of each party.

We also asked respondents about their political slant in Q.13, and whether they support Prime Minister and the president of LDP, Jyunichiro Koizumi, in Q.33. Each of these issues was mutually related to the party electorates would vote for in the House of Councilors election. The outputs from these questions would serve as ideal dependent variables in a regression analysis.

Fig. 1 shows Q.9 from the survey formatted by the FT method, which requires respondents to assign an adequate number between 0 and 100 according to their intention to vote for each party. Fig. 2 illustrates Q.17 formatted in the AHP system, in which respondents are required to conduct pair-wise comparisons across all possible combinations of parties. Figs. 3 and 4 respectively show Qs.13 and 33, formatted by the Simple-multiple-choice method, which requires respondents to choose one from among the given alternatives.

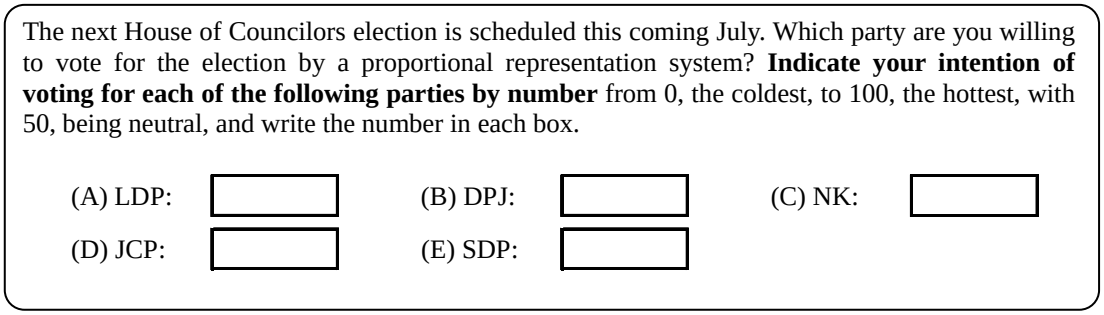

Fig. 1. Feeling Thermometer format (Question 9)

The next House of Councilors election is scheduled this coming July. Which party are you willing to vote for the election by a proportional representation system? **Compare each of the following pair of the parties, and mark the place along the segment.**

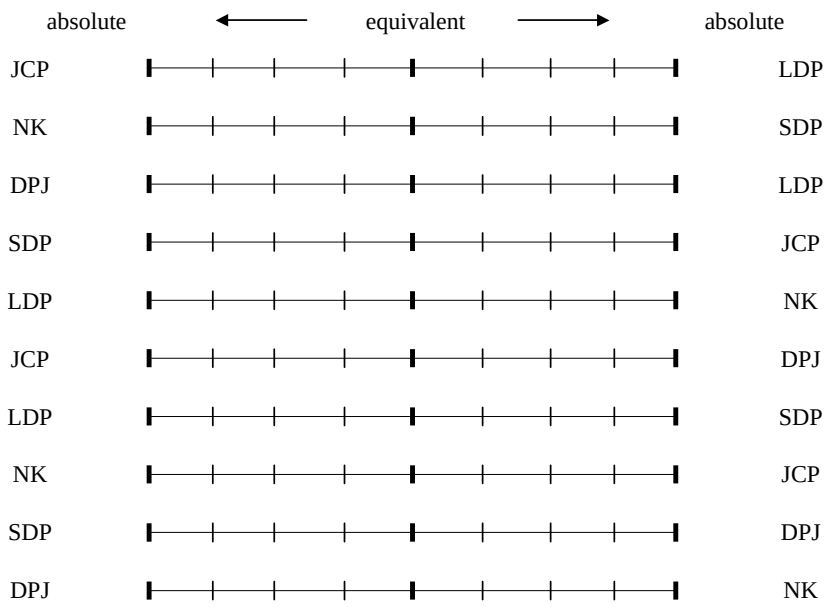

Fig. 2. AHP format (Question 17)

Which is your political slant, progressive or conservative? Choose 1 from among the 7 alternatives, and write the letter in the box.

(A) Progressive (B) Slightly progressive

(C) Neutral (D) Slightly conservative

(E) Conservative (F) DK

Fig. 3. Question 13

Do you support Prime Minister Koizumi? Choose 1 from among the 4 alternatives, and write the letter in the box.

(A) Support (B) Do not support

(C) DK

Fig. 4. Question 33

4. Results

In this section, based on the weight of each alternative derived from the AHP, we analyze responses to the FT-formatted questions of the Survey.

4.1 Comparison of aggregated rankings between the FT and the AHP

First, we focus on the difference of the aggregated ranking of alternatives between the FT and the AHP; in particular, the average of Feeling Score for each party obtained from the FT formatted question and that of the weight of each party derived from the AHP formatted question are compared. Since both rankings reflect the entire trend concerning each party of a population, they are likely to produce similar results.

Table 1 summarizes the results of two questions, Qs.9 and 17, from the Survey. The numbers in the first and the third rows are the average of Feeling Score and that of the AHP-based weight of each party aggregated across each response to the FT and the AHP, respectively. The last row represents the correlation coefficients between Feeling Score and the weight of each party among all respondents.

Table 1. Aggregated Feeling Score and weight of each party

Average	LDP	DPJ	NK	JCP	SDP
Feeling Score (FT)	53.6	66.6	19.2	14.4	14.2
Ranking	*2*	*1*	*3*	*4*	*5*
Weight (AHP)	0.34662	0.38969	0.08945	0.08007	0.09417
Ranking	*2*	*1*	*4*	*5*	*3*
Correlation Coefficient	0.81257	0.72848	0.77450	0.77443	0.75042

As shown in Table 1, Feeling Scores and the weights of the AHP imply that the most and the second most favored party among respondents are DPJ and LDP, respectively; both the FT and the AHP clearly identify the top 2 parties. On the other hand, the ranking for the remaining parties, NK, JCP and SDP, are different, even though the difference of Feeling Score or the weight of each party among those parties is quite small. Indeed, the correlation coefficients between Feeling Score and the weight of each party are high on some level. As a result, answers on the FT and on the AHP do not necessarily coincide, which yields a different aggregated ranking of alternatives. Both the FT and the AHP, however, produced similar results, and they overall seem to specify respondents' preference for parties in the choice of House of Councilors election.

4.2 Comparison of R^2s in regression analyses between the FT and the AHP

Next, we evaluate the effectiveness of each method in terms of its ability as independent variables in regression analysis; in particular, we formulate four regression models and compare R^2s. In the evaluations, as independent variables, we employ Feeling Score and the weight of each party except for those of the Social Democratic Party (SDP), because the party was minority group (The number of the member of House of Councilors was; LDP: 116, DPJ: 70, NK: 23, JCP: 20 and SDP: 5, as at the Survey was carried out. (March 2004)). Insofar as dependent variables, based on Quantification Theory I, we employ two variables; one is respondent's political slant (Q.13) in Regression Models 1 and 2, and the other is whether respondent supports Prime Minister Koizumi or not (Q.33) in Regression Models 3 and 4. As can be seen in Fig. 3, we offered six alternatives including "Don't Know" answer in Q.13. Among those alternatives, we set "Progressive" or "Slightly progressive" as 0, and "Conservative" or "Slightly conservative" as 1 in the Regression Analysis A. In the same way as for Regression Analysis A, we set "Support Koizumi" as 0, and "Do not support Koizumi" as 1, in Regression Analysis B. Thus, the actual regression models can be formulated as follows.

- Regression Analysis A
  - dependent variables; Political Slant
    "Progressive" or "Slightly progressive" = 0, "Conservative" or "Slightly conservative" = 1
  - regression equations;
    Political Slant = $a_1 + a_2*FS_{LDP} + a_3*FS_{DPJ} + a_4*FS_{NK} + a_5*FS_{JCP} + e$ (regression model 1)
    Political Slant = $b_1 + b_2*w_{LDP} + b_3*w_{DPJ} + b_4*w_{NK} + b_5*w_{JCP} + e'$ (regression model 2)

- Regression Analysis B
  - dependent variables; Support Koizumi or not
    "Support" = 0, "Do not support" = 1
  - regression equations;
    Koizumi = $c_1 + c_2*FS_{LDP} + c_3*FS_{DPJ} + c_4*FS_{NK} + c_5*FS_{JCP} + e$ (regression model 3)
    Koizumi = $d_1 + d_2*w_{LDP} + d_3*w_{DPJ} + d_4*w_{NK} + d_5*w_{JCP} + e'$ (regression model 4)

Table 2 summarizes the results obtained from Regression Analysis A. As can be seen in the table, the R^2s of regression models 1 or 2 are not large enough to predict respondents' political slant; respondent's intention for voting does not necessarily correlate with his/her political slant. By employing the weights derived from the AHP, however, R^2 was improved from 0.23755 to 0.42850. Furthermore, focusing on the p-values, the degree of precision of the regression model 2 was higher than that of the regression model 1.

Table 2. Results of Regression Analysis A (models 1 and 2)

Political slant	model 1 (Feeling Score)		model 2 (weight)	
	coefficient		coefficient	
constant	1.48488	**	1.35099	***
LDP	-1.58819	**	-0.00968	***
DPJ	-0.51939	*	-0.00175	*
NK	-0.45019	*	-0.00330	**
JCP	-1.62317	*	-0.00642	**
N = 25	R^2 = 0.23755	SE = 0.48460	R^2 = 0.42850	SE = 0.41955

p-value: 0.05 < p <= 0.1: *, 0.01 < p <= 0.05: **, p <= 0.01: ***

Table 3 summarizes the results obtained from Regression Analysis B. As shown in the table, R^2s are relatively larger that those of the Regression Analysis A; whether respondent supports Prime Minister Koizumi or not can be predicted on some level based on regression models 3 or 4. By employing the weight derived from the AHP, R^2 was again improved from 0.42293 to 0.51299. Furthermore, focusing on the p-values, the degree of precision of the regression model 4 was higher than that of the regression model 3.

Table 3. Results of Regression Analysis B (models 3 and 4)

Support Koizumi	model 3 (Feeling Score)		model 4 (weight)	
	coefficient		coefficient	
constant	-0.05558	*	-4.22262	**
LDP	0.00918	**	5.29983	***
DPJ	-0.00148	*	4.09808	**
NK	0.00481	**	6.46932	***
JCP	0.00125	*	9.04207	***
N = 17	R^2 = 0.42293	SE = 0.45130	R^2 = 0.51299	SE = 0.41459

p-value: 0.05 < p <= 0.1: *, 0.01 < p <= 0.05: **, p <= 0.01: ***

The FT method overall seem to be a good option for designing questionnaires because it assures respondents expressing their preferences for all alternatives, and it enhances the degree of freedom in answering questions for respondents. Therefore, many researchers, especially in political science field, have widely employed the FT in their survey questionnaires. Insofar as our survey is concerned, however, Feeling Scores may not be effective in predicting respondents' preferences. Either in Regression Analysis A or B, R^2s based on the FT were not large enough to predict respondents' preferences. On the other hand, the results of this section imply that the AHP could quantify respondents' preferences in terms of the distribution of the weight of each alternative; by employing weights of the AHP, R^2s in regression

analyses were improved from the use of Feeling Scores. Since Feeling Score measures the feeling for parties and the weight of the AHP measures the preference for parties, we cannot simply compare these R^2s and conclude that the weight is the better explainer than Feeling Score. The weights derived from the AHP, however, could be conjectured to function better as independent variables in regression analyses than Feeling Scores.

5. Concluding Remarks

Questionnaire design is one of the most controversial issues among survey researchers because how respondents are asked questions has a great effect on the results. One political scientist conducting a public opinion poll has remarked that different question formats yield different results, even though they are asking about the exact same content. Therefore, various ways of asking have been proposed and evaluated to date, from the perspective of appropriateness for representing each respondent's perception. In particular, the ranking method, the rating method and the multiple-choice method have often been compared and have generated much discussion. Among the aforementioned methods, the Feeling Thermometer method has been widely used because of its ease for respondents to answer and its ease in identifying for the researcher the respondents' concerns for all alternatives

In this study focused on survey research in which the questions asked of respondents involved issues (e.g., which party do respondents vote for), we examined the effectiveness of the Feeling Thermometer method by using the weight of alternatives elicited from the Analytic Hierarchy Process. The results were: (1) the Feeling Thermometer method and the AHP yielded a similar aggregated ranking of alternatives; (2) in regression analyses, Feeling Scores may not be effective in predicting respondents' preferences, while the weight derived from the Analytic Hierarchy Process could predict their preferences on some level. These results provide some evidence that the Feeling Thermometer method might provide erroneous information and thus would not be appropriate for measuring human perception in the use of questionnaire design of survey research.

References

Aupetit, B. and Genest, C. (1993), On Some Useful Properties of the Perron Eigenvalue of a Positive Reciprocal Matrix in the Context of the Analytic Hierarchy Process, *European Journal of Operational Research* 70, 263-268.

Inglehart, R. and Abramson, P. (1993), Values and Value Change of Five Continents, *paper presented at the 1993 Annual Meeting of the American Political Science Association*. *Washington D.C., September* 1-5.

Lootsma, F.A. (1989), Conflict Resolution via Pairwise Comparison of Concessions, *European Journal of Operational Research* 40, 109-116.

Lootsma, F.A. (1991), Scale Sensitivity and Rank Preservation in a Multiplicative Variant of the Analytic Hierarchy Process, *Delft University of Technology, Report of the Faculty of Technical Mathematics* 91(20).

Murphy, C.K. (1993), Limits on the Analytic Hierarchy Process from its Consistency Index, *European Journal of Operational Research* 65, 138-139.

Nikaido, F. (1968), *Convex Structures and Economic Theory*. Academic Press, London.

Saaty, T.L. (1977), A Scaling Method for Priorities in Hierarchical Structures, *Journal of Mathematical Psychology* 15, 234-281.

Saaty, T.L. (1980), *The Analytic Hierarchy Process*. McGraw-Hill.

Saaty, T.L., 1994. Highlights and Critical Points in the Theory and Application of the Analytic Hierarchy Process, *European Journal of Operational Research* 52, 426-447.

Sato, Y. (2003), Comparison between Ranking Method and the Analytic Hierarchy Process in Program Policy Analysis, *The Proceeding on the Seventh International Symposium on the Analytic Hierarchy Process 2003*, 429-439

Sato, Y. (2004), Comparison between Multiple-choice and Analytic Hierarchy Process: Measuring Human Perception, *International Transactions in Operational Research*, 11(1), 77-86.

Vargas, L.G. (1990), An Overview of the Analytic Hierarchy Process and its Applications, *European Journal of Operational Research* 48, 2-8.