A Study on Measuring the Change of the Response Results in Likert 5-Point Scale Measurement*

리커트 5점척도에서 자극에 의한 응답결과의 변화 측정에 관한 연구

Younghee Noh(노영희)**

ABSTRACT

This study examines how and which direction respondents who participated in 5-point Likert scale surveys change their initial responses when they are given an identical second survey after certain treatments. The research employs three identical questionnaires (first, second and third surveys) to analyze survey results based on group differences, kinds of treatment, survey purposes, and response change direction and the degree. This paper concludes that, first, it is significant that specialist groups do not change their initial responses compared to a general librarian group. Second, there are no differences by survey purpose; however, participants tend to change their initial responses by others' opinions rather than by previous use experiences. Third, participants who initially answered positively tend not to change their responses, and most participants who answered negatively change their initial responses in a positive direction. Fourth, when there are changes, participants change their initial responses by less than two points, and most of them change by one point. Finally, the hypothesis that middle responses change most and that participants who respond at both ends do not change their opinion was rejected by the finding that participants who answered on the negative end tend to change their initial responses in a positive direction.

초 록

본 연구에서는 리커트 5점 척도 설문에 참여했던 응답자들이 특정 자극에 의해 동일한 2차 설문이 주어졌을 때, 본인의 처음 응답결과를 어떻게, 어떤 방향으로 바꾸는지를 측정하고자 하였다. 이를 위해 3개 설문지의 1차, 2차 또는 3차에 걸친 재설문 결과들을 집단간 차이, 자극의 종류, 설문목적의 차이, 응답변화의 방향 및 변화정도의 관점에서 분석하였다. 본 연구를 통해 발견된 연구결과는 다음과 같다. 첫째, 전문가 집단과 일반 집단 중 전문가 집단이 상대적으로 의견을 바꾸지 않는 것으로 유의하게 검증되었다. 둘째, 설문목적에 따른 차이는 없는 것으로 조사되었고, 자극의 종류 중 시스템 사용경험보다는 다른 사람들의 의견이 정보로 주어졌을 때 더 많은 영향을 받는 것으로 나타났다. 셋째, 최초응답에서 긍정적인 답변에 응답했을수록 답변을 바꾸지 않는 것으로 나타났고 부정적인 답변자는 대부분 긍정적인 방향으로 의견을 바꾸는 것으로 유의하게 검증되었다. 넷째, 최초 응답을 바꾸는 경우 응답변화의 폭이 2단계 이상인 경우는 그다지 많지 않고 대부분 1단계 정도 전후로 바꾸는 경향이 있는 것으로 분석되었다. 마지막으로 중간응답자가 의견을 가장 많이 바꾸고 양극단의 경우 의견을 잘 바꾸지 않을 것이라는 예측을 하였으나 부정적인 답변자일수록 의견을 긍정적인 방향으로 바꾸는 것으로 분석되었다.

Keywords: likert-type scales, response rate, response changing, negative response, positive response, neutral point
리커트척도, 응답률, 응답변화, 부정적응답, 긍정적응답, 중간자

---

* This work was supported by the Konkuk University.
** Associate Professor in Department of Library & Information Science, Konkuk University (irs4u@kku.ac.kr)
■ 논문접수일자 : 2011년 8월 15일 ■ 최초심사일자 : 2011년 8월 16일 ■ 게재확정일자 : 2011년 9월 12일
■ 정보관리학회지, 28(3): 335-353, 2011. [http://dx.doi.org/10.3743/KOSIM.2011.28.3.335]

1. Introduction

Survey analysis is a dominant method in social science studies, and there are various methodologies developed to conduct effective surveys. For example, there are studies on survey group sampling, on the variations by response order and position when using Likert-type scale, on the meaning of the midpoint in Likert-type scale, and comparative studies of various survey delivery methods such as email, fax, post and web-based sources. However, there are few studies on different results of the survey when there are specific treatment or information intercepts during survey. In other words, scholars have analyzed user responses before and after using the system through survey method; however, they tend to present significant levels showing whether user attitude changed positively or there were no changes even after using the system. Accordingly, this study attempts to analyze the degree of changes based on the result of the first survey, and provides suggestions and discussions on how to evaluate changes of answers to each question. This study focuses on how much these treatments or information intercepts are influential to participants.

This research analyzes changes of survey results from previous studies conducted by author of this study, based on survey questionnaire method and Delphi technique (using questionnaires). Survey questionnaires used a 5-point Likert scale. In the second and third surveys, this research analyzes degree of change in responses compared to previous survey answers. Questionnaires for first, second and third surveys are identical. The research also examines degree of change by respondents who answered neutrally, and then compares it with degree of change by respondents who answered positively or negatively.

2. Literature Review

Academics have studied survey methods in various ways. In other words, there is research about the response rate of various survey delivery methods such as surveys through post mail, email and web-based sources, response rate of on-line surveys, response rate based on compensation, and response results by Likert-type survey properties such as verbal labels and numerical values. A comparative study on survey response rate among post, email and web-based source surveys showed that post surveys tend to have the highest and most prompt response rate (Cobanoglu, Warde and Moreo 2001). However, with the development of information technology communication (ITC), electronic methods have become more and more common in survey practice, and as a result, researchers began analyzing the response rate, speed and cost of surveys through email, fax and web-based sources (Cobanoglu et al. 2001). Cobanoglu and others. (2001) argued that surveys using web-based sources have the highest response rate and best cost-effectiveness. However, Sheeran and McMillan (1999) reported that web-based surveys are limited by such factors as survey population, research subject, and internet accessibility. At the same time, some researchers have attempted to use

a mixed-mode survey with email, web-based sources, telephone and fax to increase response rate (Beck 1996; Dillman and Tarnai 1998; Dilman 1999). However, in this case, it has been criticized that there are limitations such as difference of response context by survey delivery method and disparities in analysis (De Leeuw 1992; Schwarz, Hippler and Noelle-Neumann 1992; Dillman 1999).

In addition, academics have analyzed response-order effects by question order and position; survey response results vary as questions positions vary (Rugg and Cantril 1944; Payne 1951; Becker 1954; Belson 1966; Quinn and Belson 1969; Mueller 1970; Payne 1971; Brook and Uption 1974; Carp 1974; Schuman and Presser 1981). In regard to this, Crosnick and Alwin (1987) claimed that the response-order effects can be understood as primacy effects and recency effects. According to them, primacy effects occur when the placement of an item at the beginning of a list increases the likelihood that it will be selected. Recency effects are those that occur when placement of an item at the end of a list increases the likelihood that it will be chosen.

With regard to Likert scales surveys and their results, scholars have studied how presenting the scales can vary the results. In more detail, researchers such as Komorita (1963), Schwarz et al. (1991), Amoo and Friedman (2001), Armitage and Deeprose (2004), Nicholls et al. (2006), Sedlmeier (2006), and Dawes (2008) have studied scale range and its effect on response results. Weems and others (2003) and Stewart and others (2006) have analyzed how positioning negative questions and positive questions in the beginning of the survey affect results. Wyatt and Meyers (1987) attempted to analyze survey response to labeling, and Wildt and Mazis (1978) compared survey results of text labeling and numeric labeling. Recently, Hartley and Betts (2010) presented an overall analysis that presents a systematic understanding of the existing studies mentioned above.

Hartley and Betts (2010) compared four different layouts, as used in this paper: scales that are positioned numerically from ‘0’ to ‘10’, or from ‘10’ to ‘0’, and scales that are presented verbally from ‘clear’ to ‘unclear’, or ‘unclear’ to ‘clear’. As a result, more than 450 participants rated each of seven aspects of a structured abstract in a web-based study, with each one using only one of the four scale formats as listed above. The resulting data showed that the scale ‘Clear - 10 ... 0 - Unclear’ consistently lead to significantly higher ratings in all seven cases. Such findings have implications for the design of Likert- type scales and for the data that are gathered from them.

On the other hand, Garland (1991) assessed the mid-point or neutral point in Likert scales. According to him, social desirability bias can disappear when the mid-point or neutral point in Likert scales is removed, that is, measurement with only 4 points; the presence or absence of the mid-point can distort survey results. Similarly, scholars such as Matell and Jocoby (1972), Worcester and Burns (1975), Goldberg (1981), McFadden and Krug (1984), Armstrong (1987), Guy and Melissa (1997), and Cummins and Gullone (2000) have found that the presence or absence of a mid-point has an influence on survey results. Likewise, Kulas, Stachowski and Hynes (2008) conducted research on

the role of the middle response in Likert scales. They examined whether the middle response option in graphic rating scales indicates a moderate standing on a trait or an item, or a 'dumping ground' for unsure or non-applicable (N/A) responses. Their study identified middle response option dysfunction, and another study indicated respondents' use of the middle response as an N/A proxy even under the implicit 'skip if you do not know' instructional sets (Kulas et al. 2008). Although middle response category 'misuse' did not adversely affect reliability and validity in these studies, it is recommended that assessment developers (especially in on-line administration contexts) regularly include an N/A response option when administering graphic rating scales. Another study conducted by Jacoby and Matell (1971) that insisted three-point Likert scales are sufficient for surveys.

Related to this research paper, an analysis by Bignami-Van (2003) on the consistency of response by survey participants or interviewees showed the nature of the participants' individual inconsistency and its implication by using re-interview data collected during a household survey. The results showed an overall consistency among participants and little influence from inconsistent participants. Accordingly, there are re-interview studies on the reliability and stability of questions about knowledge, attitude and practice of contraception (Poti, Chakraborti, and Malaker 1962; Mauldin 1965; Green 1969; Stoeckel and Choudhury 1969; Coombs 1977). All of these studies suggested that response errors tend to be affected by characteristics of the participants, the context of the questionnaires, and the psychological nature of the participants, and their level of willingness to participate. Although there have been various studies on Likert-type surveys, there is little literature to explain whether survey results are influenced by participants' experience with the survey system or by others' opinions. Especially in Likert-type measurement surveys, there are few findings on how the responses in the middle or at both ends of the scale vary.

3. Research Questions

This paper examines the degree of change in responses across first, second and third administration of the same set of questionnaires, which used a 5-point Likert scale for measurement. It focuses on how neutral middle responses change and compares the response changes of participants at both positive and negative ends of the scale.

This study was conducted based on following research questions:

Question 1: In 5-point Likert scale measurement, do participants responding point 3, the neutral point, change their initial responses most frequently and are at the highest degree of change after treatment?

Question 2: In 5-point Likert scale measurement, do participants responding point 1 (negative) or point 5 (positive), barely change their initial responses even with any kind of treatment?

Question 3: If participants change their initial re-

sponses, how much is the degree of change?

Question 4: If participants change their initial responses, are they in the positive direction (closer to point 5) or negative direction (closer to point 1), or are they at the neutral point?

Question 5: Can the survey purpose or kind of treatment be classified as influential factors?

Question 6: Are there differences among the three survey groups?

4. Research Design and Methodology

4.1 Research Design

This study analyzed how initial responses of participants in surveys change after treatment based on existing survey results. Surveys and response results used in this research are the ones already used in previous studies by the author of this research project; the target group was librarians. The Delphi technique was used in Survey 1 (Noh 2010a) and Survey 2 (Noh 2010b), and an on-line survey method was used in Survey 3 (Noh 2010c). LIS specialists including librarians in institutions were questioned in Surveys 1 and 2, and in Survey 3 institutional librarians were questioned. Survey 1 was conducted to develop indicators for the evaluation of academic libraries, and Surveys 2 and 3 were carried out to develop metadata items.

In this paper, two different treatments were used: the average result of response from other participants and experience with system use. The treatment used in Surveys 1 and 2 were overall analysis results by question based on the average and median of the responses from the previous survey. The level of experience with system use by respondents was used as a treatment in Survey 3(see Table 1). All of these three surveys used on 5-point Likert scales, and in Surveys 1 and 2, identical questionnaires were distributed and collected three times; however, in Survey 3 identical questionnaires were distributed twice because the response rate was too low for experience with system use by respondents. Especially in Survey 3, first administration was conducted among participants who did not have experience with system use, and second administration was carried out after making these participants gained experience with the system.

〈Table 1〉 Research Data and Type

	Survey method	Target group	Context	Group title
Survey1	Delphi Method	LIS specialists including librarians	Library evaluation indicators development	Specialist group I
Survey2	Delphi Method	LIS specialists including librarians	Library metadata development	Specialist group II
Survey3	On-line Survey	Librarians by institution	Metadata item development for reference	General librarian group

Change of response by each question was analyzed to measure the degree of change in responses by survey participants. In particular, this study focused on participants who initially responded at the neutral point or at both positive and negative end points.

4.2 Methodology

This study employed a nonparametric test instead of using a parametric method to ensure normality because all variables in this study are measured based on an ordinal Likert scale. In other words, the influence of the first survey result as a treatment to the change in the second survey response was tested by the Johnkheere-Terpstra test.

The Johnkheere-Terpstra test expands the one-sided test of Wilcoxon rank sum test to K number of samples. The Johnkheere-Terpstra test is appropriate when K number of the population is classified by an ordinal group because it is a nonparametric method of testing if the location parameter of K number of groups increases gradually by order of groups. The statistic of the Johnkheere-Terpstra test is as follows:

$$J=\sum_{u=1}^{k-1}\sum_{v=u+1}^{k} U_{uv}, \text{ where } U_{uv}=\sum_{i=1}^{n_u}\sum_{j=1}^{n_v} I(X_{iu}<X_{jv})$$

Here, $I(a<b)$ shows 1 when $a<b$, or it is 0, which means an indicator function. n_u indicates number of samples in u^{th} group, and X_{iu} means i^{th} observation in u^{th} group. Rejecting null hypothesis or not is decided by comparing the J statistic of samples with the exact distribution or the approximate distribution.

5. Results

5.1 Response Rate

Two Delphi surveys and one general survey were conducted in this research. For the Delphi surveys, participants were paid 100 US dollars per person, and for the general survey 10 US dollars was given to each respondent. The response rate for the two Delphi surveys was 100 percent, and general survey had a 26.15 percent response rate. Reasons for the different response rates could be: 1) the influence of different amounts of compensation (Fox et al. 1988; James and Bolstein 1990; Yammarino et al. 1991; Church 1993; Everett, Price, Bedell, and Telljohann 1997) or 2) social desirability, acquiescence, question order effects, and primary or recency effects, according to Dillman (1999). In addition, they could be due to the degree of difficulty of the questionnaires. In this study, the response rate of Delphi surveys are higher than the general survey despite the fact that the number of questions is much greater in the Delphi surveys than in the general survey. Therefore, the discrepancy in response rates can be explained by the difference in compensation amounts.

5.2 Comparison of Specialists and General Librarians

This sub-section compares each group's average, standard deviation, and two-way contingency table, and also tests significances by using the Johnkheere-Terpstra test. The study used various analyses to answer to the research questions. In addition, this paper compared study specialist groups and a general librarian group. A comparison between specialist groups is carried out based on the same treatment but with different purposes, and another comparison between a specialist group and a general group is based on the same purpose but with different treatment.

5.2.1 Response Change by Group

The response average of specialists gradually increases and their standard deviation decreases by each survey compared to those of general librarians. In other words, librarians' average decreases and their standard deviation increases (see Table 2).

5.2.2 Response Change Treatment

In all groups, response average increases for participants who chose points 1, 2 and 3 in the analysis of response average changes by Likert scale. On the other hand, participants who chose points 4 and 5 showed a decrease in their response average. This result was more significant in the general librarian group than in specialist groups. In the case of specialist groups, the standard deviation tended to decrease when their prior answers were in the positive direction. Participants who chose point 4 showed the highest rate of response average change amongst all (see Table 3).

〈Table 2〉 Response Average and Standard Deviation by Group and by Survey

Group	N	First		Second		Third	
		Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Specialist group	932	3.8287	1.0457	3.8676	0.9874	3.8870	0.9588
General librarian group	578	4.0796	0.7612	3.9533	0.8848	NA	NA

〈Table 3〉 Response Average by Likert-5 Measure Score

Prior response	All groups			Specialist group			Specialist group I			Specialist group II			General librarian group		
	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation
1	33	1.5758	0.7084	32	1.5625	0.7156	5	2.0000	0.0000	27	1.4815	0.7530	1	2.0000	.
2	167	2.4551	0.6650	154	2.3831	0.5852	49	2.4286	0.6124	105	2.3619	0.5739	13	3.3077	0.9473
3	569	3.1810	0.5939	468	3.1239	0.5252	222	3.0721	0.4701	246	3.1707	0.5672	101	3.4455	0.7934
4	900	3.9611	0.5907	613	4.0016	0.4259	272	3.9559	0.3816	341	4.0381	0.4554	287	3.8746	0.8352
5	766	4.7624	0.5298	590	4.8610	0.3879	166	4.8855	0.3553	424	4.8514	0.3999	176	4.4318	0.7604

5.3 Two-way Contingency Table Analysis

It is difficult to accurately analyze response change by Likert scale point and its direction with response average and standard deviation by survey. Thus, two-way contingency tables of all groups, specialist groups and general librarian group was created. Tables from 4 to 10 describe response change in two-way contingency tables. In detail, Table 4 shows the analysis of response change of all groups, Table 5 for special groups, Table 8 for a specialist group that participated in the survey of library evaluation indicator (Specialist Group I), Table 9 for another specialist group that responded to the survey of metadata development (Specialist Group II), and Table 10 for a general librarian group that participated in the survey of metadata development. Table 6 and Table 7 explain how each specialist group changed responses across the first, second and third surveys.

In particular, Table 4 shows the results of second survey clearly. For instance, 51.52 percent of participants who selected point 1 in the first survey selected point 1 again in the second survey, 42.42 percent of participants changed their responses to point 2, and 3.03 percent changed to point 3. Overall, the highest percentage of participants chose the same point in both surveys, and the lowest percentage changed more than 2 points. That is, when participants changed their prior responses, mostly they changed by 1 point. As Table 4 illustrates, it is most significant with less than 0.05 based on the Johnkheere-Terpstra test of non-change responses closer to point 5 (positive).

Table 5 reflects an analysis of specialist groups in a two-way contingency table. Here, responses that changed from first survey to second and from second

〈Table 4〉 Two-way Contingency Table Analysis (All Groups)

Initial response	Re-response (Row Ratio %)					Total (Ratio %)
	1	2	3	4	5	
1	17 (51.52)	14 (42.42)	1 (3.03)	1 (3.03)	0 (0)	33 (1.36)
2	1 (5.26)	102 (61.08)	53 (8.39)	9 (1.06)	2 (0.26)	167 (6.86)
3	1 (5.26)	32 (19.16)	422 (66.77)	91 (10.72)	23 (2.99)	569 (23.37)
4	0 (0)	13 (7.78)	136 (21.52)	624 (73.5)	127 (16.54)	900 (36.96)
5	0 (0)	6 (3.59)	20 (3.16)	124 (14.61)	616 (80.21)	766 (31.46)
Total (Ratio %)	19 (0.78)	167 (6.86)	632 (25.95)	849 (34.87)	768 (31.54)	2435
Johnkheere-Terpstra Test		Statistic			1816879	
		Z			42.4967	
		p-value			〈.0001	

〈Table 5〉 Two-way Contingency Table Analysis (All Specialist Groups)

Initial response	Re-response (Row Ratio %)					Total (Ratio %)
	1	2	3	4	5	
1	17 (53.13)	13 (40.63)	1 (3.13)	1 (3.13)	0 (0)	32 (1.72)
2	1 (0.65)	100 (64.94)	46 (29.87)	7 (4.55)	0 (0)	154 (8.29)
3	1 (0.21)	25 (5.34)	368 (78.63)	63 (13.46)	11 (2.35)	468 (25.2)
4	0 (0)	3 (0.49)	46 (7.5)	511 (83.36)	53 (8.65)	613 (33.01)
5	0 (0)	1 (0.17)	6 (1.02)	67 (11.36)	516 (87.46)	590 (590)
Total (Ratio %)	19 (1.02)	142 (7.65)	467 (25.15)	649 (34.95)	580 (31.23)	1857
Johnkheere-Terpstra Test			Statistic		1138852	
			Z		42.5838	
			p-value		〈.0001	

survey to third were all considered as one single case. As a result, there were few changes when the initial responses were in the highest positive direction, and it was significant based on the Johnkheere-Terpstra test (see Table 5).

Table 6 and Table 7 demonstrate response change

〈Table 6〉 Two-way Contingency Table Analysis (Specialist Groups: First to Second Survey)

Initial response	Re-response (Row Ratio %)					Total (Ratio %)
	1	2	3	4	5	
1	17 (53.13)	13 (40.63)	1 (3.13)	1 (3.13)	0 (0)	32 (1.72)
2	1 (0.65)	100 (64.94)	46 (29.87)	7 (4.55)	0 (0)	154 (8.29)
3	1 (0.21)	25 (5.34)	368 (78.63)	63 (13.46)	11 (2.35)	468 (25.2)
4	0 (0)	3 (0.49)	46 (7.5)	511 (83.36)	53 (8.65)	613 (33.01)
5	0 (0)	1 (0.17)	6 (1.02)	67 (11.36)	516 (87.46)	590 (590)
Total (Ratio %)	19 (1.02)	142 (7.65)	467 (25.15)	649 (34.95)	580 (31.23)	1857
Johnkheere-Terpstra Test			Statistic		281872.5	
			Z		29.1946	
			p-value		〈.0001	

〈Table 7〉 Two-way Contingency Table Analysis (Specialist Groups: Second to Third Survey)

Initial response	Re-response (Row Ratio %)					Total
	1	2	3	4	5	(Ratio %)
1	6	5	0	0	0	11
	(54.55)	(45.45)	(0)	(0)	(0)	(1.18)
2	1	52	20	2	0	75
	(1.33)	(69.33)	(26.67)	(2.67)	(0)	(8.07)
3	1	9	190	30	1	231
	(0.43)	(3.9)	(82.25)	(12.99)	(0.43)	(24.87)
4	0	1	23	271	26	321
	(0)	(0.31)	(7.17)	(84.42)	(8.1)	(34.55)
5	0	0	3	26	262	291
	(0)	(0)	(1.03)	(8.93)	(90.03)	(31.32)
Total	8	67	236	329	289	929
(Ratio %)	(0.86)	(7.21)	(25.4)	(35.41)	(31.11)	
Johnkheere-Terpstra Test			Statistic		287533.5	
			Z		31.05	
			p-value		<.0001	

rates by survey, from the first survey to the second and from the second survey to the third; the results of each survey have similar trends as above. The results are all significant with the Johnkheere-Terpstra test (see Table 6 and Table 7).

Table 8 and Table 9 show the analysis results of

〈Table 8〉 Two-way Contingency Table Analysis (Specialist Group I)

Initial response	Re-response (Row Ratio %)					Total
	1	2	3	4	5	(Ratio %)
1	0	5	0	0	0	5
	(0)	(100)	(0)	(0)	(0)	(0.7)
2	0	31	15	3	0	49
	(0)	(63.27)	(30.61)	(6.12)	(0)	(6.86)
3	0	14	181	24	3	222
	(0)	(6.31)	(81.53)	(10.81)	(1.35)	(31.09)
4	0	0	26	232	14	272
	(0)	(0)	(9.56)	(85.29)	(5.15)	(38.1)
5	0	0	2	15	149	166
	(0)	(0)	(1.2)	(9.04)	(89.76)	(23.25)
Total	0	50	224	274	166	714
(Ratio %)	(0)	(7)	(31.37)	(38.38)	(23.25)	
Johnkheere-Terpstra Test			Statistic		164340	
			Z		26.2378	
			One side Pr 〉 Z		<.0001	
			Both sides Pr 〉 |Z|		<.0001	

〈Table 9〉 2 Two-way Contingency Table Analysis (Specialist Group II)

Initial response	Re-response (Row Ratio %)					Total
	1	2	3	4	5	(Ratio %)
1	17	8	1	1	0	27
	(62.96)	(29.63)	(3.7)	(3.7)	(0)	(2.36)
2	1	69	31	4	0	105
	(0.95)	(65.71)	(29.52)	(3.81)	(0)	(9.19)
3	1	11	187	39	8	246
	(0.41)	(4.47)	(76.02)	(15.85)	(3.25)	(21.52)
4	0	3	20	279	39	341
	(0)	(0.88)	(5.87)	(81.82)	(11.44)	(29.83)
5	0	1	4	52	367	424
	(0)	(0.24)	(0.94)	(12.26)	(86.56)	(37.1)
합계	19	92	243	375	414	1143
(비율%)	(1.66)	(8.05)	(21.26)	(32.81)	(36.22)	

Johnkheere-Terpstra Test	Statistic	428822
	Z	33.1597
	One side Pr 〉 Z	〈.0001
	Both sides Pr 〉 |Z|	〈.0001

two specialist groups: specialist group I and specialist group II. Results were also similar to others, and they are significant according to the Johnkheere-Terpstra test. However, in the case of specialist group I, all of the five participants who initially answered point 1 changed their choices to point 2.

Table 10 shows somewhat different results from the others: in the general librarian group, there is

〈Table 10〉 Two-way Contingency Table Analysis (General Librarian Group)

Initial response	Re-response (Row Ratio %)					Total
	1	2	3	4	5	(Ratio %)
1	0	1	0	0	0	1
	(0)	(100)	(0)	(0)	(0)	(0.17)
2	0	2	7	2	2	13
	(0)	(15.38)	(53.85)	(15.38)	(15.38)	(2.25)
3	0	7	54	28	12	101
	(0)	(6.93)	(53.47)	(27.72)	(11.88)	(17.47)
4	0	10	90	113	74	287
	(0)	(3.48)	(31.36)	(39.37)	(25.78)	(49.65)
5	0	5	14	57	100	176
	(0)	(2.84)	(7.95)	(32.39)	(56.82)	(30.45)
Total	0	25	165	200	188	578
	(0)	(4.33)	(28.55)	(34.6)	(32.53)	

Johnkheere-Terpstra Test	Statistic	73252
	Z	10.2214
	p-value	〈.0001

a tendency to change the initial response. In other words, general librarians show a low non-change response rate. Response changes from point 1 to point 2 and changes from point 2 to point 3 show the highest change rates. However, even though the differences are not great, but it is significantly verified according to the Johnkheere-Terpstra test (see Table 10). Especially in the general librarian group, observations of respondents who answered with negative points were very low compared to other groups.

5.4 Response Change Range and Degree

5.4.1 Response Change Range by Group

Comparisons of response change ranges by group (re-response - prior response), the general librarian group was more likely to change from the initial response than were the specialist groups, and the standard deviation was also higher in the general librarian group. In the case of specialist groups, response change range decreased when responses where from point 1 to point 3, standard deviation also decreased (see Table 11).

Both the range of response changes by Likert scale point and standard deviation decreased when initial responses were at positive points. Here, the range was measured as an average of prior response changes. This result was same for all groups, except respondents who chose point 4 in the general librarian group (see Table 12).

For specialist groups, overall trends are identical in each survey, and there is a higher change range between the first and second specialist groups (see Table 13).

5.4.2 Response Change Degree by Group

Tables 14 describe the degree of response change after initial response. For example, Table 14 deals with the degree of response change by initial response

〈Table 11〉 Average Response Change by Group

Group	From first to second survey			From second to third survey		
	N	Average	Standard Deviation	N	Average	Standard Deviation
Specialist	932	0.2392241	0.4903516	932	0.1679225	0.3963792
General librarian	578	0.6245675	0.6628691	NA	NA	NA

〈Table 12〉 Average Response Change by Prior Response

Prior response	All groups			Specialist group			Specialist group I			Specialist group II			General librarian group		
	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation
1	33	0.5758	0.7084	32	0.5625	0.7156	5	1.0000	0.0000	27	0.4815	0.7530	1	1.0000	.
2	167	0.4671	0.6566	154	0.3961	0.5764	49	0.4286	0.6124	105	0.3810	0.5613	13	1.3077	0.9473
3	569	0.3005	0.5432	468	0.2393	0.4836	222	0.1982	0.4322	246	0.2764	0.5237	101	0.5842	0.6967
4	900	0.3211	0.4972	613	0.1713	0.3899	272	0.1471	0.3548	341	0.1906	0.4152	287	0.6411	0.5485
5	766	0.2376	0.5298	590	0.1390	0.3879	166	0.1145	0.3553	424	0.1486	0.3999	176	0.5682	0.7604

〈Table 13〉 Specialist Group Response Change

Prior response	All Specialist Groups			Specialist Group First Survey			Specialist Group Second Survey		
	N	Average	Standard Deviation	N	Average	Standard Deviation	N	Average	Standard Deviation
1	32	0.5625	0.7156	21	0.6190	0.8047	11	0.4545	0.5222
2	154	0.3961	0.5764	79	0.4557	0.6160	75	0.3333	0.5285
3	468	0.2393	0.4836	237	0.2911	0.5404	231	0.1861	0.4118
4	613	0.1713	0.3899	292	0.1849	0.4062	321	0.1589	0.3746
5	590	0.1390	0.3879	299	0.1672	0.4243	291	0.1100	0.3448

〈Table 14〉 Response Change Degree by Initial Response (All Groups)

Initial response	Response change (row ratio %)			Total (Ratio %)
	0	1	More than 2	
1	17 (51.52)	14 (42.42)	2 (6.06)	33 (1.36)
2	102 (61.08)	54 (32.34)	11 (6.59)	167 (6.86)
3	422 (74.17)	123 (21.62)	24 (4.22)	569 (23.37)
4	624 (69.33)	263 (29.22)	13 (1.44)	900 (36.96)
5	616 (80.42)	124 (16.19)	26 (3.39)	766 (31.46)
Total (Ratio %)	1781 (73.14)	578 (23.74)	76 (3.12)	2435
Johnkheere-Terpstra Test			Statistic	971207
			Z	-5.0138
			p-value	〈.0001

for each point in the Likert scale for all groups. According to this table, response change rate of participants who did not change their initial answers to point 1 was 51.52 percent, the rate of participants who changed their initial response at 1 point level was 42.42 percent, and the rate of participants who changed their initial responses at more than 2 points level was 6.06 percent. Overall, if participants answered positively (closer to point 5), they tended not to change their initial responses. Significance levels of the results for all cases were less than 0.05 percent by the Johnskeers-Terpstra test.

6. Discussion

This study attempted to analyze the degree and the direction of changes of survey participants' initial responses by respondents' system use experiences or others' opinions. Target groups consisted of two

specialist groups and one general librarian group. In the case of the specialist groups, this study employed the Delphi method, and all three surveys were conducted based on 5-point Likert scale measurement. During the first survey administration, one of two specialist groups was polled to develop library evaluation indicators and the other was polled to develop metadata. The second survey's purpose was the same as in the survey for the general librarian group.

For the specialist groups, identical questionnaires were provided for all three surveys. Before conducting the second survey, the result of the first survey analysis (average, median and each participant's opinion in the prior survey) was shown, and before the third survey, the result of the second survey was given to respondents. In this light, the treatment here was providing an overall analysis result, that is, opinions of other people. On the other hand, there were two surveys for the general librarian group. For the first survey, general librarians participated in the survey without system use experience, and for the second survey, they all had experience using the system, thus, the treatment was system use. Based on that, this research conducted analysis on the differences among groups and differences by treatment and by survey purpose. Especially, this paper focuses on how, in what direction, and to what degree participants' responses changed when they answered closer to the positive end of the scale (point 5 on the Likert scale) or closer to the negative end (point 1 on Likert scale).

First, the analysis result on the differences between the specialist groups and the general librarian group showed that specialist groups did not change their initial responses compared with the general librarian group. Also, as the survey was repeated, response change average increased and deviation decreased in specialist groups; however, for the general librarian group, the average increased and deviation decreased. In this case, it was not clear if the differences between two groups were due to the treatment or due to the different characteristics of each group, and this remains a limitation of this research. Second, the paper compared specialist group I and specialist group II to analyze differences by survey purpose. Based on the Johnkkeere-Terpstra test, there was not a significant difference between groups. Accordingly, response change was not affected by the survey purpose, but by other factors. Third, specialist group II and general librarian group were compared to analyze differences by treatment. In the case of specialist group II, response change average increased compared to the prior survey; however, for the general librarian group, response change average increased when participants chose points 1, 2 and 3 in their prior survey, and it decreased when respondents chose points 4 and 5. In this sense, others' opinion was influential by making participants' responses change in a positive direction. At the same time, system use experience had an effect by making participants change their initial negative responses dramatically in the positive direction and change their positive responses in a negative direction slightly. Fourth, generally, participants changed their initial responses in a positive direction in the second and third surveys when they experienced certain treatments, and the result was significant based on the Johnkkeere-Terpstra test. Finally, most participants

chose the same responses as in the initial survey, and the degree of response change was higher when the original response was closer to point 5. In addition, when they changed their responses, it was at the 1-point level; they rarely changed their responses more than 2 points. In other words, respondents who received certain treatments, changed their opinions in a positive direction. Besides, the degree of response change was higher in the general librarian group than in the specialist groups, and it was significant according to the Johnkheere-Terpstra test. Additionally, participants did not change their initial responses if they answered in the positive part of the scale.

It can be said that the purpose of the surveys for the specialist groups were to gather opinions by presenting others' opinions, and the surveys for general librarian group were designed to develop an optimal system. As it was intended, those in the specialist groups changed their initial responses based on others' opinions, and this was shown as standard deviation decreased. That is, the Delphi survey was very appropriate for gathering opinions of the participants.

Before conducting this study, it was hypothesized that participants who answered at both end points would not change their responses, and on the other hand, the ones who were positioned in the middle would change their responses according to certain treatments. However, this hypothesis was rejected by the analysis of average, deviation, and group comparisons. Rather, the result showed that participants who answered negatively tended to change their initial responses in a positive direction.

Even though this study did not clearly define various other factors such as psychological changes, survey interest changes, and survey background knowledge changes, this research has important implications for further studies.

One limitation of this paper was in data collection. The data collected was not sufficient to clearly determine treatment differences and group differences. That is, it was not possible to conduct second and third surveys that can explain all of the factors influencing response change, such as survey purpose, kinds of treatment, and differences between groups. Thus, this study employed existing surveys and results that were similar to this research purpose. Thus, it is not possible to determine which factor influenced response change. Nevertheless, there were group differences in responses to the second and third surveys, and it was significant that response change was influenced by treatment. However, it is true that more data in various dimensions will be necessary for further studies.

7. Conclusion

Most existing literature on the 5-point Likert scale measurement has been related to survey methods, survey response rate by survey distribution method, middle responses, and the consistency of respondents. In this paper, response change was examined by using 5-point Likert scale measurement, and this was relatively unique compared to existing studies. This research intended to measure how initial responses of middle point or both end points on the Likert 5-point scale were changed by certain information or treat-

ment, and analyzed the direction and degree of such changes. Moreover, it studied variables such as differences in the group characteristics of specialist groups and the general librarian group and survey purposes, and their influences on response changes.

Consequently, this study has found that there are group differences between specialist groups and the general librarian group, and that the specialist groups had a tendency not to change their initial responses compared to the general librarian group. At the same time, this study has shown that respondents who were in negative direction on the first survey tended to change their responses in a positive direction significantly, and participants who initially answered with positive point responses did not change their initial responses. Middle responses did not change much as well. In terms of the degree of response change, participants changed their initial responses by 1 point when there changes.

References

Amoo, T. and H. H. Friedman. 2001. "Do numeric values influence subjects' responses to rating scales?" *Journal of International Marketing and Marketing Research,* 26: 41-46.

Armitage, C. and C. Deeprose. 2004. "Changing student evaluations by means of the numeric values of rating scales." *Psychology Learning and Teaching,* 3: 122-125.

Armstong, R. L. 1987. "The midpoint on a five-point Likert-type scale." *Perceptual and Motor Skills,* 64: 359-362.

Beck, J. 1996. Career anchors, organizational commitment, and job plateaus: An analysis of hotel executive operating committee members' career development. (Doctoral Dissertation, Purdue University, 1996). *Dissertation Abstracts International, 58(03).* (Publication Number: AAT 9725508).

Becker, S. L. 1954. "Why an order effect." *Public Opinion Quarterly,* 18: 271-278.

Belson, W. A. 1966. "The effects of reversing the presentation order of verbal rating scales." *Journal of Advertising Research,* 6: 30-376.

Bignami-Van Assche, S. G., G. Reniers, and A. A. Weinreb. 2003. "An assessment of the KDICP and MDICP data quality." *Demographic Research,* S1(2): 31-76.

Brook, D. and G. J. Upton. 1974. "Primacy effects in single trial free recall." *Applied Statistics,* 23: 414-419.

Carp, F. M. 1974. "Position effects on interview response." *Journal of Gerontology,* 29: 581-587.

Church, A. H. 1993. "Estimating the effect of incentive on mail survey response rates: A meta-analysis." *Public Opinion Quarterly,* 57: 62-79.

Cobanoglu, C. B. Warde, and P. J. Moreo. 2001. "A

comparison of mail, fax and web-based survey methods." *International Journal of Market Research,* 43(4): 441-452.

Cummins, R. A. and E. Gullone. 2000. "Why we should not use 5-point Likert scales: The case for subjective quality of life measurement." *Proceedings, Second International Conference on Quality of Life in Cities,* Singapore: National University of Singapore, 74-93.

Dawes, J. 2008. "Do data characteristics change according to the number of points used? An experiment using 5-point, 7-point and 10-point scales." *International Journal of Market Research,* 50: 61-77.

De Leeuw, E. D. 1992. *Data Quality in Mail, Telephone, and Face-to-face Surveys.* Amsterdam: TT Publications.

Dillman, D. A. 1999. *Mail and internet surveys: The tailored design method, (2nd Ed.).* New York: John Wiley and Sons.

Dillman, D. A. and J. Tarnai. 1988. "Administrative issues in mixed mode surveys." In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, II & J. Wakesberg (Eds.), *Telephone survey methodology* (pp. 509-528). New York: Wiley.

Everett, S. A., J. H. Price, A. Bedell, and S. K. Telljohann. 1997. "The effect of a monetary incentive in increasing the return rate of a survey to family physicians." *Eval Health Professions,* 20: 207-4.

Fox, R. J., M. R. Crask, and J. Kim. 1988. "Mail survey response rate: A meta-analysis of selected techniques for introducing response." *Public Opinion Quarterly,* 52: 467-491.

Friedman, H. H., Y. Wilamowsky, and L. W. Friedman. 1981. "A comparison of balanced and unbalanced rating scales." *The Mid-Atlantic Journal of Business,* 19(2): 1-7.

Garland, R. 1991. "The mid-point on a rating scale: Is it desirable?" *Marketing Bulletin,* 2: 66-70, Research Note 3.

Goldberg, L. R. 1981. "Unconfounding situational attributions from uncertain, neutral, and ambiguous ones: A psychometric analysis of descriptions of oneself and various types of others." *Journal of Personality and Social Psychology,* 41(3): 517-552.

Green, B. F. 1969. "Attitude measurement." In G. Lindzey (Ed.), *Handbook of Social Psychology.* Cambridge, Mass.: Addison-Wesley.

Guy, R. F. and M. Norvell. 1977. "The neutral point on a Likert scale." *Journal of Psychology: Interdisciplinary and Applied,* 95(2): 199-204.

Hartley, J. and L. R. Betts. 2010. "Four layouts and a finding: The effects of changes in the order of the verbal labels and numerical values on Likert-type scales." *International Journal of Social Research Methodology,* 13(1): 17-27.

Jacoby, J. and M. S. Matell. 1971. "Three-point Likert scales are good enough." *Journal of Marketing Research,* 8: 495-500.

James, J. and R. Bolstein. 1992. "Large monetary incentives and their effect on mail survey response rates." *Public Opinion Quarterly,* 56: 442-53.

Komorita, S. S. 1963. “Attitude content, intensity and the neutral point on a Likert scale.” *Journal of Social Psychology,* 61: 327-334.

Krosnick, J. A. and D. F. Alwin. 1987. “An evaluation of a cognitive theory of response-order effects in survey measurement.” *Public Opinion Quarterly,* 51: 201-209.

Kulas, J. T., A. A. Stachowski, and B. A. Haynes. 2008. “Middle response function in Likert-responses to personality items.” *Journal of Business Psychology,* 22: 251-259.

Likert, R. 1932. “A technique for the measurement of attitudes.” *Archives in Psychology,* 140: 1-55.

Matell, M. S. and J. Jacoby. 1972. “Is there an optimal number of alternatives for Likert scale items? Effects of testing time and scale properties.” *Journal of Applied Psychology,* 56(6): 506-509.

McFadden, L. S. and S. E. Krug. 1984. “Psychometric function of the “neutral” response option in clinical personality scales.” *Multivariate Experimental Clinical Research,* 7: 25-33.

Mueller, J. E. 1970. “Choosing among 133 candidates.” *Public Opinion Quarterly,* 34: 395-402.

Nicholls, M. R., C. A. Orr, M. Okubo, and A. Loftus. 2006. “Satisfaction guaranteed: The effect of spatial biases on responses to Likert scales.” *Psychological Science,* 17: 1027-1028.

Noh, Y. 2010a. “A study on developing evaluation criteria for electronic resources in evaluation indicators of libraries.” *Journal of Academic Librarianship,* 36(1): 41-52. <http://dx.doi.org/10.1016/j.acalib.2009.11.005>.

Noh, Y. 2010b. “A study on designing of metadata for constructing the Library Map Information System.” *Journal of the Korean Society for Information Management,* 27(3): 241-264.

Noh, Y. 2010c. “A Study on Metadata Elements for Web-based Reference Resources System Developed through Usability Testing.” *Library Hi Tech,* 29(2): 242-265. <http://www.emeraldinsight.com/journals.htm?articleid=1927525>.

Payne, J. D. 1971. “The effects of reversing the order of verbal rating scales in a postal survey.” *Journal of the Market Research Society,* 14: 30-44.

Payne, S. L. 1951. *The Art of Asking Questions.* Princeton: Princeton University Press.

Quinn, S. B. and W. A. Belson. 1969. *The Effects of Reversing of the Order of Presentation of Verbal Rating Scales in Survey Interviews.* London: Survey Research Centre.

Ruggg, D. and H. Cantril. 1944. “The wording of questions.” In H. Cantril (Ed.) *Gauging public opinion.* Princeton: Princeton University Press.

Schwarz, N. B., H. J. Knauper, E. Hippler, E. Noelle-Neumann, and L. Clark. 1991. “Rating scales: Numeric values may change the meaning of scale labels.” *Public Opinion Quarterly,* 55: 570-582.

Schwarz, N., H. J. Hippler, and E. Noelle-Neumann. 1992. “A cognitive model of response-order effects in survey measurement.” In N. Schwarz

& S. Sudman (Eds.), *Context effects in social and psychological research* (pp. 187-199). New York: Springer-Verlag.

Sedlmeier, P. 2006. "The role of scales in student ratings." *Learning and Instruction,* 16: 401-415.

Shcuman, H. and S. Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Sheehan, B. K. and J. S. McMillan. 1999. "Response variation in e-mail surveys: An exploration." *Journal of Advertising Journal,* 39(4): 45-54.

Stewart, T. J., E. Roberts, P. Eleazer, R. Boland, and D. Widland. 2006. "Reliability and validity issues for two common measures of medical students' attitudes toward older adults." *Educational Gerontology,* 32: 409-421.

Weems, G. H., A. J. Onwuegbuzie, J. B. Schreiber, and S. J. Eggers. 2003. "Characteristics of respondents who respond differently to positively and negatively worded items on rating scales." *Assessment and Evaluation in Higher Education,* 28: 587-607.

Wildt, A. R. and M. B. Mazis. 1978. "Determinants of scale responses: Label versus position." *Journal of Marketing Research,* XV: 261-267.

Worcester, R. M. and T. R. Burns. 1975. "A statistical examination of the relative precision of verbal scales." *Journal of Market Research Society,* 17(3): 181-197.

Wyatt, R. C. and L. S. Meyers. 1987. "Psychometric properties of four 5-point Likert-type scales." *Educational and Phychologicla Measurement,* 47: 27-35.

Yammarino, F. J., S. J. Skinner, and T. L. Childers. 1991. "Understanding mail survey response behavior: A meta-analysis." *Public Opinion Quarterly,* 55: 613-639.

Yun, G. W. and C. W. Trumbo. 2000. "Comparative response to a survey executed by post, e-mail, & web form." *Journal of Market Research Society,* 6(1): 235-239.