Augmented Web Survey with enhanced response UI for Touch-based Psychological State Estimation

Takaaki Nakagawa
Kyushu University,
nakagawa.takaaki@arakawa-lab.com

Yutaka Arakawa
Kyushu University,
arakawa@ait.kyushu-u.ac.jp

Yugo Nakamura
Kyushu University,
y-nakamura@ait.kyushu-u.ac.jp

***Abstract*—In this paper, we propose a method for estimating psychological states such as hesitation and interest during survey responses based on touch operation behavior (touch log). In addition, we propose two enhanced web survey RUIs (Response User Interfaces), (a) a slidebar type RUI and (b) a magnifier type RUI, to expand the operational differences due to hesitation and interest for more accurate estimation. To evaluate the proposed RUIs, we implemented Augmented Web Survey (AWS) as an extension of LimeSurvey and conducted two experiments. Responses were collected from 8 subjects in the experiment to evaluate (a) a slidebar type RUI and from 21 subjects to evaluate (b) a magnifier type RUI. We measured the degree of hesitation by evaluating the subjective hesitation to the questions on a 5-point scale. For the measurement of the degree of interest, the participants were asked to select their top two favorite images among the six images, and the true value was taken as the top two. As a result of comparing the operation logs of the conventional RUI and the slidebar type RUI, it was confirmed that there was a significant difference depending on the degree of hesitation. In addition, more magnification operations were obtained for the magnifier type RUI than the standard pinch-out operation suggesting that it may help estimate psychological states such as hesitation.**

***Index Terms*—Web survey, touch operation log, hesitation estimation, interest estimation, UI design**

I. Introduction

Web survey systems such as Google Form are widely used in various situations. Many companies undertake stress checks, engagement measurement, and online market research. In addition, universities use web surveys to conduct lecture evaluations. The merit of using a web survey system is that it is easier and cheaper than paper-based surveys, and it can request a large number of responses from a wide range of respondents. However, it is also known that web surveys may have some problems.

A common response format for surveys is the Likert scale [1], such as the 10-point method or the 7-point method. When the 10-point method is used, it is difficult for the respondents themselves to recognize whether the answer is 7 or 8 on the 10-point scale, and there is a possibility that their answers will be inconsistent. The second problem is the accuracy of the web survey. For example, in the case of a purchase intention survey, people who answered "I have the intention to purchase" in the web survey do not always actually purchase. To address these issues, we thought that we could increase the amount of information obtained from survey respondents by sensing the process of answering the web survey(touch log), and thereby understand their psychological state, such as hesitation(certainty) and interest when answering the survey.

Regarding the detection of hesitation based on manipulation behavior, the study by Miyazaki et al. [2] measured the mouse logs of subjects answering English word sorting questions using a PC and analyzed them using machine learning to determine the presence or absence of hesitation with 82.2% accuracy. However, this method is not compatible with smartphones, and no analysis of hesitation using touch logs has been conducted.

On the other hand, a study using touch logs was conducted by Kato et al. [3], who obtained the touch logs of subjects when they browsed online articles and analyzed the correlation with their interest level in online articles. However, the number of features analyzed by Kato et al. is small, and additional behavioral information is needed to estimate the interest level. In addition, there is a possibility that the behavior is different from the behavior when answering the web survey.

As a typical RUI used in web surveys, there is a radio button type RUI that allows users to select one of several options. Initially, we explored the possibility of detecting hesitation with the behavior of radio buttons by changing the choice after selecting it once. Initially, we explored the possibility of detecting hesitation by changing the choice of a radio button after selecting it, but even in the case of hesitation, there were many cases where the user did not change the choice after selecting it, and the difference in behavior between the case of hesitation and the case where the user was not interested was considered small. In addition to sensing touch logs, we propose a novel RUI that more easily reflects psychological states such as hesitation and interest.

First, with hesitation, we propose a slidebar type RUI. This is a RUI that allows users to answer by sliding a button over the choices. The speed and amount of movement of the button until the final choice is decided can be obtained as a touch log, increasing the amount of information compared to radio buttons. We hypothesized that hesitation about a question would appear in the movement of the slidebar until a choice was made (**Hypothesis 1**).

Next, for detecting interest, we propose a magnifier type RUI. The magnifier type RUI is a RUI in which tapping on an image causes the tapped area to be cut out and magnified in the shape of a circular magnifying glass above the tapped area, which is less expensive to operate than the standard pinch-out operation with two fingers. It is possible that there is a behavioral difference such that the images of interest take longer to zoom in than those of no interest. The following two hypotheses were formulated for the magnifier type

RUI. **Hypothesis 2**: Manipulation behavior when zooming an image reflects the user's interests. **Hypothesis 3**: the magnifier type RUI can get more zoom operations from the user than standard pinch-out operation.

In this study, to verify the above three hypotheses, we created a web survey using the proposed RUI and collected the touch logs during the responses. We analyzed the behavioral differences between the high and low groups in terms of hesitation and interest in the questions by comparing and verifying the features obtained from the touch logs with the conventional RUI. The results showed that many features of the slidebar type RUI showed significant differences between the two groups compared to radio buttons and obtained high correlation coefficients with hesitation. Although the magnifier type RUI did not show any difference in behavior depending on the interest level, it obtained more magnification operations than the standard magnification operations, which may be useful for estimating hesitation and other psychological states.

This paper is organized as follows: section Ⅱ describes related research; section Ⅲ describes the proposed augmented web survey system; section Ⅳ describes the hypothesis testing experiment of the slidebar type RUI; section Ⅴ describes the hypothesis testing experiment of the magnifier type RUI. Finally, we summarize this paper in section Ⅵ.

II. Related Research

Sakamaki et al. [4] conducted a web survey related to automobile manufacturers, and estimated the confidence level using mouse logs during the responses. The time spent by the mouse on each option was used as the confidence level of that option, and the correlation between the time spent on each option and the actual response rank of each option was analyzed. The results of the experiment showed that the response rank ±1st could be estimated with 73.3% accuracy. However, since the cursor is not always present on the screen in smartphones, it is not possible to define the time spent on the choices.

Gogami [5] et al. proposed a method for detecting Satisficing based on the degree of scrolling and text modification, focusing on the touch operation behavior during the response. Satisficing is the act of trying to achieve a goal with the least amount of effort. To detect Satisficing, screening questions are commonly used, such as DQS, which checks for violations of response instructions, and ARS, which checks for inconsistencies in responses to question pairs with the same content. However, adding screening questions can increase the number of questions and make respondents more skeptical. Gagami et al. conducted a crowd sourcing experiment with over 5600 people and found that 85.6% of Satisficing respondents which can be detected by the traditional method with additional questions.,could be detected without additional questions.

In this study, we developed an augmented web survey system for detecting hesitation and interest by extending the Operation Logger, a touch log collection system developed by Gogami et al. to collect the operation behavior of the proposed RUI. In addition, since respondents who are Satisficing are not hesitated in their answers and are not interested in the questions, the accuracy of Satisficing detection may be further improved by combining the RUI proposed in this study with the work of Gogami et al.

III. Proposed System: Augmented Web Survey

We have implemented Augmented Web Survey (AWS) as an extension of LimeSurvey, a major web survey system in the world. AWS consists of two features. One is the Operation Logger plug-in for recording user's operations developed by Gogami et al [5]. The other is some RUIs using JavaScript and CSS, which aim to amply the difference of user's operations.

A. Overview of Operation Logger

The Operation Logger collects three types of touch events as logs. There are three types: touchstart, touchmove, and touchend, which occur at the start, swipe, and end of each operation, respectively. We obtain the time, the coordinates on the screen, the amount of movement from the top of the page, and the type of touch event at these timings. This allows us to detect the "scroll speed", and "scroll length". At the timing of tapping a radio button choice, we obtain the time, the id of choice, and the id of the question. By doing so, we can detect the "response time for each question of the choice format" and "change of choice".

In addition to the above data, we extended the Operation Logger to obtain the operation log of the RUI proposed in this study.

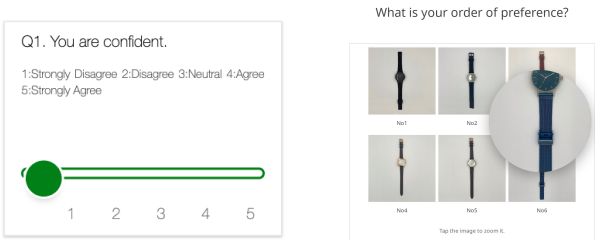
B. Proposed RUIs

First, we will explain radio buttons, which are the conventional RUI. The radio button is a RUI used in multiple-choice questions such as the n-points method, and it is a RUI where the user selects an option by tapping on the option.

We propose a slidebar type RUI as a RUI that better reflects hesitation. The slidebar type RUI is a RUI that allows users to answer by dragging the response button in the slidebar left or right with their finger. The slidebar type RUI is intended to be used for n-points method like radio buttons, and n options are displayed at the bottom of the slidebar. It is designed to select the closest option when the response button is released, and the response button automatically moves to the closest option. The initial value is an empty field on the left side of the RUI. This is because if the initial value is placed in one of the choices, the answer operation may not be obtained if the answer value matches the initial value. In addition, the slidebar type RUI cannot be answered by touch operation as in the case of radio buttons.

As the operation log for the slidebar type RUI, the Operation Logger is used to obtain the log when the button is tapped, the timing when the finger is moved while tapping the button, and the timing when the finger is released. Obtain log are time, id of the question and the answer to the question. By analyzing these logs, we can measure the response time per question, the change of answer choices, the number of reversals of the response buttons (the number of times the direction of the buttons changed), and the moving speed of the button.

The magnifier type RUI is a RUI that zooms the tapped portion of an image by tapping on it, creating a circular magnifying glass above the tapped portion. We thought that this would encourage respondents to zoom in more, since they could do so by simply tapping on the image. Basically, in a web survey such as web sites and willingness-to-pay survey, multiple images are often displayed on a single page,



(a) A slidebar type (b) A magnifier type

Fig. 1. Proposed RUIs

and we thought that the more interesting the image, the longer it would take to zoom in. We also think that the interest may be reflected in other information such as the speed of movement of the magnifier.

The magnifier type RUI acquires the time, the id of the magnified image, and the coordinates in the screen at the timing when the image is tapped and zoomed, when the finger is moved while magnifying, and when the finger is released.By analyzing these logs, we can measure the zooming time per image unit, the moving speed of the magnifier, and the moving distance of the magnifier.

IV. Evaluation of slidebar type RUI

We formulated the following hypothesis for the slidebar type answer UI. **Hypothesis 1**: hesitation about a question would appear in the movement of the slidebar until a choice was made. The purpose of the experiment described in this section is to test hypothesis 1. Specifically, we create a web survey created using the proposed RUI and radio buttons, respectively, and measure the touch log when answering using Operation Logger. The correlation between the features obtained from those touch logs and the subjective hesitation of the subject to the questions is analyzed and evaluated by comparing them for each RUI.

A. Survey contents

In this experiment, we set five questions of four arithmetic operations whose answers are 1 to 5 as questions that are less likely to be hesitating. In addition, 10 TEG [6] questions, which are used in personality tests, were set as questions that are more likely to cause hesitation.

B. Experimental methods

In this experiment, we use LimeSurvey to deliver web surveys created with a radio button and a slidebar type RUI, respectively, and collect the touch logs of the subjects' responses using Operation Logger. The operation information of each RUI is extracted from the collected touch logs, statistically analyzed, and the correlation with hesitation is analyzed.

C. subjects

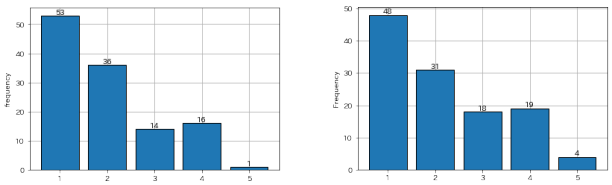
The subjects of this experiment were 8 males in their twenties. In this experiment, we shared the URL and notes of the web survey created by LimeSurvey on SNS. The shared notes are shown in (1)∼(3).

1) The survey will be answered by smart phones.
2) When answering the survey, the participants sit in a chair, hold the smartphone with their left hand, and operate it with their right index finger.
3) View a page for at least three seconds and then move on to the next page.

2) was set up to eliminate differences in operating behavior due to posture and the way the smartphone is held during smartphone operation. Regarding 3), this is because the Operation Logger may not be able to collect logs if a page transition occurs before the page has finished loading.

D. Measuring degree of hesitation

In this experiment, to measure the degree of hesitation, we asked the participants to answer the subjective degree of hesitation for each question in the survey using a 5-point method. Specifically, we asked the participants to answer hesitation in five levels: "1"∼"5". The five items were "I was hesitated (5)," "I was a little hesitated (4)," "I was undecided (3)," "I was not too hesitated (2)," and "I was not hesitated (1)," respectively.



(a) Data distribution for radio button (b) Data distribution for slidebar type RUI

Fig. 2. Distribution of hesitation data

The results of the survey in which the subjective hesitation to the questions was answered on a 5-point scale are shown in Figure 2(a) and 2(b). The vertical axis is the number of responses to the survey and the horizontal axis is the degree of hesitation about the questions. The number of respondents who answered that they were "hesitated (5)" was the lowest and the number of respondents who answered that they were "not hesitated (1)" was the highest in any RUIs. To classify these into a high group with hesitation and a low group without hesitation, "hesitated (5)" and "a little hesitated (4)" were designated as the high group, and "not so much hesitated (2)" and "not so much hesitated (1)" were designated as the low group.

E. Evaluation method

In this experiment, we compared the operation information obtained from the touch logs between the two groups with high and low hesitation in the conventional RUI and the proposed RUI, respectively, between the two groups. Specifically, characteristics such as the speed of the button and response time of each group were evaluated by statistical tests to see if there were significant differences between the two groups. The procedure for the statistical tests is shown below. First of all, the distribution of each feature of each RUI was evaluated to be normally distributed by the Shapiro-Wilk test at the 5% level of significance. In addition, the F test was used to determine whether each feature was equally distributed or not at the 5% level of significance. Since all features were not normally distributed, they were tested two-tailed at a significance level of 1% using Mann-Whitney's U test when the variances were equal and Brunner-Munzel test when the variances were not equal. The correlation between each feature and hesitation was calculated using Spearman's rank correlation coefficient. To

interpret the magnitude of the correlation coefficient, we use Guilford's rule [7]. The significance level of the correlation coefficient is considered to be 5%.

F. Results

First, the p-values for testing the features between the two groups of high and low radio buttons are shown in Table I. The correlation coefficients for each feature are also shown in Table II.

TABLE I
FEATURE COMPARISON RESULTS FOR TWO GROUPS OF RADIO BUTTON HESITATION

Features	p-value
Response time	**0.003**
Number of response changes	0.241

TABLE II
CORRELATION COEFFICIENTS BETWEEN RADIO BUTTON HESITATION AND FEATURE VALUES

Features	COR	p-value(COR)
Response time	0.578	**0.000**
Number of response changes	0.132	0.149

In terms of response time using radio buttons, there is a significant difference between the two groups, and the correlation coefficient indicates that there is a positive correlation. This indicates that the more confused the respondent is, the longer the response time becomes, which is reasonable. There is a significant difference in response time for the radio buttons, but this alone is not informative.

Next, the p-values for testing the characteristics between the two groups of high and low in the touch slidebar type RUI are shown in Table III.

TABLE III
COMPARISON RESULTS OF FEATURES IN TWO GROUPS OF HESITATION OF SLIDEBAR TYPE RUI

Features	p-value
Response time	**0.000**
Number of response changes	**0.000**
Number of button reversal	0.063
Amount of button movement	0.142
Average speed	**0.000**
Maximum speed	0.049
Minimum speed	**0.000**
Variance of speed	**0.000**

TABLE IV
CORRELATION COEFFICIENTS BETWEEN HESITATION AND FEATURE VALUES FOR SLIDEBAR TYPE RUI

Features	COR	COR(p-value)
Response time	0.74	**0.000**
Number of response changes	0.38	**0.000**
Number of button reversal	0.18	0.054
Amount of button movement	-0.09	0.350
Average speed	-0.66	**0.000**
Maximum speed	-0.45	**0.000**
Minimum speed	-0.56	**0.000**
Variance of speed	-0.55	**0.000**

For the slidebar type RUI, five features were found to be significantly different between the two groups. First, there is a significant difference in the response time, and the correlation coefficient shows that there is a strong positive correlation. This confirms that the correlation is even stronger, although a similar significant difference was found for the radio buttons. There is a significant difference in the number of response changes, and the correlation coefficient indicates that there is a positive weak correlation. This is reasonable, as it shows that the more they stray, the more they change their answers. Next, there is a significant difference in the average speed of the buttons, and the correlation coefficient indicates that there is a negative correlation. This is reasonable, as it shows that the person who is confused is more hesitant to choose an answer, and the speed of the buttons is slower overall. Furthermore, the results for the minimum speed were similar to those for the average speed. Finally, there is a significant difference in the variance of the button speed, and the correlation coefficient shows that there is a negative correlation. The reason is that there is a significantly faster behavior of the push operation when there is no hesitation, which results in larger variance. As shown above, significant differences were found for more features in the slidebar type RUI than in the radio button. In addition, the differences were also found to be larger. These results indicate that the slidebar type RUI is an RUI that reflects hesitation more than the radio button type, where the difference in behavior appears to be greater depending on the size of the hesitation. Therefore, hypothesis 1 is proven.

V. EVALUATION OF MAGNIFIER TYPE RUI

We made the following two hypotheses for the magnifier type UI. **Hypothesis 2**: Manipulation behavior when zooming an image reflects the user's interests. **Hypothesis 3**: the magnifier type RUI can get more zoom operations from the user than standard pinch-out operation. The purpose of the experiment described in this section is to test hypothesis 2 and hypothesis 3.

A. Experimental methods

In this experiment, we created a survey to rank idol groups such as "Nokizaka46" and "BTS" and products such as "watches" and "accessories". A web survey containing images of these contents is delivered by LimeSurvey with a magnifier type RUI, and the touch logs of the subjects' responses are collected using Operation Logger. We extract the operation information of each RUI from the collected touch logs, statistically analyze it, and analyze the correlation with the degree of interest.

B. Survey contents

In this experiment, four idol groups, "Nokizaka46", "KeyakiZaka46", "BTS", and "Sandaime J Soul Brothers", and two products, "Watches" and "Rings", were selected as the contents to be ranked. Since the subjects were of mixed gender, we used two groups of female idols and two groups of male idols. For the products, we used "watches" as a problem for giving gifts to men and "rings" as a problem for giving gifts to women. We extracted six images from each of these contents. The criteria for selection were the top six most popular images from each content. The images were chosen such that details could not be seen without magnification in order to induce the subject to zoom in.

We created a survey on Lime Survey with six images per page, and set a question to answer the top two positions among the six images.

C. subjects

The total number of subjects was 21, 15 males and 6 females. The subjects were 12 in their teens and 9 in their 20s. The web survey URL created by LimeSurvey and the precautions described in Section 4, Subsection C were shared via SNS (Social Network Service).

D. Measuring interest level

In this experiment, we compare the number of magnification operations performed by the standard and a magnifier type RUI. To detect whether there is a difference in behavior depending on the degree of interest in the image, the manipulation information obtained from the touch logs between the two groups with high and low interest was compared in the same way as in Section 4, Subsection F.

E. Evaluation method

In this experiment, we asked the participants to select the top two images out of six images. For the idol groups, if there is prerequisite knowledge of the group, it may affect the difference in behavior. For this reason, we asked the participants to answer the name of the idol group on the questionnaire, and those who answered correctly were eliminated from the data.

F. Results

First, we compare the number of magnification operations performed by the standard magnifier and the magnifier type RUI. The total number of standard magnification operations was 46, and the total number of magnification operations using the magnifier type RUI was 652, with the magnification using the magnifier type RUI being the larger result. Next, the calculated number of magnifications for each subject is shown in Table V.

TABLE V
COMPARISON OF THE NUMBER OF MAGNIFICATION OPERATIONS

Subjects	standard	magnifier type RUI	Subjects	standard	magnifier type RUI
No.1	1	6	No.12	0	63
No.2	1	3	No.13	1	2
No.3	0	50	No.14	0	13
No.4	1	24	No.15	5	0
No.5	3	56	No.16	0	3
No.6	0	124	No.17	8	6
No.7	0	15	No.18	0	15
No.8	0	57	No.19	0	0
No.9	0	121	No.20	0	21
No.10	5	14	No.21	21	9
No.11	0	50	average	2.2	31.0

In Table V, we can see that the magnifier type RUI has more magnification operations, although there is a bias among subjects. Therefore, Hypothesis 3 is proven.

Table VI shows that there was no significant difference between the two groups of interest in the magnifier type RUI. Therefore, Hypothesis 2 could not be proven. However, the magnifier type RUI can confirm more magnification operations than the standard magnification operations, and may be used to estimate other psychological states such as hesitation and Satisficing.

TABLE VI
COMPARISON RESULTS OF FEATURES IN TWO GROUPS OF INTEREST

Features	p-value
Magnification time	0.626
Number of magnifications	0.189
Magnifier movement amount	0.207
Average speed	0.346
Maximum speed	0.200
Minimum speed	0.999
Variance of speed	0.147

VI. CONCLUSION

In this paper, we proposed a method for estimating psychological states such as hesitation and interest during survey responses from touch logs. In addition, for more accurate estimation, we proposed an augmented web survey system that extends the RUI to expand the operation difference due to hesitation and interest and enables logging of these. We created a web survey using the proposed RUI, divided the touch logs into two groups according to their degree of hesitation and interest and verified their behavior differences. As a result, we confirmed that a proposed slidebar type RUIs are more reflective of hesitation than the conventional radio buttons. Although the magnifier type RUI did not show any difference in behavior depending on the interest level, it obtained more magnification operations than the standard magnification operations, which may help estimate hesitation and other psychological states. As a future prospect, we plan to build a machine learning model for hesitation detection by crowdsourcing a web survey with a slidebar type RUI and collecting large-scale data. In addition, we are planning to conduct an evaluation experiment to verify whether the magnifier type RUI reflects psychological states other than interest.

ACKNOWLEDGEMENT

This work was partially supported by JSPS KAKENHI Grant Numbers JP18H03233.

REFERENCES

[1] R.L Aiken, ed., "Psychological Testing and Assessment", 10th edition, Allyn and Bacon, 2000.

[2] Yoshinori Miyazaki and Naho Atsumi and Mitsumasa Zushi and Ken Norizuki, "Detecting learners' hesitation in solving word-reordering problems with machine learning for classification", EdMedia 2016–World Conference on Educational Media and Technology, pp.1638–1646, June 2016.

[3] Yuta Kato and Takeshi Iwamoto and Michito Matsumoto, "Estimate Degree of Interest in the Web Contents by using Touch Operation Log", Journal of Information Processing, vol.59, no.2, pp.508–518, 2018.

[4] Ryuji Sakamaki and Satoshi Someya, "Rank Presumption of Strength of Conviction Level Based on Respondent's Mouse-log in Web-query with Image Choices", Journal of the Visualization Society of Japan, vol.31, no.12, p.69, 2011.

[5] Masaki Gogami and Yuki Matsuda and Yutaka Arakawa and Keiichi Yasumoto, "Detection of Careless Responses in web survey Using Answering Behavior on Smartphone", IEEE Access, vol. 9, pp. 53205-53218, 2021.

[6] Shigeo Kashiwagi, "The trait theoretic evaluation of the TEG from the view point of the Five-Factor Model", The Japanese Journal of Psychology, vol.69, no.6, pp.468–477, 1999.

[7] L. Parish and J.P. Guilford, "Fundamental statistics in psychology and education", British Journal of Educational Studies, vol.5, p.191, 1957.