FACTORIAL DESIGN IN LIMESURVEY: APPROACHES OF REALIZATION

Mykola Sydorov
Taras Shevchenko National University of Kyiv, Ukraine, myksyd@knu.ua

Oleksii Sereda
Taras Shevchenko National University of Kyiv, Ukraine, as_sereda@knu.ua

Factorial design in survey is a powerful tool for multivariate analysis when one needs to investigate attitudes to any multiple factors objects and to get importance and significance of theirs. One of problems of such surveys is to select the sufficient number of vignettes created in maximum variability and orthogonality method, for example D-efficient. Bigger amount of different vignettes means higher level of their variability but needs more respondents for participation or bigger number of vignettes in one deck (questionnaire). From the other side we have to take into account the influence of low response rate, so usual for web surveys, that could bias data and even waste of some dimensions' levels. As every respondent or a group of respondents has to receive unique deck (set of vignettes) so as the number of different questionnaires (same core and different vignette part) have to be very high and paper based version of survey is less useful that electronic (web). In this case we have to reach a balance between a respondent sample size, vignette sample size, vignette deck size and questionnaire delivery method.

For questionnaires Auspurg and Hinz (2014) recommended up to 20 vignettes in one deck but we, based on paradata analysis of the number of our web researches in Ukraine, recommend no more than 6-10 vignettes (depends of vignette complexity) in one deck due to the risk of item nonresponse.

We used 3 approaches of questionnaire delivery to the respondents: delivery by emails with previously generated fixed decks, delivery by social network with random vignettes selection and emails delivery with random vignettes selection. The first and third approaches show unit response rate (about 0.2 - 0.3) and sample representation, the second one – just item response rate.

In the first approach (3 surveys) all decks were generated and associated with fixed emails by tokens before survey had been started. As a result in first of surveys with very low response rate of one of respondent's category, we lost 2 levels in one dimension (month salary) and that dimensions was not presented in regression equation at all. After that we used vignettes randomisation before deck generation – it removed levels' incompleteness but did not produce the bias in vignette frequency (should be uniform distribution). In this approach there was a possibility of sending addressed remainders to respondents for incomplete questionnaires.

The second approach was realised in pilot surveys for development of other deck generating method. There was no email addresses database and we had to generate different decks online just when respondent started the questionnaire. By this method all 60 vignettes from the sample were divided into 6 groups (6 different vignettes in deck for every respondents and 10 different alternatives for vignette in every group). As a result the distribution of vignettes was very close to uniform and there was no loss of any levels. In this approach we did not have possibility to send remainders as well as invitations.

The third approach is a result of evolution of two previous ones: there was an email database and random online generation of decks. There was not significant bias in dimension levels because of possible unit nonresponse. The distribution of vignettes was very similar to uniform. Due to existence of email database we had the possibility to send invitations and reminders. Even if a respondent did not complete the questionnaire, they could continue without regeneration of vignettes.

References

Auspurg, K.; Hinz, T. (2014) Factorial Survey Experiments. *Quantitative Applications in the Social Sciences* vol. 175, SAGE Publications, 168