

Effects of questionnaire length on response rates:

Review of findings and guidelines for future research

Mirta Galešić

Department of Psychology
University of Zagreb, Croatia
mirta.galesic@zg.tel.hr

Introduction

It is generally assumed that questionnaire length has a significant effect on survey response rate. It is one of the most frequent reasons sample members give when refusing to participate in a survey (Burchell & Marsh, 1992). However, empirical findings are equivocal. One of the reasons is that various authors use different operational definitions of questionnaire length. Studies that define length as number of questions and/or pages usually show negative relationship between questionnaire length and response rates. When questionnaire length is defined by its physical format, the same number of questions printed on smaller-sized paper, but also on larger number of pages would yield bigger response rates. Duration of interview in minutes can be positively related to its relevance for respondents, which in turn has positive effect on response rate. Many authors claim that response rates are influenced by subjective rather than objective indices of questionnaire length. Perception of length is often moderated by various other characteristics of survey situation, such as respondents' interest in topic of the survey and use of financial or other incentives.

Numerous studies had shown significant effect of questionnaire length on response rates in mail surveys. Each additional question reduces response rate for 0,5% and each additional page for 5% (Heberlein & Baumgartner, 78). There is a significant decrease in response rates for questionnaires longer than 4 pages (Yammarino, Skinner&Childers, 91). The effect of questionnaire length in data collection techniques that include personal contact is less clear. In telephone surveys, 14% of contacted sample refuses 40-minute interview and 9% refuses 20-minute interview (Collins et al, 88). In face-to-face surveys 61% of all refusals could be due to anticipated length (Burchell & Marsh, 92), but other researchers found similar refusal rates for interviews lasting 25 and 75 minutes (Frankel & Sharp, 81).

Quality of data is also commonly believed to be affected by questionnaire length. As the questionnaire lasts, respondents can become tired, annoyed, bored and/or distracted by external factors. This can decrease the “degree of effort and thought that respondent invests in answering the questions”, as quality of data is usually defined (Houston & Ford, 1976; in James & Bolstein,1990). Operationally, quality of data is measured in different ways: as

readiness to sincerely answer socially sensitive questions, variability of answers, length of answers given to open questions, percentage of unit nonresponse, percentage of “don’t know” answers etc. Findings on the effect of questionnaire length on data quality generally suggest negative influence of overly long questionnaires on data quality (Burchell & Marsh, 1992; Helgeson & Ursic, 1994; Herzog & Bachman, 1981). Again, this effect is often visible only after control for moderating factors such as level of interest in the questionnaire topic, monetary incentives or format of questions.

Questionnaire length could have significant effect on response rates in online surveys. Specifically, online surveys bring direct financial costs to the participants (phone line); are often not suitable for use of the incentive types that proved to be the most useful motivators in classic surveys such as repeated contacts or immediate monetary reward (although see Bosnjak & Tuten, 2002,); do not include personal contact with the interviewer/researcher (less/no social pressure); and have strong competitors in the fight for respondents’ attention in various other interesting content on the Net.

Studies on the effect of questionnaire length on response rates and/or quality of data in online surveys are still missing. Furthermore, a theory of survey behavior is needed in order to better understand and generalize empirical findings. Therefore we (1) examined relations between questionnaire length and unit response rates, quality of data and perceived length on a suitable data set collected on a Croatian online panel, (2) suggested a draft theoretical model which could be used in further research of the effects of questionnaire length and other characteristics of survey situation on response rates and quality of data.

1. Experiences from online panel

1.1. Proof

PROOF is an online research agency founded in summer 2000 in Zagreb, Croatia. It has conducted numerous surveys for various clients on its own online panel of 4500 Croatian Internet users. Survey topics included: Internet users' profile, use of the Internet (ways, intensity, places), interest in on-line services (ISPs, Internet banking, WAP, e-commerce), use of various products (mobile phones, cars), on-line promotion effectiveness etc.

The panel was built through advertising on most visited web pages, via newsletters of Croatian largest ISP, and by unpaid promotion in offline media. Surveys were available only to members of the panel, invited via e-mail containing unique URLs that guided respondents to their survey. This enabled automatic linking of answers with previously collected data about the respondent, interactive modification of survey questions according to respondent's characteristics, and creating specially designed - targeted samples.

1.2. Procedure

In March 2001 three surveys were conducted simultaneously, each on a separate random subsample of the panel. The topics of the questionnaires received by each subsample were as follows: ISPs - Internet Service Providers' services and promotion; Net.Basics -basic characteristics of Internet use and most visited web pages; and Net.Services - use of various Internet services (mail, chat, newsletters, games,..) and most visited web pages (the same set of questions as the second part of Net.Basics questionnaire). Due to different demands of Proof's clients, the questionnaires were unequal in length. Formal characteristics of the questionnaires are shown in Table 1.

Table 1. Formal characteristics of the questionnaires.

Questionnaire topic	ISPs	Net.Basics	Net.Services
Total number of questions	28	19	32
Number of open questions	8	11	16
Number of questions answered by an average respondent*	M=15.4 SD=1.25	M=16.2 SD=3.06	M=21.3 SD=4.78
Number of open questions answered by an average respondent**	M=2.84 SD=1.25	M=5.5 SD=2.66	M=6.7 SD=3.71

*Questionnaires contained numerous branching points.

**Following types of questions were treated as open: requests for more detailed explanation of some behavior or opinion; requests for naming three web pages that were visited most often during the last month; request for naming certain service or brand one uses. Following types of questions were considered closed: multiple choice questions – one or several possible answers, rating scales, requests for estimating certain number, such as hours&minutes of Internet use in an average workday.

Average number of questions answered by respondents in each of the three questionnaires differed significantly ($p < .001$). The shortest was the questionnaire on ISPs. Net.Basics was of relatively medium length, while Net.Services questionnaire was the longest among the three.

The questionnaires consisted mostly of rather simple questions already tested in previous Proof surveys. The panel members assigned to the three questionnaires were otherwise treated in the same way: panel members were randomly assigned to one of the three subsamples; all of them were surveyed at least once in the past; all received the same e-mail invitation (Figure 1), questionnaire with the same visual design (Figure 2), and the same incentive (participation in a sweepstake with 100 inexpensive rewards – free access to the Internet in duration from 1 month to 1 year). Length of the questionnaire was not mentioned, except for the general statement that the survey will be “short”.

Figure 1. E-mail invitation.

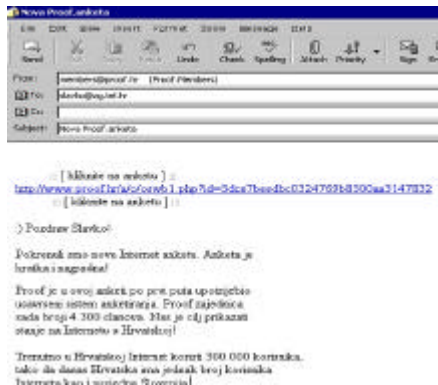
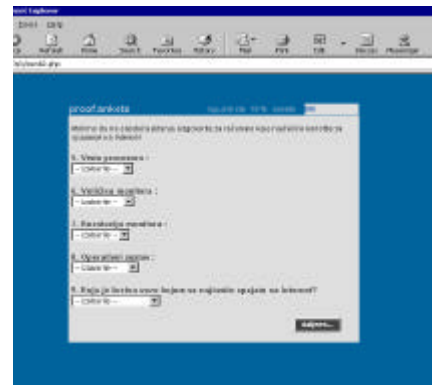


Figure 2. Questionnaire design.



Initial response rate, defined as percentage of invited panel members who answered the 1st question, was approximately equal for all three questionnaires (Table 2).

Table 2. Initial response rates for the three questionnaires.

Questionnaire topic	ISPs (shortest)	Net.Basics (medium)	Net.Services (longest)
Number of panel members invited	2272	1000	1000
Number (%) of invited members who answered the 1 st question	1096 (48%)	465 (47%)	498 (50%)

At the end of each questionnaire four additional questions – equal for all three questionnaire types - were asked (Table 3). These last four questions provided a suitable ground for observing possible effects of questionnaire length on unit nonresponse and quality of data¹.

Table 3. Set of four questions asked at the end of each questionnaire.

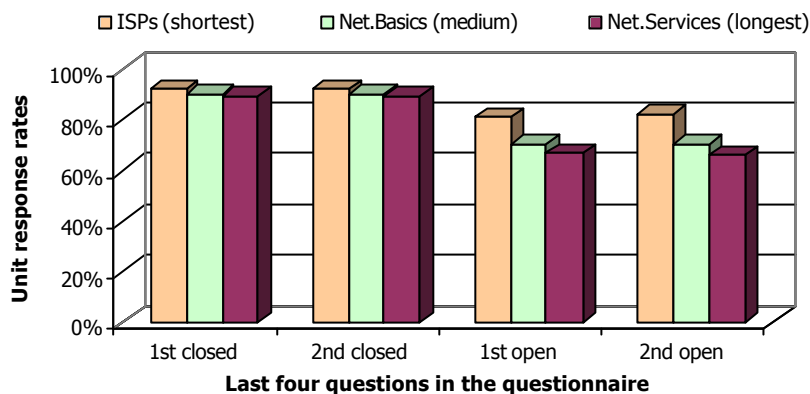
Question no	Type	Text
1	closed (rating scale)	What do you think about the length of this questionnaire?
2	closed (rating scale)	How interesting did you find this questionnaire?
3	open	Is there something Proof could do to improve the satisfaction of its members?
4	open	What do you think about Proof in general?

¹ The questions in the main body of the questionnaire were less suitable for this purpose because proportion of respondents coming to each question varied greatly due to numerous branching options in the questionnaire.

1.3. Some observations

First, we analyzed the relation between questionnaire length and *unit response rates for the last four items*. According to the findings from «classic» surveys, longer questionnaires should yield lower response rates for units at the end of the survey (for example Yammarino et al, 1991). Our results are shown in Figure 3.

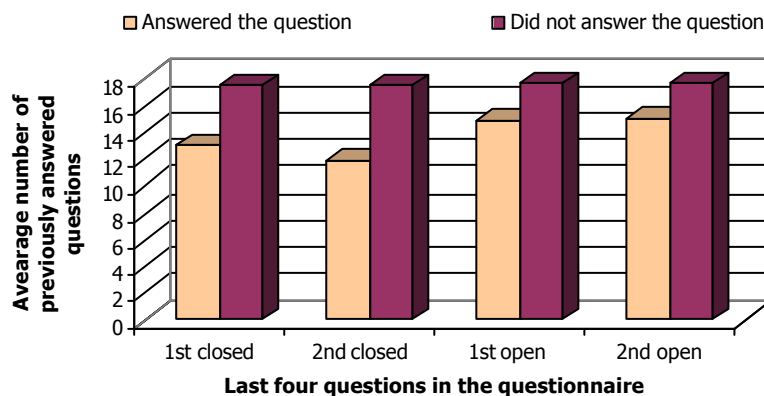
Figure 3. Last four units' response rates depending on questionnaire length



We found significant difference in unit response rates between the three questionnaires only for open questions (chi square significant on $p < .001$). No difference was found for closed questions ($p > .05$). Possible explanation could be that answering open questions demands more effort, which incites less motivated respondents to simply quit the questionnaire.

Analysis conducted on the whole data set – disregarding content differences between the questionnaires - revealed significant tendency of the respondents who answered more questions to skip the last questions in the survey (Figure 3a).

Figure 3a. Average number of questions answered before (not)answering the last four questions*



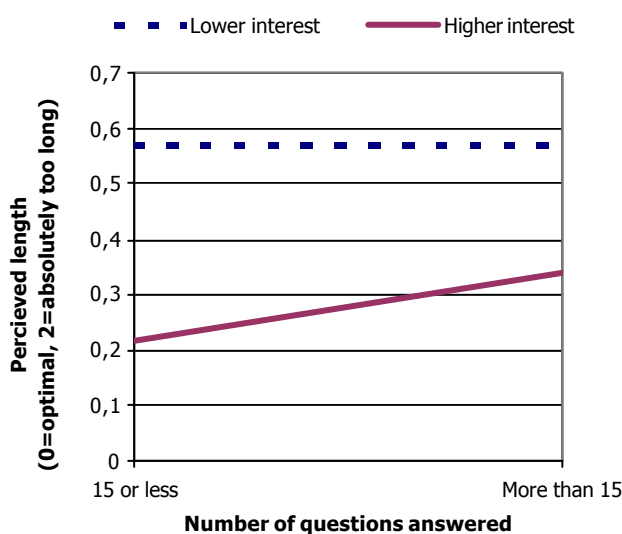
*When checked with t-test for independent samples, all differences appeared statistically significant ($p < .001$).

Our second analysis aimed to investigate the relation between objective questionnaire length – measured in number of questions actually answered - and *subjective impression of (i.e. perceived) questionnaire length*. Respondents expressed their subjective impression of questionnaire length on the following scale: «too short», «optimal», «somewhat too long», «absolutely too long». Nobody said that the questionnaire was too short. The remaining answers were coded as 0, 1 and 2 («optimal» being «0»).

Across all three questionnaire types there was an overall significant, but very small positive correlation between number of questions the respondents answered and their perception of questionnaire length ($r=0.11$, $p<.01$). Perceived length was more strongly correlated to the level of interest for the questionnaire topic ($r=-.26$, $p<.01$). The less interesting questionnaire topic was, the longer was questionnaire perceived to be. Level of interest for the questionnaire topic was not correlated to the number of questions answered ($r=.03$, $p>.05$).

Two-way ANOVA was conducted to check for the joint effect of level of interest and number of questions answered (independent variables) on perceived questionnaire length (dependent variable). The results (Figure 4) show that number of questions answered has effect on perceived length only when the level of interest is high. The respondents who had lower interest in the questionnaire topic judged the questionnaire equally long no matter of actual number of questions answered. One should be cautious, though, not to imply causal effect between interest and perceived length. The relation could be bi-directional.

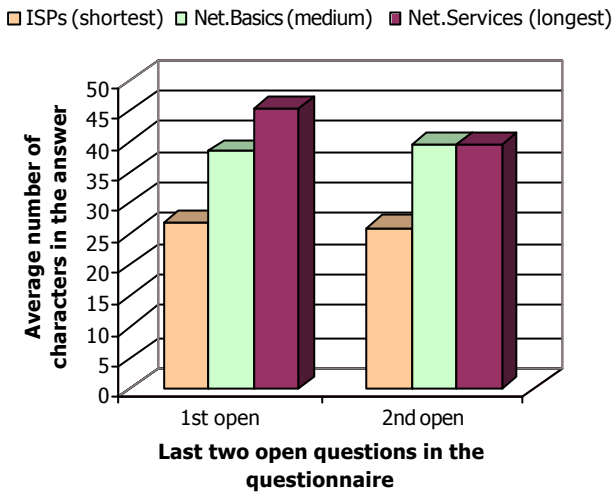
Figure 4. Joint effect of number of questions answered and level of interest on perceived length of the questionnaire*



*Independent effect of number of questions answered significant on $p<.05$; independent effect of level of interest significant on $p<.001$; interaction significant on $p<.05$.

Finally, we examined the relation between questionnaire length and *the amount of answers given to open questions*, sometimes used as a measure of data quality. We checked for the difference in average number of characters written in the answers to open questions within the three questionnaires (Figure 5).

Figure 5. Length of answers on open questions depending on the questionnaire length.*



*Differences significant for both questions on $p < .001$; posthoc Scheffe tests all significant on $p < .05$ except for 2nd open question in Net.Basics vs Net.Services).

Differences between questionnaire types were significant for both open questions ($p < .001$). There was a small, but significant correlation between average number of questions answered so far and the length of answers to both open questions ($r = .11$ for 1st, and $r = .13$ for 2nd open question; both correlations are significant at $p < .01$). This result is non-intuitive and opposite to the results for face-to-face surveys (Johnson et al, 74; Herzog & Bachman, 81). Possible explanation could be that less motivated respondents quitted before arriving to the last two open questions. Those who stayed – especially in longer questionnaires - decided to invest more effort in responding in general, which reflected in their longer answers on open questions.

1.4. Conclusion

Three questionnaires of different length but otherwise similar characteristics were applied on a large online panel. At the end of each questionnaire four equal questions were asked – two closed and two open – which enabled observations about the relation of questionnaire length and some characteristics of data. Length of the questionnaire appeared to be negatively related to unit response rates for the last four questions. On the contrary from usual findings in classic survey methods, length of the questionnaire was positively related to the length of answers given to open questions. Subjective impression of questionnaire length did not depend on its actual length as much as on general interest in its topic.

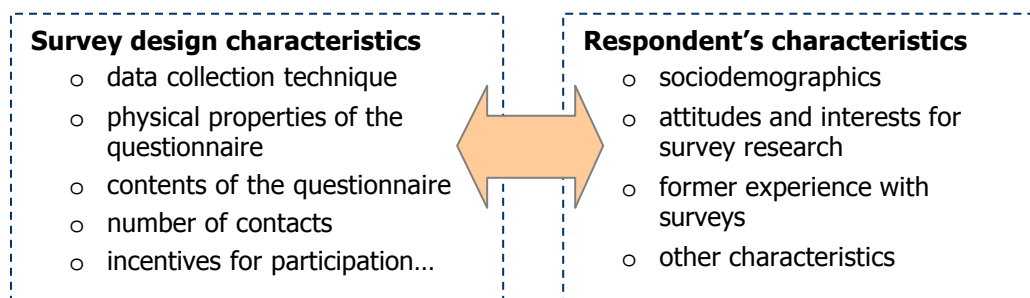
Although only side-observations, these findings open numerous questions and are a good starting point for further research. Future studies should be designed for a specific purpose of testing the effect of questionnaire length on response rates and quality of data in online surveys. But, as many researchers studying effects of survey characteristics on results of «classic» data collection techniques already noted, a theoretical model of survey behavior is needed in order to better explain and generalize the findings.

2. Draft theoretical model of behavior in survey situation

2.1. Survey situation

It is hard to forecast all characteristics of survey situation that can influence the results, especially because their influence can be modified by presence or absence of other factors. There are numerous studies on effects of various characteristics of survey situation, and many authors tried to summarize those results (Boser & Clark, 1995; Craig & McCann, 1978; Dillman, 1978; Fox, Crask & Kim, 1988; Goyder, 1982; Heberlein & Baumgartner, 1978; Kanuk & Berenson, 1975; Klein & Porst, 2000; Linsky, 1975; Yammarino, Skinner & Childers, 1991; Yu & Cooper, 1983, in Helgeson, Voss, & Terpening, 2002). Characteristics of survey situation can be roughly divided in (1) those that are determined by the survey design and (2) those that are related to respondents' characteristics (Figure 6).

Figure 6. Characteristics of survey situation.



In the first group are all those characteristics of a survey design which respondents observe and interact with during the survey process. Those are *data collection technique* (mail, telephone, face-to-face, computer assisted or online survey), *physical properties of questionnaire and other material* (for example number of pages, size and color of the questionnaire, size and form of font, amount of empty space between questions, use of graphical elements, number of questions,...), *content of the questionnaire and related materials* (relevance of the topic for the respondent, social sensitivity of questions, possible threats for respondents' privacy, content of the introductory explanation etc.), *number of contacts with the respondent* (including both contacts preceding and following the receipt of questionnaire), and various *incentives for participation*, which can be material (for example money, lottery tickets, telephone cards, sweepstakes) or non-material (for example possibility of receiving results of the survey). In telephone and face-to-face surveys, another very important factor are *characteristics of the interviewer*.

The second group consists of all those characteristics of sample members that determine the way they will perceive and react to the characteristics of survey design. These are primarily *sociodemographic characteristics* (age, gender, education level, socio-economic status, type

of settlement etc.), *attitudes and interest for surveys in general and for the topic of current survey, previous experience with surveys*, but also *other relevant characteristics*, like fear of being robbed (Groves, Cialdini, & Couper, 1992) or usual way of manipulating received mail (Helgeson i sur., 1992). Those two groups of factors mutually affect one another: respondents' characteristics determine the way they will perceive and react to the characteristics of the survey design, and the latter can be modified in interaction with the respondent (additional contacts, more incentives, different interviewer's behavior or even a different interviewer etc.).

In disproportion with the amount of studies, there are only few consistent findings about the characteristics of an «ideal» survey. The reasons are two-fold. First, researchers often manipulated only one or few factors, while trying to keep the other factors constant (Heberlein i Baumgartner, 1978). Although this seems to enable clearer conclusions, it disregards the complexity of the survey situation, in which the effect of one factor is almost always moderated by the effects of various other characteristics of survey situation. Second, most of the studies are practically inclined, without a theoretical frame that could help in explanation and generalization of the results. Efforts for explaining survey behavior within broader psychological theories are therefore more and more frequent. According to most approaches, the way people will behave in survey situation depends on their **decision** about their participation in the survey and effort to be invested in answering the questions.

2.2. Expected utility models

Existing theories used in explanation of survey behavior (social exchange theory as proposed by Dillman, 1978; “leverage-saliency” theory by Groves, Singer, & Corning, 2000) are in essence based on expected utility model of decision making. There is a wide array of theories that use the concept of expected utility in explanation of behavior in various situations. They appear under different names: attribution x evaluation, expectancy x value, instrumentality x goal etc. Although the beginnings of this approach can be noticed already in works of Pascal (1623-62) and Bernoulli (1738), these «expected utility» theories became popular among psychologists only after von Neumann & Morgenstern's seminal book «Theory of games and economic behavior» (1947). The mathematical formalization of this approach was soon developed within the framework of the subjectively-expected utility model (Edwards, 1954; Savage, 1954). Later the approach has developed in two directions: (1) research on the ways on which people make decisions in artificial situations with known and very limited number of possible actions and outcomes (Edwards, Kahneman & Tversky, Birnbaum), and (2) attempts to explain development of attitudes and behavior in complex everyday situations (led

by so called «Michigan school» whose prominent members were Atkinson and Zajonc, and «Illinois school», with Ajzen and Fishbein).

The second direction is particularly important for the explanation of behavior in survey situation. According to this direction, the potential respondents evaluate the probability/salience of various attributes of survey situation and estimate the importance of each attribute. Then they multiply the importance of each attribute with the degree of its salience (or probability of appearance) in a given survey situation. At the end they combine the results of the previous step to achieve total expected utility of participation in the survey.

2.3. Satisficing

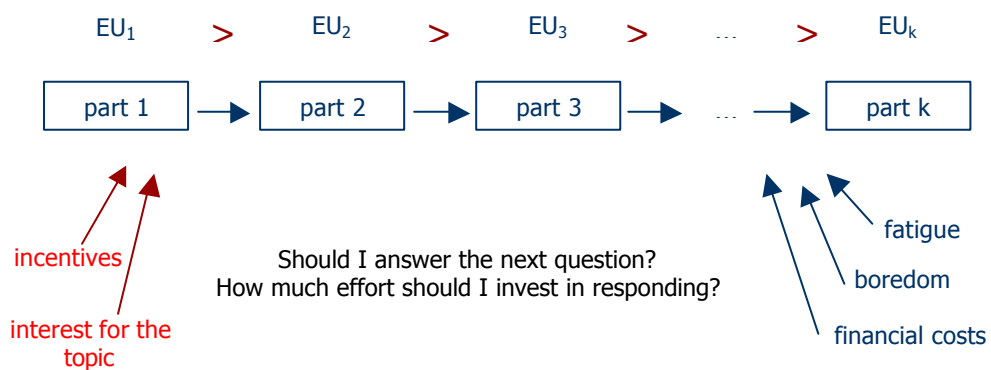
Expected utility models imply that people make decisions in an impractically long and cumbersome way, evaluating salience / probability of occurring and importance of each attribute of a given situation. The cost of such procedure in terms of time and effort would in real life very often be too high compared to the importance of the decision to be made. Simon, within his famous concept of «bounded rationality», suggests that people often do not make optimal, but satisficing decisions. Many authors differ between two basic ways of information processing: systematic and heuristic (Chaiken, 1980), central and peripheral (Petty & Cacioppo, 1979), cognitive and affective (Zajonc, 1980, 1982), controlled and automated (Shiffrin & Schneider, 1977). In situations that are new, complex, and that could have important consequences, people usually invest significant cognitive effort in information processing, trying to evaluate as many attributes of a given situation as possible before they make a decision (this is systematic, central, cognitive, controlled processing). In situations which are known, simple and whose consequences are not so important, people often choose «easier» way of processing (heuristic, peripheral, affective, automatic), relying only on one or several easy available information and using heuristics in choosing and evaluating such information.

Survey situation is usually perceived as such benign and unimportant situation, which is why potential respondents only rarely engage in detailed elaboration of reasons for and against participation in the survey (Groves, Cialdini, & Couper, 1992). More often, they make decisions about their behavior on the basis of only few aspects of survey situation, using heuristics to assess their probability / salience and importance. That means that only one factor of survey situation (for example monetary incentive or attractive visual design of the questionnaire), can be enough for a potential participant to accept invitation to a survey. Other simultaneously present factors (time of day, topic of study etc.) can, when such powerful attribute is present, have almost no influence on participant's decision.

2.4. Dynamic decision making

After making initial decision on participation in the survey, a sequence of additional decisions often occurs during the survey process. For each question, the respondent can decide (1) whether to answer the question, and (2) if yes, what amount of effort should be invested in answering it (Figure 7). The former decision will influence quantity of data in terms of total response rate, drop-out rate and unit response rate. The latter decision will influence quality of data in terms of their correspondence to actual respondent's characteristics/ behaviors. Of course, all the restrictions (satisficing, heuristics) of human decision making mentioned in the last paragraph apply here as well.

Figure 7. Dynamic decision making in survey situation.



If the respondents perceive the attributes of survey situation in the same way during the whole survey process, there is a large probability that they will stay in the process till the end, and keep invested effort on a constant level. If, on the contrary, their perception of survey situation changes (for example because of fatigue, or because they discovered that their initial estimate of questionnaire length or topic relevance was wrong), respondents can modify their initial decision and stop responding or reduce the level of effort invested in responding. In the same way, potential respondent who at first refused participation in a survey can often be convinced to convert this decision – for example, by use of additional monetary incentives.

2.5. Conclusion

Some initial guidelines were presented for the development a theoretical model of survey behavior within a general framework of decision making models. The model should take in account heuristic nature of human information processing in survey situation. It should also reflect dynamical changes in respondents' decisions about their behavior during the survey.

References

- Boser, J.A. (1987). Teacher education follow-up surveys: Variables related to response rate. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.
- Burchell, B. & Marsh, C. (1992). The effect of questionnaire length on survey response. *Quality and Quantity*, 26, 233-244.
- Craig, C.S & McCann, J.M. (1978). Item nonresponse in mail surveys: Extent and correlates. *Journal of Marketing Research*, 15, 285-289.
- Dillman, D.A. (1978). *Mail and telephone surveys: The Total Design Method*. New York: Wiley.
- Edwards, W. (1954). The theory of decision making. U: W. Edwards & A. Tversky (ur.), *Decision making*. Harmondsworth, UK: Penguin Books.
- Fox, R.J., Crask, M.R. & Kim, J. (1988). Mail survey response rate: A meta-analysis of selected techniques for inducing response, *Public Opinion Quarterly*, Vol. 52, 467-491.
- Frankel, J. (1981). The effect of interview length and question type ("effort") on perceived respondent burden. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, CA.
- Groves, R.M. (1989). *Survey errors and survey costs*. New York: Wiley.
- Groves, R.M., Cialdini, R.B., & Couper, M.P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 475-495.
- Heberlein, T. & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review*, 43, 447-462.
- Helgeson, J.G. & Ursic, M.L. (1994). The role of affective and cognitive decision-making processes during questionnaire completion. *Public Opinion Quarterly*, 11(5), 493-510.
- Herzog, A.R. & Bachman, J.G. (1981). Effects of questionnaire length on response quality. *Public Opinion Quarterly*, 45, 549-559.
- James, J.M. & Bolstein, R. (1990). The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Public Opinion Quarterly*, 54, 346-61.
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kanuk, L. and Berenson, C. (1975). Mail surveys and response rates: A literature review. *Journal of Marketing Research*, 12, 440-53.
- Klein, S. & Porst, R. (2000). Mail Surveys: Ein Literaturbericht. *ZUMA-Technischer Bericht*, 10.
- Layne, B.H. & Thompson, D.N. (1981). Questionnaire Page Length and Return Rate. *The Journal of Social Psychology*, 113, 291-292.
- Linsky, A.S. (1975). Stimulating responses to mailed questionnaires: a review. *Public Opinion Quarterly*, 39, 82-101.
- Petty, R.E. & Cacioppo (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, 37, 1915-1926.
- Savage, L.J. (1954). Historical and critical comments on utility. U: W. Edwards & A. Tversky (ur.), *Decision making*. Harmondsworth, UK: Penguin Books.
- Shiffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127-190.
- Yammarino, F.J., Skinner, S.J., & Childers, T.L. (1991). Understanding mail survey response behavior: A meta-analysis. *Public Opinion Quarterly*, 55, 613-639.