Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study

ELISABETH DEUTSKENS *   Ec.Deutskens@mw.unimaas.nl
*Maastricht University, Faculty of Economics and Business Administration, Department of Marketing,
P.O. Box 616, 6200 MD, Maastricht, the Netherlands*

KO DE RUYTER   K.deruyter@mw.unimaas.nl
Maastricht University, Maastricht, the Netherlands

MARTIN WETZELS   M.g.m.wetzels@tm.tue.nl
Eindhoven University of Technology, Eindhoven, the Netherlands

PAUL OOSTERVELD   paul.oosterveld@millwardbrown.nl
Millward Brown/Centrum, Amsterdam, the Netherlands

Abstract

This study examines the effect of the timing of follow-ups, different incentives, length, and presentation of the questionnaire on the response rate and response quality in an online experimental setting. The results show that short questionnaires have a higher response rate, although long questionnaires still generate a surprisingly high response. Furthermore, vouchers seem to be the most effective incentive in long questionnaires, while lotteries are more efficient in short surveys. A follow-up study revealed that lotteries with small prizes, but a higher chance of winning are most effective in increasing the response rate. Enhancing questionnaires with visual elements, such as product images, lead to a higher response quality and generate interesting interaction effects with the length of the questionnaire and the incentives used. Finally, the timing of the follow-up has no significant influence on the response rate.

Keywords: online marketing research, questionnaire design, response rate, response quality

The rapid growth of the Internet has opened new opportunities for collecting and disseminating research information worldwide. Market researchers have long recognized the advantages of Internet-based surveys, the most important of these being lower costs and faster response (e.g., Ilieva et al., 2002). Existing research on online surveys has mainly focused on comparing response rates to traditional mail surveys or combining existing evidence into a meta-analysis of online survey response rates (e.g., Cook et al., 2000; Ilieva et al., 2002; Sheehan, 2001; Shermis and Lombard, 1999). Some authors have started to test the influence of certain format or design parameters (Couper et al., 2001; Dillman et al., 1998; Lozar Manfreda et al., 2002; Sheehan and McMillan, 1999). These studies are often narrow in scope and limit themselves to the assessment of response rates

* Corresponding author.

only. There has been little methodological research around influences on response rates as well as response quality in online surveys. Before online surveys can be used on a large scale for academic as well as market research, it is necessary to experimentally determine the impact of format and design parameters on nonresponse error as well as on response accuracy and questionnaire completeness.

While empirical evidence on response rate and response quality of online surveys is still limited, past research on traditional mail surveys has gathered extensive evidence on the influence of several format and design parameters on response rate and response quality. The literature on mail surveys verifies that the most important factors for maximizing response rate and response quality are follow-up mailings and incentives. In addition to this, the length and presentation of the questionnaire have been examined extensively (Church, 1993; Dillman, 2000; Fox et al., 1988; Heberlein and Baumgartner, 1978; Kanuk and Berenson, 1975; Yammarino et al., 1991; Yu and Cooper, 1983). Yet, the virtual environment of online surveys has added new aspects to the discussion and administration of these factors. We argue that results with regard to offline surveys are not per se generalizable to the online setting. On the one hand, online surveys share a lot of characteristics with traditional mail surveys (e.g., answering questions presented via text) (Kiesler and Sproull, 1986). On the other hand, they differ in the method of contacting, the medium, and the mode of responding (Tourangeau et al., 2000). Thus, the interactive and impersonal nature of the Internet apparently causes differences in the effectiveness of format and design parameters between online and offline surveys. The aim of this article is to assist researchers in selecting the most effective incentives for online surveys, in choosing the appropriate length, and deciding upon the optimal timing of the follow-up contacts as well as on the optimum design of the online survey.

The remainder of the article is structured as follows: Firstly, in order to develop our hypotheses, empirical evidence from traditional mail surveys is combined with scarce empirical evidence from online surveys. Next, we describe the experimental study that we conducted, which aimed to test these hypotheses. A follow-up study provides some additional insights into the effectiveness of lotteries. Finally, we conclude this paper with a discussion and implications.

1.   Literature Review and Hypotheses Development

Previous research has examined a large number of factors that increase response rates and improve data quality. In terms of online surveys, the most important factors to consider are follow-ups, incentives, and length and presentation of the questionnaire (Church, 1993; Dillman, 2000; Fox et al., 1988; Heberlein and Baumgartner, 1978; Kanuk and Berenson, 1975; Yammarino et al., 1991; Yu and Cooper, 1983). In the following paragraphs we will develop our hypotheses based on evidence from mail as well as online survey literature.

1.1. Follow-ups

Follow-up contacts have been consistently reported as being the most powerful technique for increasing response rates, both in mail and online surveys (Dillman, 2000; Fox et al., 1988; Heberlein and Baumgartner, 1978; Schaefer and Dillman, 1998; Yammarino et al., 1991). Dillman (2000) suggests the use of four contacts with a participant, but even single follow-ups have been reported to increase the response rate significantly (Heberlein and Baumgartner, 1978). Sending out multiple follow-ups in an online survey is virtually costless, however, it should be done with great care. Repeated follow-ups have diminishing returns and may be considered as spam, thereby irritating or annoying potential respondents without noticeably increasing response rates (Solomon, 2001).

Sending the follow-up right after the majority of respondents have reacted to the initial mailing has been identified as essential for maximizing the response rate (Dillman, 2000). At the same time, the question of the exact timing between the initial and subsequent request has been left largely unresolved. Essentially, the response rate takes on the form of an inverted U-shape.[1] However, the central question is which distribution characteristics the response curve for online surveys adopts, as this determines the best possible timing of the follow-up. The literature review from Ilieva et al. (2002) reveals that the average response time in online surveys is 5.59 days, which is substantially faster than the 12.21 days in mail surveys. Hence, follow-ups could be conducted earlier in online settings than in traditional settings, for example after one week (Dillman, 2000). As the optimal timing of the follow-up is imperative for maximizing the response rate, we hypothesize that a follow-up, which is sent right after the majority has responded to the initial mailing, will increase the overall response rate in an online survey compared to a later follow-up.

With respect to response quality, neither research on mail surveys nor experiential or intuitive reasoning suggests an effect of the timing of follow-ups on response quality. Therefore, we do not expect, a priori, a direct effect of the timing of follow-ups on the response quality of an online survey. Hence we hypothesize:

H_{1a}: Earlier follow-ups will increase response rates compared to later follow-ups.
H_{1b}: The timing of the follow-up will not influence response quality.

1.2. Incentives

In addition, the use of monetary incentives in general, and small prepaid financial incentives in particular, have been declared as being effective in increasing the response rate in offline and online surveys (Church, 1993; Dillman, 2000; Fox et al., 1988; Heberlein and Baumgartner, 1978; Kanuk and Berenson, 1975; Yammarino et al., 1991; Yu and Cooper, 1983). However, the intangible nature of the Internet has raised new (administrative) problems, as cash incentives cannot be attached to a virtual questionnaire. Therefore, empirical research is needed to identify alternative online incentive systems, such as vouchers, lotteries or donations, and examine their effect on response rates. As empirical evidence from mail surveys verifies that cash is more effective than lotteries or charitable donations (Furse and Stewart, 1982; Warriner et al., 1996), we hypothesize that

the response rate in an online survey will be highest if vouchers, the closest online equivalent to cash, are used as an incentive, followed by lotteries and finally donations.

Self-perception theory of respondent behavior stipulates that if external cues, such as incentives, are present respondents feel less committed and hence provide lower quality responses (Hansen, 1980). However, the mail literature does not provide evidence that the response quality varies for specific types of incentives. As there is no evidence (or a logical reason) to assume that any of the three incentives will affect response quality, we hypothesize that the response quality is the same regardless of whether vouchers, lotteries, or donations are used as an incentive. Therefore, we postulate:

H$_{2a}$: Response rates will be highest if vouchers are used as an incentive, followed by a lottery and finally donations.

H$_{2b}$: Response quality will be equal for vouchers, lotteries or donations.

1.3. Length

The relationship between length and response rate and quality may be different for Internet-based research, as it is not clear what 'long' means in the online environment. Rosenblum (2001), for example, indicates that online surveys should consist of approximately 20 questions, which would generally be considered too short for substantial market and academic research. Therefore, we need to investigate whether response rate and response quality are indeed lower for longer online surveys. In terms of mail surveys, the length of the questionnaire has been repeatedly investigated (Dillman, 2000; Fox et al., 1988; Heberlein and Baumgartner, 1978; Kanuk and Berenson, 1975; Yammarino et al., 1991; Yu and Cooper, 1983). Common sense suggests that longer questionnaires will obtain lower response rates than shorter questionnaires, as they demand more time from the respondent. Likewise, one would also assume a negative relation between the response quality and length of the questionnaire. However, the literature about traditional mail surveys provides mixed results. Whereas several studies show that survey length does not influence response (Linsky, 1975; Yu and Cooper, 1983), a number of studies reveal that there is a negative relation between survey length and response rate as well as response quality (cf. Heberlein and Baumgartner, 1978; Yammarino et al., 1991). Therefore, we hypothesize:

H$_{3a}$: Short questionnaires will increase response rates compared to long questionnaires.

H$_{3b}$: Short questionnaires will exhibit a higher response quality compared to long questionnaires.

1.4. Presentation of the Questionnaire

Previous research on paper-and-pencil surveys suggests that the design of the questionnaire may be extremely important in obtaining unbiased answers from respondents, as respondents evaluate both the verbal and the visual elements of the questionnaire. Dillman (2000) states that "respondent-friendly" design improves mail survey response, but

that there is a lack of agreement on what exactly constitutes respondent friendly design. Research on specific graphical aspects, such as color of the questionnaire, does not provide clear results about their effect on response rate and response quality (Fox et al., 1988; Kanuk and Berenson, 1975).

The Internet has added a new dimension to the design of questionnaires as it offers a wide area of new design opportunities. Simple questionnaires can, for example, be enhanced with cinematic and interactive images, such as 3D presentations of products. This results in web questionnaire design challenges for the researcher. On the one hand, pictures may enhance the attractiveness of the survey and may make it a more enjoyable experience for the respondent. In this event, it is likely that the response quality would be higher for surveys that are graphically enhanced, as respondents enjoy the process of filling in the questionnaire and therefore put in more effort and answer more seriously. On the other hand, these advanced features make the questionnaire more difficult to access and complete and lead to longer download times, which could consequently reduce the response rate (Dillman, 2000). Therefore, we hypothesize:

H$_{4a}$: Visually enhanced questionnaires will decrease response rates compared to text-based questionnaires.

H$_{4b}$: Visually enhanced questionnaires will improve response quality compared to text-based questionnaires.

2.   An Empirical Study

2.1.   Research Setting

The questionnaire used in this research was a standardized, multi-client attitude and usage study, which has been conducted in the Netherlands via a traditional mail surveys since 1991. The questionnaire contained 19 product categories (such as beer, shampoo, cheese, olive oil and toast) in the long version and 9 product categories in the short version of the survey. Every product category included a number of brands for which respondents had to indicate whether they are familiar with the brand, how often they use it and how they perceive the price and quality of this brand. Participants in this experiment were recruited from a database, which consists of names and e-mail addresses from people who had participated in a prior telephone survey and consented to being contacted again for other research purposes. Respondents received an e-mail invitation for this research, which included a short introduction to the study, a request to participate and the hyperlink to the web questionnaire. With one click on this link, respondents were directed towards the questionnaire. By using a unique ID for each respondent, respondents did not have to fill out a username and password before entering the questionnaire.

2.2.   Independent Variables

To test the hypotheses, we designed a between-subjects, fixed effects factorial design with four factors. These were: type of incentive (voucher, lottery, donation), length of the
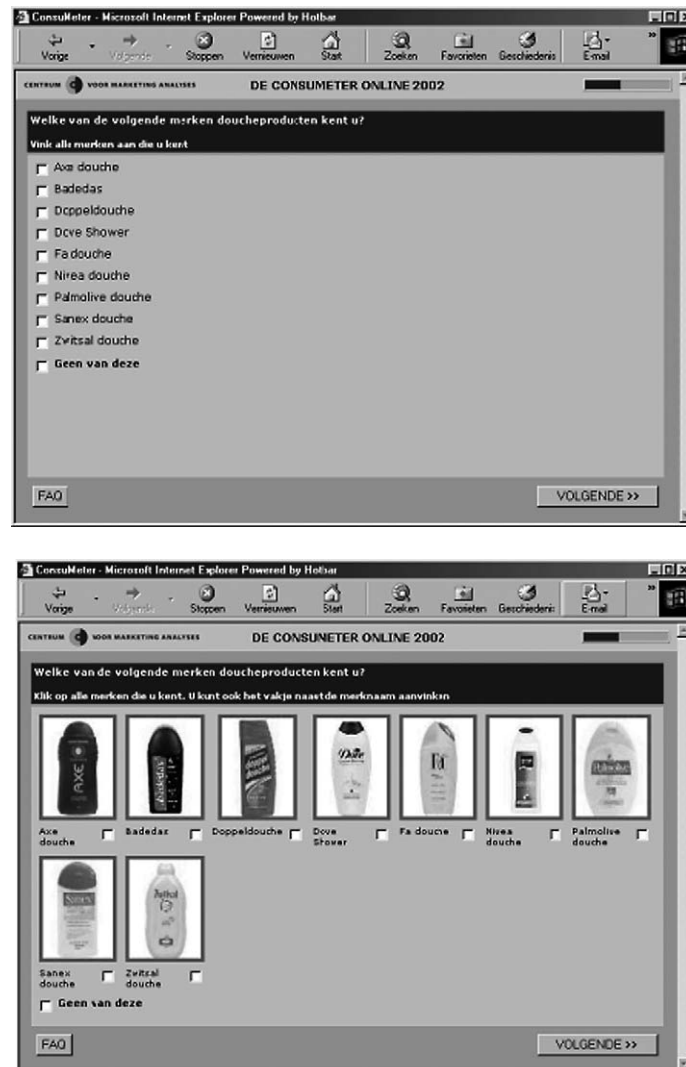
Figure 1.   Textual and Visual Version of the Questionnaire.

questionnaire (long, short), presentation of the questionnaire (textual, visual), and timing of the follow-up (early, late). This resulted in a $3 \times 2 \times 2 \times 2$ full factorial design with 24 different cells.

First, we used three different incentives in this experiment: *Vouchers*: a 2€ voucher in the short version, or a 5€ voucher in the long version for an online book and CD store. We opted for an online book and CD store, as "books and CD's continue to be the most popular items bought online for the 3rd consecutive year" (Greenspan, 2002). *Charitable*

donations: we noted the overall amount that we would donate if everybody in the survey participated (500€). Respondents could chose between the World Wide Fund for Nature (WWF), Amnesty International, or a cancer association. *Lottery*: respondents had the chance of winning one of 5 vouchers of 25€ in the short version and 50€ in the long version for an online book and CD store. The second independent variable was *length* of the questionnaire, where the long version took approximately 30–45 minutes to finish, while the short version could be completed within 15 to 30 minutes. Thirdly, the presentation of the questionnaire was either *textual* or *visual*. To utilize the design possibilities on the Internet, we presented one random half of the respondents with only the name of the product, while the other half of the respondents obtained in addition to the brand name also a picture or logo of the product (see Figure 1). Finally, we randomly divided participants, who did not fill-in the questionnaire after the initial invitation, into two groups, an *early* and a *late* follow-up group. Based on initial empirical evidence that the mean response time to online surveys is 5.59 days (Ilieva et al., 2002), the 'early' group received the reminder after one week. The 'late' group received the follow-up after two weeks, as suggested by the mail literature (Dillman, 2000).

2.3. *Dependent Variables*

The two dependent variables that were analyzed in this study were response rate and response quality. Response rate was defined as the percentage of the contacted sample that has answered and returned the questionnaire. We only considered the net response rate, i.e. the percentage of questionnaires that actually reached the respondent, thus excluding undeliverable e-mails.

With respect to response quality, two specific components of data quality were examined, namely the completeness and accuracy of respondent answers (Goetz et al., 1984; Hansen, 1980; McDaniel and Rao, 1980). Since no open-ended questions were used in this survey, completeness was assessed by the number of non-opinions ("don't know" answers) given by each respondent (Goetz et al., 1984) and the number of semi-completed questionnaires. "Don't know" answers were asked after respondents had indicated that they know a product, but were not able or willing to answer the price, quality and usage questions with respect to that product. Thus, "don't know" answers were considered undesirable, as respondents were either inconsistent with previous answers, or they wanted to speed up the process by indicating "don't know." Semi-completed questionnaires refer to questionnaires where respondents started to fill them in, but did not complete them.

The second aspect of response quality was bias or inaccuracy (Goetz et al., 1984). As early as 1978, Hansen and Scott noted that "bias" is an umbrella term that covers a broad range of quality problems. Some authors look at bias as a deviation from a " 'known' and presumably truthful response" (Hansen, 1980, p. 79). Yet, Hansen (1980) suggested in his JMR article, where he identified a negative link between incentives and response quality, that the distribution or summary of responses of one subgroup should be compared with those of another subgroup. Therefore, we examined the response distribution (Hansen, 1980) by comparing means and variances of the answers, as this method is especially

appropriate "when different methods are used in an attempt to stimulate response to a survey" (p. 79).

3. Results

Data collection took place from mid April until the end of May 2002. The profile of the respondents shows that there is a balance in gender, with 51.5% male and 48.5% female respondents. Younger and higher educated people are somewhat over-represented. For example, 36.2% are between 15 and 34 years old, 40.0% are in the range of 35–49 years, still 22.1% are 50–64 years, but only 1.8% of the respondents are 65–74 years old. With respect to education, 50.4% completed higher education, 42.6% got a medium level of education, and only 5% lower education. After only three days, we had received more than half of the final responses (52.9%). On average, it took respondents 6.6 days to complete and return the questionnaire, which is slightly higher than the 5.59 days reported by Ilieva et al. (2002). Interestingly, we observed that respondents in the lottery group responded with 5.7 days significantly faster than the donation group with 6.7 days and the voucher group with 7.4 days ($F(2,727) = 3.81$, $p = 0.023$). In addition, we employed the Games–Howell procedure for multiple comparisons of the response time (Games and Howell, 1976) and found that the voucher and lottery group differed significantly ($t = 2.84$, $p = 0.013$).

In total, 5413 e-mails were sent out. One third of those e-mails (1836) were undeliverable, which left us with 3577 usable e-mails, from which we received 730 completed questionnaires. This yielded a net response rate of 20.4%. The response rate in the online survey differed significantly between the different cells of the design ($\chi^2(23) = 73.97$, $p < 0.001$; $G^2(23) = 77.24$, $p < 0.001$). The short, visual version of the questionnaire with lottery as incentive and a late reminder had the highest response rate with 31.4%, while the long, visual version with donation to charity as incentive and an early reminder only had a response of 9.4%. Table 1 summarizes the response rates for the main effects.

In order to rule out non-response error, we carried out a time trend extrapolation test. The assumption is that respondents who respond less readily are similar to non-respondents (Armstrong and Overton, 1977). The early respondents included all those who filled in the questionnaire on the first day after they had received the e-mail invitation, whereas the late-respondent group consisted of the people who responded after 10 days or later. Only a negligible number of variables (less than 1%) used in the questionnaire showed a significant difference between early and late respondents. Those variables were distributed among all items of the questionnaire so that no consistent pattern could be discerned. Therefore, we may conclude that our data did not suffer from nonresponse problems.

3.1. Follow-ups

In line with our expectations, the early follow-up had a higher response rate (21.2%) than the late follow-up (19.5%). However, this difference was not significant ($\chi^2(1) = 1.70$,

Table 1. Response Rates for the Different Format and Design Parameters

Factors							
Timing of the follow-up		Type of incentive		Length of the questionnaire		Presentation of the questionnaire	
Early	21.2%	*Voucher*	22.8%	*Long*	17.1%	*Textual*	21.9%
Late	19.5%	*Donation*	16.6%	*Short*	24.5%	*Visual*	19.0%
Lottery	22.8%						
Testing hypotheses							
$\chi^2(1) = 1.70$ $(p = 0.192)$		$\chi^2(2) = 20.26$ $(p < 0.001)^a$		$\chi^2(1) = 29.94$ $(p < 0.001)$		$\chi^2(1) = 4.55$ $(p = 0.033)$	
$G^2(1) = 1.70$ $(p = 0.192)$		$G^2(2) = 20.63$ $(p < 0.001)^b$		$G^2(1) = 29.72$ $(p < 0.001)$		$G^2(1) = 4.54$ $(p = 0.033)$	

[a]Pearson chi-squared statistic.
[b]Likelihood-ratio statistic.

$p = 0.192$; $G^2(1) = 1.70$, $p = 0.192$) and hence, H_{1a} was not supported. As hypothesized, we did not find a significant difference in the response quality for the early and the late follow-up. Neither the analysis for completeness, nor for accuracy, revealed significant differences at $\alpha = 0.05$. Thus, H_{1b} was supported.

3.2. Incentives

Vouchers and lotteries with 22.8% had a significantly different response rate ($\chi^2(2) = 20.26$, $p < 0.001$; $G^2(2) = 20.63$, $p < 0.001$) to the donation to a charity group with 16.6% (where 61% choose for the cancer association, 25% for the WWF, and 15% for Amnesty International, adjusted residual = |4.50|). Thus, H_{2a} was only partially supported. As hypothesized, donations had a lower response rate than both vouchers and lotteries. However, the overall response rate of vouchers was not higher than the response rate of lotteries. ANOVA revealed that there were no significant differences in the number of "don't know" answers at $\alpha = 0.05$, and also the number of semi-completed questions did not vary between the different incentive groups ($\chi^2(2) = 2.69$, $p = 0.261$; $G^2(2) = 2.67$, $p = 0.264$). In addition, the analysis of accuracy did not show significant differences at $\alpha = 0.05$. As we did not find a discrepancy in the response quality for the different types of incentives, H_{2b} was supported.

3.3. Length

As expected, the short version of the questionnaire had a significantly higher response rate ($\chi^2(1) = 29.94$, $p < 0.001$; $G^2(1) = 29.72$, $p < 0.001$) than the long version of the questionnaire, with 24.5% and 17.1%, respectively. Hence, H_{3a} was supported. The analysis of the number of "don't knows" in the long and the short version revealed that

there are proportionally more "don't know" answers in the long version at $\alpha = 0.05$. Also the analysis of the number of semi-completed questionnaires showed that there are more semi-completes in the long version than in the short version (220 semi-completed questionnaires in the long version versus 131 in the short version; $\chi^2(1) = 22.58$, $p < 0.001$; $G^2(1) = 22.79$, $p < 0.001$). Furthermore, respondents stopped relatively earlier in the long version (after 41.56%) than in the short version (54.29%) of the questionnaire ($p < 0.001$). For the second aspect of response quality, bias or inaccuracy, ANOVA did not reveal a significant difference in means and variances at $\alpha = 0.05$ for the length of the questionnaire. Aggregating the findings from the completeness and accuracy analysis, we found partial support for H$_{3b}$.

3.4. *Presentation of the Questionnaire*

The visual presentation of the questionnaire had, with 19.0%, a significantly lower response ($\chi^2(1) = 4.55$, $p = 0.033$; $G^2(1) = 4.54$, $p = 0.033$) than the textual presentation with 21.9% and therefore H$_{4a}$ was supported. In terms of presentation of the questionnaire, 5 out of the 19 product categories produced a significant difference at $\alpha = 0.05$ in the number of "don't know" answers. Respondents in the textual group indicated more often "don't know" than respondents in the visual group. In the analysis of the number of semi-completed questionnaires in the visual and textual version of the questionnaire, we did not find a significant difference ($\chi^2(1) = 0.551$, $p = 0.458$; $G^2(1) = 0.551$, $p = 0.458$), although participants in the long version of the questionnaire stopped earlier in the visual version (45.39% answered questions versus 52.13%), whereas respondents in the short version of the questionnaire stopped earlier in the textual version (56.95% versus 60.63%). Comparing the means and variances for the visual and textual version did not show a significant difference at $\alpha = 0.05$. Hence, the evidence for H$_{4b}$ was equivocal. Respondents in the textual version indicated more often "don't know," but there were no significant differences in the means and variances and the analysis of the semi-completes did not reveal a main effect for the presentation of the questionnaire.

3.5. *Interaction Effects*

In the previous sections we used contingency table analysis for testing the main effects of the factors in our study (Everitt, 1992). To further refine our results, we also tested for interaction effects of incentives, timing of follow-ups, length and presentation of the questionnaire, on response rate and response quality. For testing the impact on the response rate, we employed logit modeling (Agresti, 1990; Everitt, 1992). We used the HILOGLINEAR procedure in SPSS release 11.0 using a backward elimination approach. The partial association tests indicate that two two-factor interactions are significant at $\alpha = 0.05$, i.e. INCENTIVE · LENGTH and INCENTIVE · PRESENTATION. In the short version, lotteries were most effective (27.9%), while vouchers and donations had with 23% the exact same response rate. In the long version, vouchers performed best (22.6%), followed by

lotteries (18.8%) and donations (12.7%). With respect to the presentation of the questionnaire, we see that in the textual version of the questionnaire, the response rates are about the same (lottery: 22.9%, voucher 22.4% and donations 20.5%), while the response rate for the donation group drops to 14% in the visual version. In addition, we used the LOGLINEAR procedure implemented in SPSS release 11.0 to conduct a hierarchical logit analysis using nested models. Our results suggested that omitting the two-factor interactions terms from the model significantly decreased the fit of the model ($\chi^2(18) = 26.34$, $p = 0.092$; $G^2(18) = 26.61$, $p = 0.087$; $\Delta\chi^2(9) = 21.47$, $p = 0.011$, $\Delta G^2(9) = 21.74$, $p = 0.009$). Inspection of the z statistic for the individual parameter estimates confirmed that the two-factor interaction terms (INCENTIVE · LENGTH) and (INCENTIVE · PRESENTATION) were significant at $\alpha = 0.05$. Our final model retained the "main" effects (INCENTIVE) and (LENGTH) and the two-factor interaction terms (INCENTIVE · LENGTH) and (INCENTIVE · PRESENTATION). Our results suggest a good fit to the data ($\chi^2(16) = 10.92$, $p = 0.814$; $G^2(16) = 10.94$, $p = 0.813$).

The interaction effects of response quality were analyzed with the GLM Univariate procedure to carry out a series of ANOVAs. However, our analysis did not reveal significant interaction effects for the means and variances at $\alpha = 0.05$. Also the number of "don't knows" and semi-completes were not significantly different at $\alpha = 0.05$ between the different treatment groups.

3.6. Additional Study on Incentives

Our results indicate that lotteries present a very interesting new form of incentives. Participants in the lottery group of the first experiment had the chance of winning one of 5 vouchers of 25€ in the short version and 50€ in the long version of the questionnaire for an online book and CD store. When selecting the prize type and structure, we based our choice on evidence from the mail literature as well as best practice from a large marketing research agency. Yet, it could also be examined whether an attractive commodity, such as a DVD player, as a first prize in the lottery would be preferred to an online voucher. To further explore whether we used the most effective lottery type, we conducted a follow-up study where we focus on one of our most interesting findings.[2] In total, 497 students were contacted via e-mail and asked to participate in an online survey on the perceived quality of examination facilities at a large university. The questionnaire contained 4 open-ended questions and 40 closed questions, which were evaluated on a 7-point Likert scale ranging from "totally disagree" (1) to "totally agree" (7). Respondents were randomly assigned to one of the lotteries, where they were informed that they had the chance of winning either one out of ten vouchers of 25€, one out of five vouchers of 50€ or one DVD player. By focusing on lotteries only, we manipulated the values of the lotteries and separated it from the length of the questionnaire. Hence, we could ensure that the overall expected monetary value between all treatment groups was the same.

The overall response rate was 28.2%. The final sample consisted of 42.7% female and 57.3% male respondents. With respect to the effectiveness of the different types of lotteries, our findings indicated that the lottery with 10 · 25€ had a significantly higher response rate

(35.5%) than the lottery with $5 \cdot 50€$ (26%) and the DVD player (22.5%) ($\chi^2(2) = 7.43$, $p = 0.024$; $G^2(2) = 7.35$, $p = 0.025$). In terms of response quality, the means and variances of the closed questions were analyzed, as well as the number of words and arguments that respondents used in their answers to open-ended questions. No significant differences at $\alpha = 0.05$ could be identified for the variances of the closed questions. However, the means were significantly different for three questions, where two times the mean of the DVD-player as incentive was significantly lower than the mean of the 25€ version ($\alpha = 0.050$ and 0.035, respectively), and once the DVD was significantly higher than the 25€ and 50€ versions ($\alpha = 0.005$). No significant differences were found in the variances of the number of arguments and the variances and means of the number of words. However, for two open-ended questions, the mean number of arguments differed significantly ($\alpha = 0.03$ and 0.02, respectively), where the mean was lowest for the DVD player as an incentive. Thus, lotteries with smaller prizes but a higher chance of winning are most effective in increasing the response rate and are also favorable in terms of response quality.

4. Discussion and Implications

This article examines the effect of several design parameters on response rate and response quality in online surveys. With respect to the optimal timing for sending the follow-up, our results do not give a clear answer. There seems to be a slight preference for sending the follow-up after one week instead of after two weeks, although the results were not significant. Nevertheless, we propose that reminders are sent quite early in order to fully utilize the fast turnaround times of online questionnaires in comparison to mail surveys. Our results indicate, in line with previous research, that monetary interests (vouchers and lotteries) are stronger than altruistic appeals (donations) in increasing the response rate of online surveys. This effect is especially important for long questionnaires. Respondents in the long version demanded an incentive that seemed to reimburse them for the time and effort they devoted to filling in the questionnaire, as can be seen in the remarkably low response rate of 12.7% for the donation group. Furthermore, respondents in the short version appeared to be more risk-taking, as lotteries were by far the most effective type of incentive. Interestingly, vouchers and donations received the exact same response rate in the short version, which could be due to the fact that respondents in the short version had to devote less time and effort and were hence more willing to act altruistically and accept incentives that did not compensate themselves. Overall, we propose that lotteries are probably the most effective reward in an online environment, as they lead to the highest response rate in the short version and still a respectable response in the long version, while being much more cost-efficient than vouchers. In addition, the response time in the lottery groups was more than 1.5 days faster than in the voucher group, as respondents in the lottery might have inferred that by responding early, they have a higher chance of winning a price. To explore the effectiveness of lotteries further, we conducted a follow-up study that examined whether an attractive commodity such as a DVD-player would be preferred to different levels of online vouchers. In line with our prior reasoning, we found

a significantly lower response rate for the DVD player and a slight difference in response quality. Hence, a lottery with smaller prizes but a higher chance of winning is most effective. The effectiveness of using a commodity such as a DVD player is likely to depend on the attractiveness as well as usefulness of that prize to consumers.

As hypothesized, the short version of the questionnaire had a higher response rate than the long version. However, the myth that online surveys should not contain more than 20 questions (Rosenblum, 2001) may be corrected, as even the short version with a response time between 15–30 minutes was still relatively long for an online survey. Hence, response rates of almost 25% in the short version and still around 17% in the long version illustrate that even relatively long questionnaires yield a considerable response rate. Most importantly, the length of the questionnaire did not have a negative effect on the quality of responses. Only the number of "don't knows" and semi-completes were slightly higher in the long, and, in particular, in the long, visual version of the questionnaire. Thus, it seems feasible to conduct long and elaborate surveys via the Internet, especially when respondents are adequately rewarded.

Examining the main effects of the presentation of the questionnaire on the response rate, we find that the response rate is significantly lower for the visual than for the textual version of the questionnaire, although this difference is relatively small. However, when incorporating interaction effects, the main effect becomes insignificant and hence the interaction effects are much more meaningful. The analysis of the number of semi-completes revealed that participants in the long version of the questionnaire abandoned the questionnaire significantly earlier in the visual version, while respondents in the short version of the questionnaire stopped earlier in the textual version. This may suggest that respondents in the long and visual version of the questionnaire were faced with long downloading times due to the large number of pictures, which in turn had a negative impact on their motivation to continue with the questionnaire and hence on their response rate. For the short and visual version on the other hand, the download time did not seem to be problematic, as the short-visual-lottery condition had the highest overall response rate. In any case, as computer and Internet connection speed increases in the future, the negative effect of longer downloading times of visual presentations will diminish over time. Interestingly, the effectiveness of the different types of incentives differed largely between the respondents in the textual and visual version. Whereas all incentives performed almost equally well in the textual version, our results indicated that in the visual version, the response rate for the donation grouped dropped considerably from 20.5% to only 14%. This supports our assumption that respondents in the visual version were subject to longer downloading times, and therefore demanded an incentive that compensated for the time and effort they spent on the questionnaire. However, respondents who had completed the visual version of the questionnaire answered less often with "don't know" than respondents in the textual version. It seems that once respondents recognize a product on the basis of the name and the picture, they have a more active association of the relevant features of the product and hence are better able to give their opinion about the price and quality of a product.

4.1. Limitations and Future Research Guidelines

Future research on Internet-based surveys should be directed at confirming the effects of the timing of follow-ups, incentives, presentation and length of the questionnaire across different questionnaires and populations. Our studies were conduced in the Netherlands, which is especially suitable for this type of research in Europe, since according to Nielsen Netratings (NUA Internet Surveys, 2003) it is rated one of the most mature Internet markets (together with Sweden, Hong Kong, and Australia). Furthermore, the Netherlands is the fourth country in terms of connection Internet Speed on home PC (Hupprich and Bumatay, 2002), outperforming countries such as the US and UK. Hence, the findings from this article serve as a good signal of how incentives, visual presentation, the timing of the follow-up, and length of the questionnaire will influence the response rate and response quality in countries where the Internet penetration is still lower.

For this study, we had access to a large database, which contained e-mail addresses from respondents to an earlier telephone survey. Further research should test whether our findings hold if a sample is recruited differently, for example with the help of an online panel, through site intercepts or pop-ups, or by using a multimode strategy. In our study, we opted to use incentives in all cells of the experiment, as we felt that the length and complexity of the questionnaire necessitated some form of compensation for the respondents, especially as the research was very general and participants did not associate themselves with the topic of the questionnaire. Hence, the experiment should be repeated with a no-incentive group in a context where respondents hold a relation to the research, as for example in a customer satisfaction survey, where incentives are presumably not necessary and could lead to a response rate that is only marginally lower than for the incentive group. In our experiment, respondents in the lottery group were not notified about the number of participants in the lottery so that they were unable to calculate the expected value of the prize. Given that online surveys are usually sent to many people, respondents in the lottery condition might have perceived their chances of winning, and thus the expected value of their prize, as much lower. Future studies should ensure that respondents receive sufficient information to calculate the expected value and whether it is equal across the different incentive groups.

Caution should be exercised in applying the results of this paper to other types of studies where respondents might have different motivations for responding, for example surveys about the attitudes, opinions, and lifestyles of online respondents. In addition, the insignificant findings with respect to the timing of the follow-ups could indicate that neither the early follow-up after one week, nor the late follow-up after two weeks were ideal. Therefore, we suggest that future studies focus on this aspect and re-examine the optimal timing of follow-ups in online surveys. Insights into these issues will advance the knowledge base about online surveys and thereby help to empirically assess the potential of the Internet to supplement traditional means of conducting research by ensuring an adequate response rate and response quality.

Notes

1. We would like to thank one anonymous reviewer for raising this interesting point.
2. We would like to thank both the editor as well as two anonymous reviewers for the suggestion that we should explore the effectiveness of incentives in more depth.

References

Agresti, Allan. (1990). *Categorical Data Analysis*. New York: Wiley.

Armstrong, J. Scott and Jerry S. Overton. (1977). "Estimating Nonresponse Bias in Mail Surveys," *Journal of Marketing Research*, 14(3), 396–402.

Church, Allan H. (1993). "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis," *Public Opinion Quarterly*, 57(1), 62–79.

Cook, Colleen, Fred Heath, and Russell Thompson. (2000). "A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys," *Educational & Psychological Measurement*, 60(6), 821–836.

Couper, M. P., M. Traugott, and M. Lamias. (2001). "Web Survey Design and Administration," *Public Opinion Quarterly*, 65(2), 230–253.

Dillman, Don A. (2000). *Mail and Internet Surveys, The Tailored Design Method*. New York: Wiley.

Dillman, Don A., Robert D. Tortora, John Conradt, and Dennis Bowker. (1998). "Influence of Plain vs. Fancy Design on Response Rates for Web Surveys," *Joint Statistical Meetings*, Dallas, Texas.

Everitt, Brian S. (1992). *The Analysis of Contingency Tables*. London: Chapman and Hall/CRC.

Fox, Richard J., Melvin R. Crask, and Jonghoon Kim. (1988). "Mail Survey Response Rates," *Public Opinion Quarterly*, 52(4), 467–491.

Furse, David H. and David W. Stewart. (1982). "Monetary Incentives versus Promised Contribution to Charity: New Evidence on Mail Survey Response," *Journal of Marketing Research*, 19(3), 375–380.

Games, Paul A. and James F. Howell. (1976). "Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study," *Journal of Educational Statistics*, 1(2), 113–125.

Goetz, Edward G., Tom R. Tyler, and Fay Lomax Cook. (1984). "Promised Incentives in Media Research: A Look at Data Quality, Sample Representativeness, and Response Rate," *Journal of Marketing Research*, 21(2), 148–154.

Greenspan, Robyn. (2002). "E-Shopping Around the World." Retrieved August 20th, 2002, from the World Wide Web: *http://cyberatlas.internet.com/markets/retailing/article/0,,6061_1431461,00.html*

Hansen, Robert A. (1980). "A Self-Perception Interpretation of the Effect of Monetary and Nonmonetary Incentives on Mail Survey Respondent Behavior," *Journal of Marketing Research*, 17(1), 77–83.

Hansen, Robert A. and Carol A. Scott. (1978). "Alternative Approaches for Assessing the Quality of Self Report Data." In Keith Hunt (ed.), *Advances in Consumer Research*, Vol. 5. Chicago: Association for Consumer Research, 99–102.

Heberlein, Thomas A. and Robert Baumgartner. (1978). "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature," *American Sociology Review*, 43(4), 447–462.

Hupprich, Laura and Maria Bumatay. (2002). "Hong Kong Leads the World in High-Speed Internet Connections, According to Nielsen/Netratings' Global Internet Trend Survey," *Nielsen/NetRatings*. Retrieved May 27th, 2003, from the World Wide Web: *http://www.nielsen-netratings.com/pr/pr_020815.pdf*

Ilieva, Janet, Steve Baron, and Nigel M. Healey. (2002). "Online Surveys in Marketing Research: Pros and Cons," *International Journal of Market Research*, 44(3), 361–382.

Kanuk, Leslie and Conrad Berenson. (1975). "Mail Survey and Response Rates: A Literature Review," *Journal of Marketing Research*, 12(4), 440–453.

Kiesler, Sara and Lee S. Sproull. (1986). "Response Effects in the Electronic Survey," *Public Opinion Quarterly*, 50(3), 402–413.

Linsky, Arnold S. (1975). "Stimulating Responses to Mailed Questionnaires," *Public Opinion Quarterly*, 39(1), 82–101.

Lozar Manfreda, Katja, Zenel Batagelj, and Vasja Vehovar. (2002). "Design of Web Survey Questionnaires: Three Basic Experiments," *Journal of Computer-Mediated-Communication*, 7(3). Retrieved January 20th, 2003, from the World Wide Web: *http://www.ascusc.org/jcmc/vol7/issue3/vehovar.html*

McDaniel, Stephen W. and C. P. Rao. (1980). "The Effect of Monetary Inducement on Mailed Questionnaire Response Quality," *Journal of Marketing Research*, 17(2), 265–268.

NUA Internet Surveys. (2003). "Nielsen Netratings: Global Net Population Increases." Retrieved May 27th, 2003, from the World Wide Web: *http://www.nua.ie/surveys/index.cgi?f=VS&art_id=905358729&rel=true*

Rosenblum, Jeff. (2001). "Give and Take," *Quirk's Marketing Research Review*. Retrieved June 14th, 2002, from the World Wide Web: *http://www.quirks.com/articles/article_print.asp?arg_articleid=705*

Schaefer, David R. and Don A. Dillman. (1998). "Development of a Standard E-Mail Methodology," *Public Opinion Quarterly*, 62(3), 378–397.

Sheehan, Kim B. (2001). "E-mail Survey Response Rates: A Review," *Journal of Computer-Mediated-Communication*, 6(2). Retrieved January 15th, 2003, from the World Wide Web: *http://www.ascusc.org/jcmc/vol6/issue2/sheehan.html*

Sheehan, Kim B. and S. J. McMillan. (1999). "Response Variation in E-Mail Surveys: An Exploration," *Journal of Advertising Research*, 39(4), 45.

Shermis, Mark D. and Danielle Lombard. (1999). "A Comparison of Survey Data Collected by Regular Mail and Electronic Mail Questionnaires," *Journal of Business & Psychology*, 14(2), 341–354.

Solomon, David J. (2001). "Conducting Web-Based Surveys," *Practical Assessment Research & Evaluation*, 7(19). Retrieved January 9th, 2002, from the World Wide Web: *http://ericae.net/pare/getvn.asp?v=7&n=19*

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. (2000). *The Psychology of Survey Response*. Cambridge University Press.

Warriner, Keith, John Goyder, Heidi Gjertsen, Paula Hohner, and Kathleen McSpurren. (1996). "Charities, No; Lotteries, No; Cash, Yes," *Public Opinion Quarterly*, 60(4), 542–562.

Yammarino, Francis J., Steven J. Skinner, and Terry L. Childers. (1991). "Understanding Mail Survey Response Behavior," *Public Opinion Quarterly*, 55(4), 613–639.

Yu, Julie and Harris Cooper. (1983). "A Quantitative Review of Research Design Effects on Response Rates to Questionnaires," *Journal of Marketing Research*, 20(1), 36–44.