IMPROVING MAIL SURVEY RESPONSE RATES USING CHOCOLATE AND REPLACEMENT QUESTIONNAIRES

MIKE BRENNAN
JAN CHARBONNEAU

Abstract  This study examined the effectiveness of two procedures aimed at improving response rates for mail surveys: an incentive sent with either the first mail-out or first follow-up, and the inclusion of a replacement questionnaire, with or without an incentive, with either the first or second follow-up. The survey involved a sample of 1,600 New Zealand residents aged 18 years or older, randomly selected from the 2005 Electoral Roll and randomly assigned to one of four treatment groups. The response rates after two follow-ups ranged from 62.3 to 66.5 percent. Using a chocolate as an incentive with the first mail-out was effective in generating a significantly higher initial response than the control; however, after two follow-up mail-outs, differences were not significant. Sending a replacement questionnaire plus a chocolate with the first follow-up generated a significantly higher response rate than merely sending a letter, and the effect of this procedure persisted through to the end of the survey. Overall, these results provide further compelling evidence of the importance of using follow-up mail-outs for improving mail survey response rates.

Introduction

Mail surveys are widely used in survey research but have long been criticized for generating low response rates. While it has been convincingly demonstrated that response rates are a function of procedure rather than mode, and that high response rates are possible in mail surveys (Dillman 1972, 1978, 2000; Brennan 1992a), there is clear evidence that survey response rates in general are steadily declining, with mail surveys being no exception (Gendall 2000; de Leeuw and de Heer 2002; Bednall and Shaw 2003; CMOR 2003;
Curtin, Presser, and Singer 2005). Although there have been numerous investigations into procedures for improving mail survey response rates, there is clearly a need for on-going research in this area.

Over the years, the effects on response rates of virtually every aspect of the mail survey have been examined. For example, researchers have tested elements of the envelope (size, color, franked versus stamped, type of stamp, type of address, handwritten versus typed address, logo); the cover letter (tone, signature, paper type, honorific, content); questionnaire (colour, size, topic, layout); contact procedures (number of contacts, use of replacement questionnaire, type of contact, e.g., letter, postcard or telephone call; precontact and follow-up); and the use of incentives (for reviews, see Kanuk and Berenson 1975; Linsky 1975; Dillman 1978, 1991, 2000; Duncan 1979; Yu and Cooper 1983; Harvey 1987; Fox, Crask, and Kim 1988; Brennan 1992a; Church 1993). While numerous elements have been found to make small differences to response rates, two have consistently been found to contribute significantly to higher response rates: the use of incentives and the use of multiple attempts to contact nonrespondents (Kanuk and Berenson 1975; Linsky 1975; Yu and Cooper 1983; Harvey 1987; Dillman 2000).

Of the wide range of response inducements that have been tested, cash incentives have consistently been found to be the most effective (Armstrong 1975; Kanuk and Berenson 1975; Linsky 1975; Goodstadt et al. 1977; Heberlein and Baumgartner 1978; Hansen 1980; Furse, Stewart, and Rados 1981; Yu and Cooper 1983; Fox, Crask, and Kim 1988; Gajraj, Faria, and Dickinson 1990; James and Bolstein 1990; Church 1993; Jobber, Saunders, and Mitchell 2004). Indeed, it is possible for monetary incentives to be so effective at eliciting responses that the savings in postage from the reduction in follow-up mail-outs can substantially offset the cost of the incentive (Brennan, Hoek, and Astridge 1991; Brennan 1992b; Gendall, Hoek, and Brennan 1998). Unfortunately, in some countries, such as New Zealand, it is no longer legal to send monetary incentives via the mail, requiring the use of alternatives.

To be practical, mail survey incentives need to be inexpensive, easy to process and send through the mail, and attractive enough to encourage participation. On the basis of these criteria, many previously tested incentives are unsuitable. Some are quite expensive (books, lottery tickets, phone cards), various items are awkward to process (books, golf balls, pens), or others have more limited appeal (discount coupons, key rings, tie pins, trading stamps). Prize draws such as lottery tickets and “scratch-and-win” have the added drawback that nonrespondents may win. Among those examined, postage stamps, tea bags, coffee sachets, and chocolates appear to be suitable candidates; however, evidence to date suggests that tea bags, coffee sachets, and stamps are not very effective (Gendall, Hoek, and Brennan 1998; Gendall, Leong, and Healey 2005; Brennan, Charbonneau, and Hercus 2007), leaving chocolates as the option with the most potential.

To date, only one study has been found examining the effectiveness of chocolates as an incentive. Gendall, Leong, and Healey (2005) used gold-foil-covered
chocolate coins of various denominations which increased response rates by 2.7–5.1 percentage points. They also concluded that the coins were a cost-effective option. While it is not clear whether these results are due to the incentives being chocolates or because they are well-recognized novelty coins, this positive result suggests that further investigation of chocolates as an incentive is warranted.

Another effective way of improving mail survey response rates is to use several follow-up contacts (Scott 1961; Linsky 1975; Dillman 1991, 2000). It is not so clear, however, whether it is more effective to send out a replacement questionnaire than merely sending a follow-up letter. Von Reisen (1979) found replacement questionnaires to be more effective than follow-up postcards. Futrell and Lamb (1974) found replacement questionnaires to be more effective than follow-up letters; however, Heberlein and Baumgartner (1978) concluded that they were not. Brennan (1992a) found that sending a replacement questionnaire with the second reminder was no more effective in that wave than just sending a letter. Brennan (2004) found that sending a replacement questionnaire had no effect on the response rate to that wave when an incentive had been offered in the first mail-out, but had a significant effect when no incentive had been offered previously. After a second follow-up, however, the response rates for all treatment groups were similar. Given the conflicting results and the paucity of reported information concerning the most effective time to send a replacement questionnaire, this issue also requires further examination.

The purpose of this paper is to re-examine the effects on response rate, response speed, and the cost-effectiveness of two procedures: including a non-monetary incentive in either the first mail-out or first follow-up mail-out and sending a replacement questionnaire with either the first or second follow-up mail-out.

Method

SAMPLE

A sample of 1,600 New Zealand residents was randomly selected from the 2005 electoral roll of all eligible voters over the age of 18. Each selected respondent was randomly assigned to one of four treatment groups (henceforth referred to as T1, T2, T3, and T4) as set out in table 1. For the analysis, sample sizes were adjusted for undelivered mail (gone, no address) and ineligibles (deceased).

PROCEDURE

The survey was conducted between August 31 and October 21, 2005. The first mailing was sent on August 31, 2005. The second mailing was sent to nonrespondents 12 days after the first mailing on September 12, 2005, with
Table 1. Research Design

<table>
<thead>
<tr>
<th>Treatment groups</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>W1 Choc.</td>
<td>W2 Choc. + Q</td>
<td>W2 Choc. − Q</td>
</tr>
<tr>
<td>First mailing</td>
<td>L + Q</td>
<td>L + Choc. + Q</td>
<td>L + Q</td>
<td>L + Q</td>
</tr>
<tr>
<td>Second mailing</td>
<td>L</td>
<td>L</td>
<td>L + Choc. + Q</td>
<td>L + Choc.</td>
</tr>
<tr>
<td>Third mailing</td>
<td>L + Q</td>
<td>L + Q</td>
<td>L</td>
<td>L + Q</td>
</tr>
</tbody>
</table>

NOTE.—L = letter; Q = questionnaire; Choc. = chocolate incentive attached to the letter.

the third mailing sent 14 days after that on September 26, 2005. The overall response rate (RR1) after two follow-up mail-outs was 64.3 percent (AAPOR 2008).

SURVEY INSTRUMENT

The eight-page questionnaire on the topic of “Reality TV” was in the form of an A4 international standard size booklet (A3 international standard equivalent folded), in one of four bold colors (green, blue, red, and purple) (please see the supplementary data online). These colors were balanced across treatments and a post-survey analysis revealed no interactions between the questionnaire color and the treatments reported in this paper. A reply-paid envelope was provided whenever a questionnaire was supplied.

The one-page cover letter was printed on the university letterhead. It explained what the survey was about and why it was being conducted; assured respondents that the survey was simple and not trying to trick them or sell them anything; explained how they were selected; stressed confidentiality; explained the purpose of the ID number; informed them of the reply-paid envelope; acknowledged that they were busy and expressed appreciation in advance for their assistance. Shortened forms of this letter, emphasizing the importance of a response and appreciation of their assistance, were used in the follow-up mail-outs. All envelopes had the university logo printed in the top left-hand side corner.

The incentive was a small (45 mm × 55 mm × 6 mm), flat, individually foil-wrapped high-quality milk chocolate bar. This chocolate remains solid at temperatures considerably higher than those likely to be encountered by the mail so there was minimal likelihood of the chocolate melting. The colors of the packaging (blue and gold) coincidentally matched the colors of the university letterhead used for the cover letters, making for an appealing package. The incentive (when used) was attached to the letter with double-sided adhesive tape, and the following statement was added to the letter: “As a token of our appreciation, we hope you will enjoy the attached sample of
Table 2. Response Rates

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 1 + 2</th>
<th>Wave 3</th>
<th>Wave 1 + 2 + 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>1 Control</td>
<td>382</td>
<td>34.0</td>
<td>252</td>
<td>21.8</td>
<td>194</td>
</tr>
<tr>
<td>2 W1 Choc.</td>
<td>392</td>
<td>41.3*</td>
<td>228</td>
<td>15.8</td>
<td>186</td>
</tr>
<tr>
<td>3 W2 Choc. + Q</td>
<td>391</td>
<td>35.3</td>
<td>251</td>
<td>36.3</td>
<td>157</td>
</tr>
<tr>
<td>4 W2 Choc. − Q</td>
<td>389</td>
<td>31.9</td>
<td>264</td>
<td>22.7</td>
<td>201</td>
</tr>
<tr>
<td>Overall</td>
<td>1,554</td>
<td>35.6</td>
<td>995</td>
<td>24.3</td>
<td>738</td>
</tr>
</tbody>
</table>

Note.—Bold numbers represent the response rates for the Control and most effective treatment group for that wave.

*p < .05; **p < .01; #p < .001.

Whittaker’s chocolate.” Whittaker’s is a long-established, well-known New Zealand chocolate manufacturer.

Results

EFFECT OF THE INCENTIVE

The cumulative response rates across each mailing, henceforth referred to as waves, are reported in table 2. The following analysis will examine both the main effects due to the treatments compared to the Control and the cumulative effects across waves.

In Wave 1, the inclusion of the chocolate incentive for Group 2 generated a significant increase in the response rate over that of the Control of 7.3 percent (z = 2.09, p = .018). As the survey progressed, however, this effect disappeared. By the end of Wave 2 (see Wave 1 + 2 in table 2), the increase due to the incentive was down to 2.1 percentage points. By the end of the survey (see Wave 1 + 2 + 3 in table 2), the increase was only 0.7 percent. Neither of these differences was statistically significant.

In Wave 2, two treatments are of interest: T3 (questionnaire + chocolate) and T4 (letter only + chocolate). Comparing T1 (Control), T3, and T4, it is apparent that simply sending a chocolate with a reminder letter is not an effective way to improve response rates as the response rate for T4 was only slightly higher than that of the Control (22.7 percent versus 21.8 percent). Sending a chocolate with a replacement questionnaire, however, is a very effective way of increasing response rates to the first reminder. Compared to the Control (T1), the chocolate plus questionnaire treatment (T3) lifted response rates by almost 15 percentage points within the second wave and by over 10 percent by the end of Wave 2 (see Wave 1 + 2 in table 2) (Wave 2: z = 3.584, p = .0001; Wave 1 + 2: z = 2.836, p = .0023). An important finding however, is that by the end of the
survey, after two reminders, two of the three treatments involving incentives (T2 and T4) had only marginally higher response rates than the Control (.07 and 3.0 percentage points) and these differences were not statistically significant. However, the effect of sending the chocolate and replacement questionnaire with the first reminder (T3) persisted, maintaining a slight but statistically significant increase in the response rate after two reminders of 4.2 percentage points ($z = 1.217, p = .015$).

EFFECT OF PROVIDING A REPLACEMENT QUESTIONNAIRE

Given the ineffectiveness of the incentive in T4 (letter + chocolate), it would seem likely that most if not all of the increase in the response rate in T3 was due to the inclusion of a replacement questionnaire, not the chocolate (see table 2). Since the study did not test sending a questionnaire without an incentive in the first follow-up, however, this remains conjecture. These results suggest that, in terms of overall response rate, it does not matter whether a replacement questionnaire is sent with the first follow-up mail-out or the second.

SPEED OF RESPONSE

While the overall response rate is of interest, speed of response is also important as this affects the cost-effectiveness of the survey since it alters the cost of follow-up mail-outs. Figure 1 shows the cumulative daily responses across the entire survey (along the top) and for each wave (along the bottom). The regular flat spots on the curves coincide with weekends when no processing was done. Mail arriving one to two days after a mail-out was treated as belonging to

Figure 1. Cumulative Daily Responses: By Wave and Overall.
the preceding wave in order to account for the return mail posted but not yet received on the day of the follow-up mail-out. Mail arriving after this time was assumed to have been prompted by the follow-up mail-out.

It is apparent that the incentive used in Wave 1 (T2) was very effective at prompting a speedy response over the first five days, after which the rate dropped off and was similar for all treatment groups up until the first reminder. In contrast, sending a replacement questionnaire in Wave 2 (T3) produced a gradual but sustained increase in responses across the wave. After the second follow-up, however, responses to T3 (which received only a letter) dropped off quite quickly whereas the responses of groups T1, T2, and T4 (which all received a replacement questionnaire) were more sustained. These cumulative daily response curves demonstrate clearly the effectiveness of the incentive for producing a quick initial response and the value of sending out replacement questionnaires to sustain responses and generate a higher overall response rate.

COST-EFFECTIVENESS

The relative cost of each of the treatments is shown in table 3. For these calculations, the sample size for each of the three treatments was adjusted to that of the Control to allow for valid comparisons. The cost of materials, postage, processing the incentives, and preparing the mail-outs are factored into the cost calculations as outlined in Brennan, Seymour, and Gendall (1993).

The overall costs of each treatment were within 7–8 percent ($94–111) of the cost of the Control, but the average cost per response was even closer (0–6 percent). Indeed, T3, where a replacement questionnaire and chocolate were sent with the first reminder, actually cost just one cent more per response than the Control. The effectiveness of T3 (Q + choc. first reminder) is also evident when one considers the cost of the additional responses generated. T3 generated 17 more responses (8 percent) than the Control (T1) at almost the same cost as the average cost per return ($5.48 versus $5.47) and was just slightly less cost-effective ($E = .98; see table 3). In contrast, the extra responses generated by T2 (chocolate in the first mail-out) were quite expensive at $28 each and these extra three responses were probably not worth the extra cost and effort. As demonstrated by the $E$ values (see table 3), T4 ($E = .59$) was more cost-effective than T2 ($E = .17$) but not nearly as cost-effective as T3 ($E = .98$) or T1 ($E = 1$).

Discussion

This research demonstrates that sending a chocolate with the first mail-out is an effective way of generating a high initial response, although this advantage disappears if two follow-up mail-outs are used. Sending a chocolate with the first follow-up mail-out, however, does not improve the response rate unless
Table 3. Cost-effectiveness of the Incentive

<table>
<thead>
<tr>
<th>Treatment</th>
<th>A Cost ($)</th>
<th>B Response (n)</th>
<th>C Cost/response ($)</th>
<th>D Additional cost versus control ($)</th>
<th>E Additional responses versus control (n)</th>
<th>F Cost/extra response ($)</th>
<th>G Cost-effectiveness (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Control</td>
<td>1,305.72</td>
<td>239</td>
<td>5.47</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.00</td>
</tr>
<tr>
<td>2 W1 Choc.</td>
<td>1,404.40</td>
<td>242</td>
<td>5.80</td>
<td>98.68</td>
<td>3</td>
<td>28.01</td>
<td>.17</td>
</tr>
<tr>
<td>3 W2 Choc. + Q</td>
<td>1,400.11</td>
<td>256</td>
<td>5.48</td>
<td>94.39</td>
<td>17</td>
<td>5.58</td>
<td>.98</td>
</tr>
<tr>
<td>4 W2 Choc. − Q</td>
<td>1,416.92</td>
<td>251</td>
<td>5.65</td>
<td>111.22</td>
<td>12</td>
<td>9.10</td>
<td>.59</td>
</tr>
</tbody>
</table>

NOTES.—N = 382; standardized so all four treatment groups had same initial sample size.

E = Incremental response/incremental cost, where incremental response = ((response rate for treatment − response rate for control)/response rate for control) × 100,
and incremental cost = ((cost of treatment − cost of control)/cost of control) × 100. (See Brennan, Seymour, and Gendall 1993.)

E > 1 = more cost-effective than control.
a replacement questionnaire is sent as well. This result suggests that it was the replacement questionnaire rather than the incentive that improved the response rate although we cannot rule out the possibility of an interaction effect between the replacement questionnaire and incentive and this requires further investigation.

While a replacement questionnaire generated a higher response to both the first and second follow-up mail-outs than simply sending a follow-up letter, it is not possible to conclude that including a replacement questionnaire with either the first or second reminder is actually necessary when two follow-ups are used since the effect of sending only a letter in both follow-ups was not tested. This procedure also deserves further investigation as it has significant implications for the cost of a survey given that mailing a questionnaire is typically more expensive than mailing a letter (double the cost in this case) and incurs the additional cost of the questionnaire. In addition, replication is clearly needed to determine whether these results for the chocolate incentive and replacement questionnaire generalize across different survey topics and different survey populations.

A key conclusion from this study is that the most effective procedure for obtaining a high response rate is to send at least two follow-up mail-outs. While the differences in the response rate due to incentives and/or a replacement questionnaire were not large by the end of the survey, there was, in fact, a substantial increase in response rates across waves for all treatments. On average, the first reminder increased the response rate by 15.6 percentage points, while the second reminder increased the response rate by a further 13.1 percentage points. These results provide further evidence to support the claim that the key to achieving a respectable response rate is to increase the number of contacts, in other words, to send out at least two follow-up mail-outs (Dillman 1972, 1978, 2000; Kanuk and Berenson 1975; Linsky 1975; Heberlein and Baumgartner 1978; Brennan 2004), reinforcing the advice that has been provided by Dillman and others for many years.

Supplementary Data

Supplementary data is available online at http://poq.oxfordjournals.org/.

References


