Australasian Journal of Market & Social Research

December 2010 • Volume 18 Number 2 • ISSN 1832 7362

CONTENTS

Australian MARKET & SOCIAL RESEARCH SOCIETY AMSRS

AUSTRALASIAN JOURNAL OF MARKET & SOCIAL RESEARCH

AJMSR is published by the Australian Market & Social Research Society

ISSN 1832 7362

Australian Market & Social Research Society
Level 1
3 Queen Street
Glebe NSW 2037
Australia

Phone: +61 (0) 2 9566 3100
Fax: +61 (0) 2 9571 5944
Email: publications@amsrs.com.au
Web: www.amsrs.com.au

Australasian Journal of Market & Social Research

The Australasian Journal of Market & Social Research (AJMSR) is the official journal of the Australian Market & Social Research Society (AMSRS). It is published twice a year in June and December.

All members of the AMSRS receive a complimentary subscription to AJMSR as part of their membership.

Submissions

Original papers are invited. There is an editorial policy to ensure that a mix of theoretical and practical papers is presented. Papers should be of some relevance to at least a segment of the population of practicing market or social researchers. Purely academic marketing papers are probably more suitable elsewhere. All papers are refereed by independent assessors.

Authors are required to supply their articles electronically in Microsoft Word format, doubled spaced with article references in Harvard referencing style. Diagrams and figures also need to be provided electronically. In writing papers, authors should consider the style of articles in previous issues.

From this December issue onwards all editions of AJMSR will be referenced by EBSCO, an on-line database of journal articles accessible by researchers worldwide. As a writer we need your permission to enable this to occur. Could you please indicate your agreement to your article being referenced in this way when you submit it for publication.

Contributors of paper will receive an electronic copy of the issue in which their article appears.

AJMSR is an open publication. All expressions of opinion are published on the basis that they are not to be regarded as expressing the official opinion of the Australian Market & Social Research Society. The AMSRS accepts no responsibility for the accuracy of any opinions or information contained in this publication and readers should rely on their own enquiries in making decisions.

All papers for submission should be forwarded to:

Dr. David Bednall, Associate Professor
School of Management and Marketing
Faculty of Business and Law
Deakin University (Burwood)
221 Burwood Highway
Burwood VIC 3125 Australia
Email: dbednall@deakin.edu.au

Contact details

All enquiries regarding AJMSR matters should be directed to:

Australasian Journal of Market &
Social Research
Australian Market & Social Research Society
Level 1, 3 Queen Street
Glebe NSW 2037 Australia
Phone: +61 (0) 2 9566 3100
Fax: +61 (0) 2 9571 5944
Email: publications@amsrs.com.au

AMSRS recognises the contribution of Associate Professor David Bednall and Deakin University.

Sharing Your Ideas

While the AJMSR has a long history, a review of its articles shows that only a few members of the industry contribute regularly. There are a number of reasons apparent. First there is the mundane but important issue of having enough time to write the article. Agreed this is even more an issue now with shorter timelines, international parent companies and other demands. One possible way of alleviating this is to work with academics who are seeking to publish. As a practitioner you have ready access to an invaluable commodity that academics crave– data of all kinds, waiting for analysis.

A second reason is that the inventiveness of the industry in developing data collection and analysis techniques generates commercially valuable IP. To share this with competitors, especially those who have not paid for its development, is difficult. Using data belonging to clients can also be an issue. But there are many topics, such as the behaviour of survey panellists or managing large social research studies, which could easily be shared with other members of the industry. Similarly case studies (e.g. from the Effectiveness Awards), histories or meta-analyses are possible. Having an article published in a refereed journal like AJMSR means you get external validation of your work.

Third, some people are reluctant to submit their work to a refereeing process because they are uncertain about how well they would fare. There are two responses to this. First, you are welcome to seek advice from your editor about the suitability of what you are working on, prior to submission. The aim here is to foster useful and insightful articles about our industry and its practices. Second, peer review (typically by an industry practitioner and by an academic) can give great feedback that incites you to re-examine your ideas and analysis.

This edition of the journal has become the de facto "Mike Brennan" special issue. This is important as he continues to make a major contribution to our understanding of survey response and non-response. On the other hand, it would have been better to have more industry articles to accompany these contributions.

Hopefully over this relatively quiet period you are sitting down wondering how you can contribute!

Associate Professor David Bednall
dbednall@deakin.edu.au

The Effect of a Chocolate Incentive on Mail Survey Response Rates

Mike Brennan, Massey University

ABSTRACT

This study examined two methods for improving survey response rates: a chocolate incentive sent in either the first mail-out or first follow-up, and a replacement questionnaire sent with either the first or second follow-up. The survey involved a stratified random sample of 1800 people aged 18 years or older selected from the 2006 New Zealand electoral roll. After three mail-outs, the overall responses rates for the six treatment groups ranged from 37% to 46%. The chocolate incentive significantly increased the response rate to the first mail-out, but this advantage disappeared after two follow-up mail-outs. The incentive did not significantly increase the response rate when sent with either a reminder letter or a replacement questionnaire. Sending a replacement questionnaire with the first follow-up and a letter only in the second follow-up did not produce a significantly different response rate than sending a letter with the first follow-up and a replacement questionnaire with the second. After two follow-up mail-outs, the differences between the six treatment groups were not statistically significant. These results confirm the importance of follow-up mail-outs. They also confirm previous findings that a chocolate incentive can significantly boost response speed when sent with the first mail-out but is ineffective when sent with a follow-up, and unnecessary if two follow-ups are used. The results also suggest that it doesn't matter whether a replacement questionnaire is sent with the first follow-up or the second. The most cost-effective approach was to not use an incentive at all, send a letter only with the first follow-up, and send a replacement questionnaire with the second follow-up.

INTRODUCTION

Mail surveys are an important and long established research instrument, but the procedures used need ongoing development to deal with changing conditions. It has been well documented that response rates to surveys have been declining for many years, and these observations have sparked widespread concern over the possible effect this trend may have on non-response bias (Gendall 2000; de Leeuw and deHeer 2002; Bednall and Shaw 2003; CMOR 2003; Curtin, Presser, and Singer 2005). What is clear is that there is an ongoing need to address this issue. However, while numerous procedures have been tried in an attempt to boost response rates in mail surveys, only two appear to be consistently effective: the use of a pre-paid monetary incentive, and the use of multiple mail-outs (Kanuk & Berenson 1975; Linsky 1975; Duncan 1979; Harvey 1987; Dillman 1991; Singer, Groves & Corning 1999). Unfortunately, using a cash incentive is not always legal or appropriate. Furthermore, it is not clear what form a follow-up mail-out should take. Is it better to send a replacement questionnaire, or will a simple letter do? This paper examines the effect on response rate of using a pre-paid (ex ante) non-monetary incentive (a chocolate) in either the first mail-out or first follow-up, and either a letter or replacement questionnaire in subsequent follow-ups. The paper also examines the cost-effectiveness of these techniques.

While it is well established that sending a pre-paid cash incentive is an effective way to increase mail survey response rates (Armstrong 1975; Linsky 1975; Hansen 1980; Yu & Cooper 1983; Gajraj, Faria & Dickinson 1990; Brennan 1992; James & Bolstein 1992; Church 1993; Jobber, Saunders & Mitchell 2004), sending cash in the mail is forbidden in some countries, such as New Zealand, so an alternative form of incentive is required. Furthermore, cash may also be inappropriate in some circumstances, such as in surveys conducted for voluntary organisations, where a monetary incentive might be perceived by respondents as "wasting" valuable funds, or in surveys about financial matters or surveys targeting low income respondents, where a cash incentive may be viewed as "coercive".

While a large number of different types of non-monetary incentives have been tested, few have proven to be particularly suitable for mail surveys, or even effective. As noted by Brennan and Charbonneau (2009), some incentives that have been used are relatively expensive (books, lottery tickets, phone cards), others are awkward to process (books, golf balls, pens, pencils) and some have limited appeal (discount coupons, key rings, tie pins, trading stamps, bookmarks). Stamps, coffee sachets and tea bags meet the ideal criteria for a mail survey of being small and inexpensive, but do not appear to be effective (see Robinson & Agasim 1951; Brennan 1958; Pucel, Nelson & Wheeler, 1971; Whitmore 1976; Nederhof 1983; Snyder & Lapovsky 1984; Gajraj, Faria & Dickinson 1990; Jobber, Mirza & Wee 1991; Brennan, Seymour & Gendall 1993; Gendall, Hoek & Brennan 1998; Kalafatis & Madden 1995; Arzheimer & Klein 1999; Brennan, Charbonneau & Hercus 2007). However, recent studies suggest that chocolate might be a suitable option.

Effect of chocolate as an incentive

Three studies to date have specifically examined the effects of chocolate as an incentive. Gendall, Leong, and Healey (2005) tested the effectiveness of gold-foil covered chocolate "coins" as an incentive in a survey of the general public. These coins are typically used as novelty treats for children and come in three denominations and two sizes (small 20 cent; large 50 cent; large $2). They found that the coins produced modest increases in response rate (2.7% for the small coins and 5.1% for the large), and were cost-effective.

Brennan and Charbonneau (2009) tested a more up-market chocolate than Gendall et al. (2005) – a small high quality foil-wrapped, branded, milk chocolate square. This incentive significantly boosted the response rate to the first mail-out, although the overall response rate after two follow-up mail-outs was similar to that of the control. Sending a chocolate in a first follow-up to non-responders who had not previously received an incentive also significantly lifted the response rate, but only for respondents who also received a replacement questionnaire; the incentive had no effect when sent with a follow-up letter only.

Brennan and Xu (2009) used the same type of chocolate incentive employed by Brennan and Charbonneau (2009). They also found that the chocolate significantly increased the response rate for the first mail-out. However, a chocolate sent in the first follow-up to non-responders who had not previously received an incentive produced a significantly lower response rate than not sending an incentive at all. Thus, while the results for the first mail-out confirm the results of Brennan and Charbonneau (2009), the results for the first follow-up contradict their findings. Even so, the results of these studies suggest that a chocolate incentive can effectively increase the response rate of the first mail-out, and do so in a cost-effective manner.

Timing of incentives

Incentives are generally sent with the first mail-out. However, the studies of Brennan and Charbonneau (2009) and Brennan and Xu (2009) raise questions about whether an incentive can or should be used in a follow-up. Church (1993) concluded from a meta-analysis of 38 studies that there was strong support for including pre-paid incentives in the first mail-out, but does not comment on the effectiveness of including incentives in subsequent mail-outs. Furse, Stewart and Rados (1981) found no difference in overall response rates whether a cash incentive was sent in the first mail-out or follow-up mail-out, or both. These results suggest that sending the incentive with the first mail-out is an effective way of obtaining faster responses, but that sending the incentive with the follow-up mail-out may be more cost effective. However, the results of Brennan and Xu (2009) suggest that care needs to be taken with the timing of the incentive as there is a possibility that sending the incentive with the first follow-up rather than the first mail-out can have a negative effect and actually reduce responses, a point that has been made by others (e.g., Kalafatis & Madden 1995; Gendall, Leong, & Healey 2005). The present study examines this issue.

Follow-up mail-outs

While incentives may help boost responses, it has been well established that the most effective way to improve response rates is to use follow-up mail-outs (Dillman 1978, 1991, 2000; Singer, Groves & Corning 1999). However, it is less clear whether each mail-out should contain a replacement questionnaire or whether a letter is sufficient. Since using replacement questionnaires in follow-up mail-outs has cost implications, a relevant question is whether it is necessary to send a replacement questionnaire rather than simply a reminder letter, and if so, whether it is more effective to send the replacement questionnaire with the first or the second follow-up mail-out. In addition, there are unanswered questions about the effect of the timing of the incentive and the combining of an incentive with a particular type of follow-up that needs consideration. The present study examines this issue also.

The evidence regarding the effectiveness of different types of follow-up is scant and equivocal. Both Futrell and Lamb (1974) and Brennan, Charbonneau & Hercus (2007) found a replacement questionnaire to be more effective than a reminder letter, but Heberlein and Baumgartner (1978) did not. Brennan (2004) obtained a significantly higher response after two mail-outs for a replacement questionnaire than for a reminder letter when no incentive was offered, but no found difference between a replacement questionnaire and a reminder letter when an incentive was offered (a prize draw offered in the first and second mail-out for $100 worth petrol vouchers). Brennan and Charbonneau (2009) found a replacement questionnaire with a chocolate was more effective than a reminder letter either with or without a chocolate, but did not examine the effect of a replacement questionnaire without a chocolate. In the studies by Brennan et al., differences in the effects of the treatments had disappeared after a second follow-up. The observations suggest an interaction between the use of an incentive and the type of follow-up (replacement questionnaire or letter). They also suggest that the number of follow-ups is what has the most important influence on response rates. This issue requires further attention, and is also addressed in this paper.

In summary, using chocolate as a prepaid incentive in the first mail-out appears to offer an effective way of increasing the response rate, although this advantage does not appear to persist if more than one follow-up is used. However, evidence to date is mixed regarding the effectiveness of chocolate as an incentive when used with the first reminder instead of the first mail-out. Furthermore, the apparent diminishing effects of incentives across mail-outs raises the question of whether an incentive is even necessary if multiple mail-outs are used. And from a practical perspective, a researcher needs to know which of the various combinations of techniques is the most cost-effective. The purpose of this paper is to examine these issues. Specifically:

1. Is a chocolate incentive an effective way of increasing the response rate in a mail survey?
2. Should a replacement questionnaire be sent with the first or second follow-up?
3. What is the most cost-effective combination of these techniques (incentive/no incentive; replacement questionnaire/letter) for boosting mail survey response rates?

Table 1: Research Design

	T1 QLQ	T2 QQL	T3 Q(L+C)Q	T4 Q(Q+C)L	T5 (Q+C)LQ	T6 (Q+C)QL
First mailing	Q	Q	Q	Q	Q+C	Q+C
Second mailing	L	Q	L+C	Q+C	L	Q
Third mailing	Q	L	Q	L	Q	L

Note: Q = Questionnaire L = Letter C = Chocolate incentive attached to letter
QLQ means Questionnaire in mail-out 1, Letter only in mail-out 2 and Questionnaire in mail-out 3.
(Q+C) means Questionnaire + Chocolate ; (L+C) means Letter + Chocolate
T1 –T6 were balanced with respect to Questionnaire version (6 versions), respondent age group (5 categories) and respondent gender (2)

METHOD

A mail survey of 1800 members of the general public was conducted between June 25 and August 31, 2007. Six quota samples of 300 people each, balanced in terms of gender and age-group (5 categories), were randomly selected from the 2006 Electoral Roll of a major South Island city in New Zealand. These six samples were used for experimental treatments presented in the questionnaires, so each was allocated a slightly different version of the questionnaire. Members of each of these six samples were then systematically allocated to one of six mail survey treatment groups, such that each of the groups was balanced in terms of gender, age-group and questionnaire version. The topic of the survey was "New Zealanders attitudes towards new products" and contained questions about five "new" products or services. The questionnaire was 12 A4 pages long. After removing Gone/No Address's and ineligibles, the final sample size was 1646, with 691 valid responses.

The incentive used in some mail treatments was a small, flat (44mm x 45mm x 6mm) foil-wrapped milk chocolate (Whittaker's), which was attached to the cover letter with double sided adhesive tape. The chocolate is hard, and does not easily break or melt in the mail.

The cover letter explained what the survey was about, requested participation, and where the incentive was used, ended with the statement: "Please accept the attached chocolate as a small token of appreciation." The research design is shown in Table 1.

RESULTS

The response rates for each wave, and the cumulative response rates across waves, are displayed for each treatment in Table 2. Waves 1, 2 and 3 represent the responses to the first, second and third mail-outs respectively; Wave 1+2 and Wave 1+2+3 represent the cumulative response rates following Wave 2 and Wave 3 respectively.

Effect of incentive on response rates

In Wave 1, treatments T1 to T4 are identical (questionnaire with no incentive) and serve as controls, whereas treatments T5 and T6 both involve an incentive (questionnaire with chocolate). A comparison of the average response rate for the controls (T1,T2,T3,T4) with the average response rate for the incentives (T5, T6) indicates that the incentive produced a (statistically) significantly higher response rate than the control (control: 24.5%, incentive: 29.7%, z = 2.281, p = .023).

In Wave 2, sending a chocolate incentive with a reminder letter was no more effective than sending a reminder letter without an incentive (T3 : T1, 11.6% cf. 12.0%, z =.116, p = 0.908), sending a replacement questionnaire with an incentive was no more effective than sending a replacement questionnaire without an incentive (T4 : T2, 15.2% cf. 13.90%, z =.366, p = 0.714). By the end of Wave 2 (Wave 1+2), the differences between the six treatments were not statistically significant ($\chi 2$ = 6.604, d.f. = 5, p = .252). The effect of using an incentive in Wave 3 was not examined.

After two follow-up mail-outs (Wave 1+2+3), the differences between the six treatments were not statistically significant ($\chi 2$ = 5.273, d.f. = 5, p = .383). This result suggests that the incentive worked by prompting quicker responses from people likely to respond to the survey anyway, rather than by attracting respondents who wouldn't otherwise respond, since the incentive had no significant effect on the overall response rate after two follow-up mail-outs had been used.

Table 2: Effect of Incentive and Replacement Questionnaire on Response Rates

	Wave 1 N %	Wave 2 N %	Wave 1+2 N %	Wave 3 N %	Wave 2+3 N %	Wave 1+2+3 N %
T1: QLQ	269 22.3	209 12.0	269 31.6	184 13.6	209 23.9	269 40.9
T2: QQL	270 23.0	208 13.9	270 33.7	179 4.5	208 17.8	270 36.7
T3: Q(L+C)Q	274 24.5	207 11.6	274 33.2	183 14.2	207 24.2	274 42.7
T4: Q(Q+C)L	274 28.1	197 15.2	274 39.1	167 5.4	197 19.8	274 42.3
T1+T2+T3+T4	1087 24.5					
T5: (Q+C)LQ	281 31.3	193 10.9	281 38.8	172 11.6	193 21.2	281 45.9
T6: (Q+C)QL	278 28.1	200 14.5	278 38.5	171 7.6	200 21.0	278 43.2
T5+T6	559 29.7					
Mean	1646 26.2	1214 13.0	1646 35.9	1056 9.6	1214 21.3	1646 42.0

Note: The table reports the response rates to a particular mail-out or set of mail-outs.
QLQ means Questionnaire in mail-out 1, Letter only in mail-out 2 and Questionnaire in mail-out 3.
(Q+C) means Questionnaire + Chocolate ; (L+C) means Letter + Chocolate

Effect of replacement questionnaire on response rate

The effects of sending a replacement questionnaire are examined in three sets of paired treatments: T1 with T2; T3 with T4; and T5 with T6. The results are presented in Table 2.

While sending a replacement questionnaire with the first reminder letter produced a higher response rate for that wave (Wave 2) than sending a reminder letter without a replacement questionnaire, the difference was not statistically significant for any of the three pairs of treatments (T2:T1, 13.0% cf. 12.0%, z = .602, p = 0.547; T4:T3, 15.2% cf. 11.6% z = 1.072, p = 0.284; T6:T5, 14.5% cf. 10.9%, z = 1.075, p = 0.282), thus the differences may be due to chance.

Caution is required when examining, in isolation, the effects of sending a replacement questionnaire with the second follow-up (Wave 3), as the treatments in Wave 2 differed for each member of the three pairs of comparative treatments (T1:T2; T3:T4; T5:T6). For example, in Wave 2, T1 received a letter whereas T2 received a questionnaire, while in Wave 3, T1 received a questionnaire whereas T2 received a letter. Thus to determine whether it is more effective to send a replacement questionnaire with the first reminder or the second, it is necessary to combine the results of Wave 2 and Wave 3 (Wave 2+3 in Table 2). The differences are statistically non-significant across all six treatments (($\chi 2$ = 3.666, d.f. = 5, p = .589), and also for the matching pairs of treatments (T1:T2, 23.9% cf. 17.8%, z = 1.540, p = .124; T3:T4, 24.2 % cf. 19.8%, z = 1.055, p = .291; T5:T6, 21.2% cf. 21.0%, z = .059, p = .953).

The differences between the cumulative responses after three mail-outs are also statistically non-significant (T1:T2, 40.9% cf. 36.7%, z = 1.006, p = .315; T3:T4, 42.7% cf. 42.3%, z = .086, p = .931; T5:T6, 45.9% cf. 43.2%, z = .652, p = .515). This series of results suggests that if a replacement questionnaire is sent in a follow-up mail-out, the effect on response rate is similar regardless of whether it is sent with the second mail-out or the third.

Combined effects of incentive and replacement questionnaire

The issue of which combination of incentive and replacement questionnaire is most effective is addressed by comparing the final responses rates of each treatment with that for T1, which represents the standard approach we generally use (T1: QLQ - see Table 2). None of the differences between treatments and control were statistically significant (T1:T2, 40.9% cf. 36.7%, z = 1.006, p = .315; T1:T3, 40.9% cf. 42.7%, z = .427, p = .670; T1:T4, 40.9% cf. 42.3%, z = .341, p = .733; T1:T5, 40.9% cf. 45.9%, z = 1.185, p = .236; T1:T6, 40.9% cf. 43.2%, z = .538, p = .591).

These results justify the use of the standard procedure (T1: QLQ), which is to send a reminder letter with the second mail-out, and a replacement questionnaire with the third, as in T1. However, given the findings of Brennan and Charbonneau (2009), and the result for T5, one might also send a chocolate incentive with the first mail-out because of the faster initial response generated by the incentive, and the potential savings this may generate. Of course, these savings will depend on the response rates to Waves 1 and 2, as the higher these are, the fewer follow-up mail-outs required and the greater the savings on postage and printing.

Cost-effectiveness

A consideration when selecting a procedure for eliciting a high response rate is not merely the response rate achieved, but the cost effectiveness of the procedure. One way to evaluate cost-effectiveness is to compare the cost per return. Another, more informative, approach is to compare incremental response the procedure generates (i.e., the difference in response from the control) with the incremental cost of the procedure (i.e., the difference in cost relative to the control) (Brennan, Seymour & Gendall 1993).

These measures are presented in Table 3. Due to the variation in responses to identical treatments in Wave 1 for T1, T2, T3 and T4, the Wave 1 response rate for each group was replaced by the mean response rate of the four treatment groups for the calculation of costs and returns in each wave. Similarly, the Wave 1 response rate for T5 and T6 was replaced by the mean response rate of these two treatment groups. Costs (postage, printing, stationery, incentives), were based on the number of mail-outs required for each wave (that is, the number of non-responders to the previous mail-out) plus the number of responses to each mail-out (return mail costs). Postal charges for reply-paid responses are based on actual returns, not the number of mail-outs. Letters cost 45 cents each, while questionnaire mail-outs and survey returns (A4 envelopes) cost 90cents each. The incentives were 20 cents each.

The results in Table 3 indicate that none of the treatments were as cost-effective as the control (T1). While sending an incentive in the first mail-out (T5 and T6) generated a slightly higher response rate, the average cost per return was also higher. The best of these procedures was T5, which involved sending a chocolate in the first mail-out, a letter only in the second, and a replacement questionnaire in the third. However, compared with the control, T1, this procedure cost 3.7% more per response ($7.89 cf. $7.63), and cost 8.7% more to generate a 4.7% higher response. Thus for this survey, the most cost-effective approach was the control, which used the same set of procedures as T5, but without the incentive (i.e., first mail-out = Q + no incentive; first reminder = letter + no incentive; second reminder = questionnaire + no incentive).

Table 3: Relative Cost-effectiveness of Treatments

	n	Cost ($)	Return	cost/return	IR%	IC%
T1: QLQ (Control)	100	327.83	43	7.63	-	-
T2: QQL	100	334.50	38	8.82	-11.6	2.0
T3: Q(L+C)Q	100	357.08	43	8.23	0.0	8.9
T4: Q(Q+C)L	100	362.96	39	9.20	-9.3	10.7
T5: (Q+C)LQ	100	356.41	45	7.89	4.7	8.7
T6: (Q+C)QL	100	361.46	45	8.11	4.7	10.3

Note. In waves 1, 2 and 3, the average wave 1 response rate value for T1...T4 was used for T1,T2,T3 and T4 and the average Wave 1 response rate value for T5 and T6 was used for T5 and T6.
IC% = percentage incremental cost of the treatment compared to that of the control (T1).
IR% = percentage incremental response rate for the treatment compared to that of the control (T1).

DISCUSSION

The results confirm earlier findings that a chocolate incentive is effective when used in the first mail-out, but not when used in the first follow-up mail-out, regardless of whether it accompanies a reminder letter or a replacement questionnaire (Brennan& Charbonneau 2009; Brennan & Xu 2009). The results are also consistent with previous observations that the effects of the incentive tend to be restricted to the wave in which it is used. That is, after three mail-outs, the response rates for the treatments that received an incentive were not significantly higher than those that did not, even when the incentive generated a significantly higher return to the mail-out in which it was used. This diminishing effect across mail-outs is consistent with observations reported by others for both monetary incentives (Hopkins, Hopkins & Schon 1988; Brennan, Hoek & Astridge 1991; James & Bolstein 1992) as well as non-monetary incentives (Nederhof 1983). The implication is that an incentive is not needed if at least three mail-outs are going to be used, but an incentive sent with the first mail-out will increase responses to that mail-out, thereby reducing the costs of follow-up mail-outs. Thus, if the early response rate is high enough, the use of the incentive can become cost-effective.

In the present study, the use of the incentive was less cost-effective than the control, a finding also reported by Brennan and

Charbonneau (2009). Thus while chocolate can boost response rates to mail-out in which it is used, evidence to date suggests that it does not come close to the effectiveness of a pre-paid monetary incentive. For example, in the study of Brennan, Hoek and Astridge (1991) a small monetary incentive (50 cents) boosted response rates to the first wave to the extent that the treatment cost less than the control because of the reduced cost of follow-up mail-outs. In their study, the 50 cent incentive not only produced the highest response rate, but was also the most cost-effective procedure.

The finding that sending a replacement questionnaire with the first follow-up failed to generate a significantly higher response than simply sending a reminder letter is consistent with the conclusions of Herberlein and Baumgartner (1978) and the findings reported in Brennan (1992), although contrary to the findings of Futrell and Lamb (1974), Brennan and Charbonneau (2009), and Brennan (2004). Sending a replacement questionnaire in the second follow-up also had only a small effect. However, given that the proportion of unanswered questionnaires that may be "lost" or disposed of by potential respondents will increase as time passes, it does seem sensible to send a replacement questionnaire. Evidence to date suggests that it is better to send it with the second follow-up rather than the first.

The results in the present study also clearly demonstrate the importance of multiple mail-outs. The response rate for the overall control, T1, which did not use an incentive, increased nine percentage points after the second mail-out, and increased a further nine percentage points after the third. The response rates were: 22.3% (Wave1), 31.6% (Wave 1+2); 40.9% (Wave 1+2+3) – see Table 2. The average across all six treatments was a 9.7 percentage point increase from Wave 1 to Wave 1+2, and a further 6.1 percentage points increase from Wave 1+2 to Wave 1+2+3. These increases due to an additional mail-out far exceed those generated by either the incentive (5 percentage points in Wave 1; 2 percentage points in Wave 2) or the use of a replacement questionnaire (2 percentage points in Wave 2). Thus these results provide further support for the advice of Dillman (1978, 1991, 2000), and others, to use multiple mail-outs.

The strength of the present study is that the sample was stratified by age category, gender, and questionnaire version, and obtained reasonable sample sizes in each treatment, thus reducing sampling error. However, the study is limited by the fact that the survey was essentially a vehicle for a series of experiments, thus the topic and tasks involved were almost certainly more demanding and less intrinsically rewarding to respondents than typical surveys of the general public. It is therefore not surprising that the response rate was low; 42% overall compared with 56% for Brennan and Charbonneau (2009), for samples drawn from a similar population. However, given that demanding surveys are precisely when procedures for lifting response rates are needed in order to minimise non-response bias, the failure of the procedures involving the incentive and/or replacement questionnaire to make a significant difference is disappointing. Indeed, these results appear to argue against claims that incentives work best with respondents for whom the topic has less salience (e.g., Berlin et al. 1992, Baumgartner & Rathbun 1997, Martinez-Ebers 1997, Groves, Singer and Corning 1999). As with all methodological studies, what are required are further replications across different survey populations, and involving different topics, to help define the boundary conditions for cost-effective use of this incentive.

In conclusion, the results of the present study clearly demonstrate why multiple follow-up mail-outs are critical in mail surveys, and confirm the effectiveness of sending a chocolate incentive in the first mail-out as a means of boosting responses to that mail-out. On the basis of the present study, the message for survey researchers is to use at least two follow-up mail-outs, and if this not feasible, consider using a chocolate incentive with the initial posting.

References

Armstrong, J.S. 1975. Monetary incentives in mail surveys. *Public Opinion Quarterly*, 39 (spring), 111-116.

Arzheimer, K. & Klein, M. 1999. The effect of material incentives on return rate, panel attrition and sample composition of a mail panel survey. *International Journal of Public Opinion Research*, 11, 368-377.

Baumgartner, R. & Rathbun, P. 1997. Prepaid monetary incentives and mail survey response rates. *Paper presented at the Annual Conference of the American Association of Public Opinion Research*, Norfolk, Virginia.

Bednall, David H.B., and Mike Shaw. 2003. "Changing Response Rates in Australian Market Research." *Australian Journal of Market Research* 11(3): 31-41.

Berlin, M., Mohadjer, L., Waksberg, J., Kolstad, A., Kirsch, A., Kirsch, I., Rock, D. & Ymamoto, K. 1992. An experiment in monetary incentives. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 393-398.

Brennan, R.D. 1958. Trading stamps as an incentive in mail surveys. *Journal of Marketing*, 22 (3), 306-307.

Brennan, M. 1992. The effect of a monetary incentive on mail survey response rates: New data. *Journal of the Market Research Society*, 34 (2), 173-177.

Brennan, M. 2004. A test of two procedures for increasing responses to mail surveys. *Marketing Bulletin*, 15 (Research Note 3), 1-9. http://marketing-bulletin.massey.ac.nz

Brennan, M. Charbonneau, J. & Hercus, A. 2007. Coffee, tea or replacement questionnaire: Improving mail survey response rates from volunteers. *Australasian Journal of Market and Social Research*, 15 (1), June , 11-17.

Brennan, M. & Charbonneau, J. 2009. Improving mail survey response rates using chocolate and replacement questionnaires. *Public Opinion Quarterly*, 73 (2), 368-378.

Brennan, M., Hoek, J. & Astridge, C. 1991. The effects of monetary incentives on the response rate and cost-effectiveness of a mail survey. *Journal of the Market Research Society*, 33 (3), 229-241.

Brennan, M., Seymour, P. & Gendall. P. 1993. The effectiveness of monetary incentives in mail surveys: Further data. *Marketing Bulletin*, 4, 43-51. http://marketing-bulletin.massey.ac.nz

Brennan, M. & Xu X. 2009. The effect of questionnaire colour, a chocolate incentive and a replacement return envelope on mail survey response rates. *Australasian Journal of Social and Market Research*, 17 (1), June, 29-35.

Church, A.H. 1993. Estimating the effect of incentives on mail survey response rates: A meta-Analysis. *Public Opinion Quarterly*, 57, 62-79.

Curtin, R, Presser, S and Singer, E. 2005. Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 69, 87-98.

CMOR. 2003. Respondent Cooperation and Industry Image Study. The Council for Marketing and Opinion Research, Port Jefferson, New York. http://www.cmor.org/resp_coop_studies.cfm

de Leeuw, Edith, and Wim de Heer. 2002, Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In *Survey Nonresponse. ed*. Robert M Groves, Don A Dillman, John E Eltinge, and Roderick J. A. Little, 41-54. New York: Wiley.

Dillman, D.A. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons.

Dillman, D.A. 1991. The design and administration of mail surveys. *Annual Review of Sociology*, 17, 225-49.

Dillman, D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Duncan, W. J. 1979. Mail questionnaires in survey research: A review of response inducement techniques. *Journal of Management*, 5 (Winter), 39-55.

Furse, D.D., Stewart, D.W. & Rados, D.L. 1981. Effects of foot-in-the-door, cash incentives and follow-ups on survey response. *Journal of Marketing Research*, 18, 473-478.

Futrell, C.M. & Lamb, C.W. 1974. Effect on mail survey return rates of including questionnaires with follow-up letters. *Perceptual and Motor Skills*, 52, 11-15.

Gajraj, A.M., Faria, A.J. & Dickinson, J.R. 1990. A comparison of the effect of promised and provided lotteries, monetary and gift incentives on mail survey response rates, speed and cost. *Journal of the Market Research Society*, 32 (1), 141-162.

Gendall, P. 2000. Responding to the problem of non-response. *Australasian Journal of Marketing Research*, 8 (1), 3-18.

Gendall, P, Leong, M. & Healey, B. 2005. *The effect of pre-paid non-monetary incentives in mail surveys*. In proceedings of ANZMAC 2005 Conference, December 5-7, Freemantle, Australia.

Gendall, P., Hoek, J. & Brennan, M. 1998. The teabag experiment: More evidence on incentives in mail surveys. *Journal of the Market Research Society*, 40 (4), 347-352.

Groves, R.M., Singer, E. & Corning, A.D. 1999. Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, 64, 299-268.

Hansen, R.A. 1980. A self-perception interpretation of the effect of monetary and nonmonetary incentives on mail survey respondent behaviour. *Journal of Marketing Research*, XVII, pp77-83.

Harvey, L. 1987. Factors affecting response rates to mailed questionnaires: A comprehensive literature review. *Journal of the Market Research Society*, 29 (3), 341-53.

Heberlein, T.A. & Baumgartner, R. 1978. Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review* 43, 447-62.

Hopkins, K.D., Hopkins, B.R. & Schon, I. 1988. Mail surveys of professional populations: The effects of monetary gratuities on return rates. *Journal of Experimental Education*, 56, 173-175.

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, 346-361.

James, J.M. & Bolstein, R. 1992. Large monetary incentives and their effect on mail survey response rates. *Public Opinion Quarterly*, 56, 442-453.

Jobber, J.M., Mirza, H. & Wee, K.H. 1991. Incentives and response rates to cross-national business surveys: a logit model analysis. *Journal of International Business Studies*, fourth quarter, 711-721.

Jobber, D., Saunders, J. & Mitchell. V-W. 2004. Pre-paid monetary incentive effects on mail survey response. *Journal of Business Research*, 57, 347-50.

Kalafatis, S.P. & Madden, F.J. 1995. The effect of discount coupons and gifts on mail survey response rates among high involvement respondents. *Journal of the Market Research Society*, 37 (2), 171-184.

Kanuk, L. & Berenson, C. 1975. Mail surveys and response rates: A review. *Journal of Marketing Research*, XII (November), 440-453.

Linsky, A.S. 1975. Stimulating responses to mailed questionnaires: A review. *Public Opinion Quarterly*, 39 (spring), 82-101.

Martinez-Ebers, V. 1997. Using monetary incentives with hard-to-reach populations in panel surveys. International Journal of *Public Opinion Research*, 9, 77-86.

Nederhof, A. J. 1983. The effects of material incentives in mail sturveys: Two studies. *Public Opinion Quarterly*, 47, 103-111.

Pucel, D.J., Nelson, H.F. & Wheeler, D.N. 1971. Questionnaire follow-up returns as a function of incentives and responder characteristics. *Vocational Guidance Quarterly*, 19, 188-193

Robinson, R.A. & Agasim, P. 1951. Making mail surveys more reliable. *Journal of Marketing*, 15, 415-424.

Singer, E., Groves, R.M. & Corning, A.D. 1999. Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation. *Public Opinion Quarterly*, 63 (2), 251-260.

Snyder, M. & Lapovsky, D. 1984. Enhancing survey response from initial non-consenters. *Journal of Advertising Research*, 24 (3), 17-20.

Whitmore, W.J. 1976. Mail survey premiums and response bias. *Journal of Marketing Research*, XII (February), 46-50.

Yu, J. & Cooper, H. 1983. A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research*, XX (1), February, 36-44.

Address correspondence to :

Mike Brennan
Department of Communication,
Journalism & Marketing
Massey University
Private Bag 11222,
Palmerston North, New Zealand.
Ph +66 6 356 5576
Email: m.brennan@massey.ac.nz
(corresponding author)

The Effect of an Incentive (Chocolate) and a Replacement Questionnaire on Sample Composition, Item Non-response and Response Distribution in a Mail Survey

Mike Brennan and Jan Charbonneau, Massey University

Abstract

This paper examines the effect on sample composition, item non-response and response distribution of using a chocolate incentive and a replacement questionnaire to boost response rates in a mail survey. The data are from a previous study in which a chocolate incentive generated a significantly higher response to the first mail-out (41.3% cf. 34.0%), while a replacement questionnaire accompanied by a chocolate incentive generated a significantly higher response to the second mail-out 36.3% cf. 21.8%. While these results are encouraging, the point of using an incentive is to boost response rates in the belief that this will reduce both non-response bias and measurement error. It does not make sense to commit resources to improving response rates if this has a negative (or no) effect on total survey error. The analysis presented in this paper found the differences between survey method treatment groups with regard to sample composition, item non-response rates and response distributions to be small and not statistically significant. These findings suggest that using a chocolate incentive and a replacement questionnaire are unlikely to have any significant unintended effects on total survey error in mail surveys of the general public.

INTRODUCTION

Mail surveys are widely used in survey research, but declining response rates and increasing item non-response rates are both major issues for survey researchers (Beatty and Hermann, 2002; Bednall and Shaw, 2003; CMOR, 2003; Curtin, Presser and Singer, 2005; de Leeuw, 2001; de Leeuw and de Heer, 2002; Yan, Jans and Curtin, 2006) because of concern that these factors may make a significant contribution to total survey error. As a consequence, researchers are constantly searching for methods to boost survey participation and completion rates.

One of the main procedures used to lift response rates is to use an incentive, either pre-paid (ex ante) or promised (ex post), and both types are widely used in survey research. Pre-paid monetary incentives in particular have been found to be very effective for improving response rates (Brennan, 1992). However, while finding ways to obtain respectable response rates is a priority, it is also important to examine the effects incentives may have on sample composition, response bias and measurement error (Groves and Couper, 1998; Groves, 2006a; Groves et al., 2006b; Olson, 2006a). As noted by Olson and Kennedy (2006b): "Bringing in respondents who were difficult to contact or who had previously refused to participate in the survey may decrease nonresponse bias and increase the representativeness of the sample compared with the population. However, if these respondents are also those who are the most likely to give faulty answers, and the increase in measurement error exceeds the reduction in nonresponse bias, then the mean square error of those estimates will have increased, relative to not having included those respondents in the sample" (p.4181).

In spite of concern that incentives may affect sample composition and lead to biased results (e.g., Furse and Stewart, 1982; Hansen, 1980; Nederhof, 1983; Groves et al., 2006; Olson, 2006a; Robertson and Bellenger, 1978; Ryu, Couper and Marans, 2006; Willimack et al., 1995; Wotruba, 1966), the evidence that incentives create sample bias is not strong. Nederhof (1983) found no bias in age, gender, marital status, religion, political creed and research experience associated with the use of an incentive, but did find significant differences for education and occupation. Furthermore, while evidence on these matters to date is scant, there is some evidence that response distributions are not necessarily affected, even if demographic variables are (e.g., Finn, Wang and Lamb, 1983; Hansen, 1980; Mizes, Fleece and Roos, 1984; Nederhof, 1983; Ryu, Couper and Marans, 2006; Whitmore, 1976). Goritz (2004) concluded that promised incentives (redeemable bonus points, cash lottery and prize lottery) did not affect sample composition, response quality or survey outcome when used with panel members.

There is also some evidence that using an incentive may even help reduce non-response bias associated with the effects of respondent interest (or disinterest) in the survey topic (Groves et al., 2006), and there is some evidence that incentives may help reduce item non-response (McDaniel and Rao 1980). Both Wotruba (1966), and James and Bolstein (1990), reported a higher level of completed questionnaires, as well as a higher response rate, when a pre-paid (monetary) incentive was used, while McDaniel and Rao (1980) found a monetary incentive significantly decreased item omission and response error, and improved completeness of answers.

While encouraging, most of the research to date on the effects of incentives on response and non-response bias has examined the effects of monetary incentives. However, in some countries, such as New Zealand, it is not legal to send cash via the mail, so non-cash incentives have to be used. While most non-cash pre-paid (ex ante) incentives are not particularly effective, one that does show promise is chocolate (Gendall, Leong, and Healey, 2005; Brennan & Xu, 2009; Brennan & Charbonneau 2009; Brennan, 2008). Gendall et al. (2005) reported an overall increase in response rates of 2.7-5.1 percentage points for novelty chocolate coins. Brennan (2008), Brennan & Xu (2009) and Brennan & Charbonneau (2009) found that an up-market foil-wrapped chocolate square produced higher response rates than the control (no incentive) for the first mail-out (5.2, 8.5 and 7.3 percentage points respectively), although in two of these studies, the early boost to response rates did not persist when two follow-up mail-outs were used (Brennan, 2008; Brennan & Charbonneau, 2009).

Brennan and Charbonneau (2009) also found that sending a replacement questionnaire rather than just a letter in a follow-up mail-out significantly boosted the response rate to that mail-out (by 14.5 percentage points). The two key findings in that study were: i. the incentive sent in the first mail-out boosted the response rate relative to the control (41.3% cf. 34.0%, $z = 2.09$, $p=.018$), but the effect did not persist after two follow-up mail-outs; ii. Sending a replacement questionnaire with a chocolate in the second mail-out boosted the response rate to that mail-out compared with a letter-only control (36.3% cf. 21.8%, $z = 3.58$, $p=.000$), but the effect on the overall response rate after a further mail-out was not significant. It would appear that the effect of the replacement questionnaire plus chocolate was due to the replacement questionnaire rather than the chocolate, as a control using just a chocolate with a reminder letter failed to boost the response rate (achieved just 22.7%).

While the studies of Brennan (2008), Brennan and Xu (2009) and Brennan and Charbonneau (2009) demonstrate that chocolate and a replacement questionnaire can effectively boost early response rates, and can be cost-effective, there remains the question of whether these treatments have any effect, positive or negative, on survey error. For example, do they influence the types of respondents recruited, affect completion rates, or affect data quality? The purpose of the present study is to examine the effects of a chocolate incentive on sample composition, item non-response and response distributions, using data from a previous study (Brennan & Charbonneau, 2009).

METHOD

Sample

An eight page questionnaire on the topic of Reality TV and product placement was sent in a mail survey to 1600 New Zealand residents randomly selected from the 2005 electoral roll. The overall response rate after two follow-up mail-outs was 64.4% (valid responses / [1600 – returned gone no address] = 1001 / [1600 – 46]). The overall response rates for the four treatment groups (see Table 1) were: T1, 62.3%; T2, 63%; T3, 65.5%; T4, 65.3% (see Brennan & Charbonneau 2009 for more details).

Procedure

Respondents were randomly assigned to one of the four treatment groups set out in Table 1. Group T1 was the overall Control. Group T2 received a non-monetary incentive (chocolate) in the first mailing. Non-respondents in Group T3 received a chocolate incentive as well as a replacement questionnaire with the first reminder. Non-respondents in Group T4 received a chocolate incentive but no replacement questionnaire with the first reminder.

The sample characteristics for each treatment group are reported in Tables 2, 3 and 4.

Table 1: Research Design

	TREATMENT GROUPS			
	T1: QLQ	**T2: (Q+C)LQ**	**T3: Q(Q+C)L**	**T4: Q(L+C)Q**
First mailing	Q	Q + C	Q	Q
Second mailing	L	L	Q + C	L + C
Third mailing	Q	Q	L	Q

Note. Q = Questionnaire plus letter L = Letter only C = Chocolate incentive attached to letter

Survey Instrument

The questionnaire was in the form of an A4 booklet (A3 folded). A reply paid envelope was provided. The questionnaire contained questions on product placement in Reality TV, Internet and computer usage and TV viewing in general, and a suite of demographic questions.

The incentive was a small (45mm by 55mm by 6mm), individually foil-wrapped milk chocolate bar. It was attached to the cover letter with double-sided adhesive tape, with the following statement added to the letter: "As a token of our appreciation, we hope you will enjoy the attached sample of Whittaker's chocolate."

Question Sets

The question sets cited in Table 6 are as follows:

Set1: "Reason watch TV", provided a list of 18 reasons why people watch reality TV. Respondents were asked to rate each item in terms of their importance to them as a reason for watching reality TV, using a 5-point Not-at-all-important to Extremely Important scale.

Set 2: "Beliefs re Reality TV" was a list of eight statements about reality TV shows. Respondents were asked how much they agreed with the statements, using a 5-point Strongly Agree to Strongly Disagree scale.

Set 3: "Response to Reality TV" asked respondents how much they agreed with each of nine statements about reality TV, using a 5-point Strongly Agree to Strongly Disagree scale.

Set 4: "TV viewing" asked respondents how often they watched each of 10 types of Reality TV programs, using a three point scale (Regularly, Occasionally or Never).

Table 2: Effect of Incentive on Post Survey Sample Composition: Age

		AGE OF RESPONDENT									
TREATMENT	N	16-25 %	26-35 %	36-45 %	46-55 %	56-65 %	66+ %	df	$\chi 2$	p	ES_1
T1 QLQ	226 (382)	12.8 (13.1)	19.0 (16.8)	18.1 (19.4)	19.0 (18.8)	15.9 (15.2)	15.0 (16.8)	5	1.36	>.05	.08
T2 (Q+C)LQ	221 (392)	10.0 (12.2)	16.7 (16.6)	21.7 (19.6)	21.7 (20.4)	14.5 (15.2)	15.4 (16.8)	5	1.75	>.05	.09
T3 Q(Q+C)L	247 (391)	8.5 (12.3)	15.8 (16.6)	24.7 (21.5)	19.4 (20.5)	15.0 (13.0)	16.6 (16.1)	5	5.03	>.05	.14
T4 Q(L+C)Q	237 (389)	11.4 (12.3)	13.5 (16.5)	24.1 (22.9)	24.1 (22.1)	15.6 (12.9)	11.4 (13.4)	5	4.07	>.05	.13
T1,2,3,4:	$\chi^2 = 11.77$, df = 15, p = .70; $\Phi_{C(931)} = .07$ $ES_2 = .11$ (based on final sample only)										

Notes: *Row significance test results compare actual response frequencies with expected frequencies calculated using initial sample frequencies. The initial sample sizes and distributions are displayed in brackets ( ).*
ES_1= Effect size = SQRT((C^2)/(1-C^2)) where C = SQRT((χ^2)/(χ^2 + N))
ES_2= Effect size = Φ_C = Cramer's Phi Coefficient = SQRT(χ^2/(N(k-1)) (where k = smaller of r or c)

RESULTS AND DISCUSSION

Sample composition

An issue with using an incentive is that it may bias the results prompting disproportionate responses from different age or gender segments of the sample (Ryu, Couper and Marans, 2006). For this study, information was available about the gender and age of respondents in the initial sample. This data was used to compute the expected frequencies for the analysis of the post survey response distributions for age and gender. For age, the census database provided a five year age range category containing date of birth.

As can be seen in Table 2 (age) and Table 3 (gender), the differences between the initial and final sample compositions did not change significantly for any of the four treatment groups: the Chi-squares are non-significant at the .05 level, and the effect sizes, measured using Cramer's Phi Coefficient (Welkowitz, Cohen & Ewen 2006) are very small. Furthermore, the differences in the post survey distributions across the four treatment groups were also very small and insignificant (Age: $\chi 2$ = 11.77, df = 15, p = .70, Φc = .07; Gender: $\chi 2$ =6.74, df = 3, p = .08, Φc = 0.05), given that an Effect Size of <.1 for a Chi-Square is considered to be of negligible practical significance (Cohen 1987, pp.224-225). Thus one can conclude that the treatments did not affect sample composition with regard to age and gender.

The final sample composition of the four treatment groups with regard to ethnic group and education is presented in Table 4. The original sample composition for these variables is unknown. As for age and gender, there is little variation across the four groups, and the Effect Sizes are small. These results suggest that the use of a chocolate incentive / replacement questionnaire does not bias respondent sample composition, at least with regards to these four variables. The findings for age and gender are consistent with previous findings regarding the effects of incentives on sample composition (e.g., Furse and Stewart, 1982; Hansen, 1980; Nederhof, 1983; Robertson and Bellenger, 1978; Ryu, Couper and Marans, 2006; Willimack et al., 1995; Wotruba, 1966).

Table 3: Effect of Incentive on Post Survey Sample Composition: Gender

TREATMENT	N	M %	F %	df	$\chi 2$	p	ES
T1 QLQ	229 (382)	47.6 (49.0)	52.4 (51.0)	1	0.17	>.05	.03
T2 (Q+C)LQ	227 (392)	42.7 (47.4)	57.3 (52.6)	1	2.03	>.05	.10
T3 Q(Q+C)L	251 (391)	47.8 (54.0)	52.2 (46.0)	1	3.82	>.05	.12
T4 Q(L+C)Q	244 (389)	41.0 (47.0)	59.0 (53.0)	1	3.60	>.05	.12
T1,2,3,4:	χ^2 =3.47, df = 3, p = .32, $\Phi_{c(951)}$ = .06 ES= .12 (based on final sample only)						

Notes: Row significance test results compares actual response frequencies with initial sample frequencies (expected frequencies)
ES= Effect size = $SQRT((C^2/(1-C^2))$ where $C = SQRT((\chi^2)/(\chi^2 + N))$

Table 4: Comparison of the Four Treatment Groups across Respondent Characteristics

	T1 QLQ	T2 (Q+C) LQ	T3 Q(Q+C) L	T4 Q(L+C) Q	Total	n	$\chi 2$	df	p	ES
Ethnic group										
European	71.9	74.3	75.6	75.5	74.4	945	5.06	6	.53	.05
Maori	13.2	9.3	13.2	12.0	11.3					
Other	13.9	19.2	7.7	12.8	13.3					
Education										
<4 yrs Secondary	30.3	37.2	33.7	30.8	29.1	943	6.08	9	.73	.08
>4 yrs Secondary	30.3	27.9	26.9	31.3	19.5					
<4 yrs Tertiary	18.4	16.8	21.7	20.8	18.5					
>4 yrs Tertiary	21.1	18.1	17.7	17.1						
N	237	245	262	257	1001					

Note. ES = Effect Size = Φ_c *= Cramer's Phi Coefficient = SQRT(*χ^2*/(N(k-1)) (where k = smaller of r or c)*

Item non-response bias

Another concern is that an incentive may affect item non-response, and thereby affect the interpretation of responses to survey questions. In this survey, there were seven demographic questions, three questions on computer use, five questions relating to general television viewing, including one question with 10 scale items, and seven questions relating to Reality TV, including one question with 18 items, one with nine items, and one with eight items. Two single item questions were excluded from the analysis because the high level of non-response suggested that people simply skipped the questions because they were not applicable to them (e.g., number of children at home; use of computer at work). The item non-response rates for the single-item questions are reported in Table 5.

A notable result is that item non-response did not differ significantly across treatment groups for any of the 16 single item questions, and the Effect Sizes are all very small, as one would expect. Although T2 (chocolate sent with the first mail-out) consistently produced slightly higher item non-response across questions, suggesting a systematic effect, the differences between treatment groups are not statistically significant.

The absence of significant differences in item non-response across treatment groups suggests that the incentive did not significantly reduce item non-response.

The item non-response for the multi-item questions, and for the 45 items overall, are reported in Table 6. For each individual, an item non-response score was produced for each scale set by counting the number of omitted items in the set. ANOVA's were used to compare treatment groups.

Table 5: Item Non-response for Single Response Questions

	T1 QLQ	T2 (Q+C)LQ	T3 Q(Q+C)L	T4 Q(L+C)Q	Total	$\chi 2$	p	ES
	%	%	%	%	%			
Demographics								
Income	8.4	9.4	8.4	14.0	10.1	6.02	.11	.08
Age	4.6	9.4	5.3	7.8	6.8	5.61	.13	.08
Gender	3.4	7.3	4.2	5.1	5.0	4.52	.21	.07
Ethnic group	3.8	7.8	5.0	6.2	5.7	3.92	.27	.06
Education	3.8	7.8	5.0	6.6	5.8	4.11	.25	.06
Household size	4.7	7.8	5.4	5.5	5.9	2.40	.49	.05
Computer use								
Time spent on web	2.1	4.5	3.1	3.5	3.3	2.22	.53	.05
Use computer at home?	2.5	2.9	1.1	1.6	2.0	2.50	.48	.05
TV/movie viewing								
Hire movie DVDs?	6.3	8.6	6.5	6.6	7.0	1.26	.74	.04
Time spent watching TV.	3.4	6.1	2.3	3.9	3.9	5.22	.16	.07
Record TV movies?	8.4	9.0	12.6	9.7	10.0	2.91	.41	.05
TV prog. recording freq.	1.3	1.6	1.5	2.3	1.7	.94	.82	.03
Reality TV								
Time spent watching RealityTV	11.0	13.5	17.2	12.1	13.5	4.79	.19	.07
Purchased a Reality TV product?	6.3	7.3	6.5	5.4	6.4	.76	.86	.03
Discussed Reality TV products?	6.3	7.3	6.1	5.8	6.4	.54	.91	.02
Visited a Reality TV website?	6.3	7.3	6.1	5.4	6.3	.79	.85	.03
N	237	245	262	257	1001			

Note. The values reported in Table 5 are the % of responses missing for each question. For all items, df = 3. ES = Effect Size = Φc = Cramer's Phi Coefficient = SQRT(χ^2/(N(k-1)) (where k = smaller of r or c)

Table 6: Mean Item Non-response for Questions with Item Sets

	T1 QLQ	T2 (Q+C)LQ	T3 Q(Q+C)L	T4 Q(L+C)Q	df	F	p	ES
	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)				
Set 1. Reason watch TV (18 items)	1.76 (.33)	2.05 (.35)	1.64 (.30)	1.37 (.28)	3,997	.80	.49	n.a.
Set 2. Beliefs re RealityTV (8 items)	.56 (.13)	.72 (.14)	.50 (.12)	.50 (.12)	3,997	.69	.56	n.a.
Set 3. Response to RealityTV (9 items)	.44 (.12)	.43 (.12)	.44 (.12)	.53 (.12)	3,997	.21	.62	n.a.
Set 4. TV viewing (10 items)	.08 (.04)	.12 (.04)	.19 (.06)	.08 (.03)	3,997	1.31	.27	.00
Total set (45 items)	3.94 (.60)	4.84 (.67)	4.11 (.56)	3.80 (.56)	3,997	.60	.67	n.a.
N	237	245	262	257				

Note. *SE = standard error*
ES = Effect Size = $\eta^2 = (df_B(F-1))/(df_B F + df_W)$ *(only meaningful when F>1)*

As with the single item questions reported in Table 5, the item omission tended to be higher for the T2, but the differences between treatment groups were all non-significant at the .05 level. It would therefore seem that the use of an incentive and replacement questionnaires did not have a significant effect on item non-response. It is also clear that the incentive did not reduce item non-response, contrary to the findings reported by McDaniel and Rao (1980).

Response Distribution

A key concern is whether the use of an incentive might affect respondents' responses to the questions that address the substantive issues in the survey. In this study we cannot directly assess measurement error, as we would need to know what the "correct" answers were, and this information is not available. Instead, differences in response distributions across groups are taken as a surrogate measure of measurement error. The rationale used is that if the treatment groups do not differ on demographics, or on item non-response, then a difference across groups in the responses to the substantive questions would be indicative of error introduced by the treatment, rather than due to respondent differences. i.e., differences in response distributions are used as an indirect measure of measurement error. As we do not know the "correct" response for the population, we cannot tell which treatment gives the "best" result – we can only determine whether the different treatments produce different response distributions. This is a problem if the response distributions differ to the extent that we might draw different inferences from the results.

The mean responses for single item questions are shown for each treatment group in Table 7. None of the differences in response distributions across the four treatment groups were statistically significant (at $p<.05$) for any of the behavioural variables listed.

Table 7: Comparison of the Four Treatment Groups for Respondent Behaviour

	$\chi 2$	df	p	ES
Computer use				
Use computer at work (yes/no)	0.81	3	.85	.03
Time spent on web (6)	8.95	15	.88	.06
Use computer at home (yes/no)	2.54	3	.47	.05
Reality TV				
Time spent watching Reality TV (3)	8.84	6	.18	.07
Purchased a product on Reality TV (3)	2.87	6	.82	.04
Discussed Reality TV products (3)	4.27	6	.64	.05
Visited a Reality TV website (2)	2.03	3	.57	.05
TV/movie viewing				
Time spent watching TV (3)	8.39	6	.21	.07
Hire movie DVDs (5)	8.33	12	.76	.06
Record TV movies (5)	12.59	12	.40	.07
Record TV programmes (5)	6.41	12	.89	.05

Note. Number of response categories shown in brackets ( ).
ES= Effect size = Φc (Cramer's Phi Coefficient)

Respondents were also presented with a list of 10 types of TV Reality shows (e.g., romance, adventure, family situations, group living, personal advancement, physical improvement, home gardening, pet shows, community services, celebrity), and asked to indicate how often they watched each type (regularly/occasionally/never). One item "*Watch community service shows*" revealed a significant difference between groups ($F_{(990,3)}$ =2.81 p=.04), but the Effect Size was small ($\acute{\eta}^2$ =.008).

Three sets of attitude questions (comprising 35 statements in total) were also used to explore respondents' attitudes towards Reality TV. Only one set contained items that differed significantly across treatment groups (see Table 8). The groups differed on two items with regards to how strongly they agreed or disagreed with the statements (5 point scale: agree strongly /agree/neutral/disagree/disagree strongly)("*Members of the public usually overact for the cameras*" $F_{(927,3)}$ =3.60, p=.012, $\acute{\eta}^2$ =.009; "*These programmes give real people a chance to speak on TV about what matters to them*" $F_{(923,3)}$ =2.90 p=.034, $\acute{\eta}^2$ =.006) . However, the effect sizes for these two were extremely small and not practically significant (see rightmost column in Table 8).

For a question listing 18 reasons for watching Reality TV, no items differed across groups in terms of importance (4 point scale: not at all important/somewhat important/important/extremely important). And in a question comprising nine statements about Reality TV, none of the items differed across groups with regards to how strongly they agreed or disagreed with the statements (5 point scale).

Table 8: Comparison of the Four Treatment Groups for Respondent Attitudes

	T1 QLQ	T2 (Q+C)LQ	T3 Q(Q+C)L	T4 Q(L+C)Q	df	F	p	ES
Real life stories are more entertaining than fiction	2.67 (.069)	2.64 (.073)	2.77 (.071)	2.63 (.068)	3, 927	.880	.451	n.a.
Members of the public usually overact for the cameras	2.30 (.061)	2.20 (.054)	2.45 (.054)	2.38 (.052)	3, 927	3.670	.012	.009
I don't like watching TV programmes where real people face difficult physical situations	3.03 (.064)	2.96 (.071)	3.05 (.065)	3.08 (.063)	3, 923	.560	.641	n.a.
These programmes give real people a chance to speak on TV about what matters to them	2.86 (.065)	2.78 (.063)	3.01 (.060)	2.80 (.062)	3, 923	2.904	.034	.006
Members of the public usually act the same on TV as in real life	3.69 (.060)	3.66 (.062)	3.68 (.053)	3.59 (.056)	3, 926	.666	.573	n.a.
I don't like watching TV programmes where real people face difficult emotional situations	2.88 (.065)	2.74 (.069)	2.84 (.064)	2.82 (.064)	3, 926	.870	.456	n.a.
I think these programmes are really useful as they give you all sorts of information about life	3.20 (.076)	3.06 (.077)	3.22 (.072)	3.12 (.069)	3, 926	.924	.428	n.a.
I think Reality TV takes advantage of people who are in the programmes	2.51 (.068)	2.53 (.065)	2.59 (.062)	2.55 (.065)	3, 929	.252	.860	n.a.

Note. ES = Effect Size = $\hat{\eta}^2 = (df_B(F-1))/(df_B F + df_W)$ *(only meaningful when F>1)*

Given the large number of statements (45) and with an alpha of .05, one would expect two significant results due to chance (45 x.05 = 2.25), so it is quite possible that the statistically significant differences are simply due to sampling error.

Overall, there is no strong evidence that the chocolate incentive had any significant effect on measurement error, either positive or negative.

CONCLUSION

The results of this study demonstrate that using a chocolate incentive had no statistically significant effect on sample composition (age, gender, education, ethnic group). Thus this incentive does not appear to introduce sample bias.

The findings do not support the idea that an incentive may reduce item non-response (McDaniel and Rao 1980). In this study there was a slight increase in item non-response associated with the use of the incentive, although the difference between treatments was statistically non-significant and the effect sizes were small, suggesting that these differences were due to sampling error. Even so, it would seem prudent to monitor for this effect in surveys using this incentive by using a hold-out sample as a control, as there is clearly a potential for an interaction between the type of incentive used and the topic of a survey.

Another important finding was that the treatments, and by implication, the incentive, did not significantly affect response distributions, and so presumably had minimal impact on measurement error. It should be noted, however, that there was no reason to suspect that the incentive should have had a confounding effect in this particular survey, given the survey topic and the nature of the questions. But given that it is conceivable that a chocolate incentive could influence responses to topics relating to diet, health or confectionary, the potential for such effects needs to be kept in mind. Further research across a wider range of topics is needed to determine whether significant topic/incentive interactions are likely.

Given the positive effect on increasing the speed of response exhibited by both the chocolate (first mail-out) and the replacement questionnaire (second mail-out) (see Brennan & Charbonneau 2009) and the minimal negative effects on sample bias, item non-response, measurement error, or survey cost (see Brennan & Charbonneau, 2009), a chocolate incentive sent with the first mail-out, and a replacement questionnaire rather than a reminder letter sent with the first reminder, appear to be acceptable procedures for mail surveys of the general public.

REFERENCES

Beatty, P., Hermann, D., 2002. To answer or not to answer: Decision processes related to survey item nonresponse. In Groves, R.M, Dillman, D.A., Eltinge, J.E., Little, R.J.A. (Eds.), *Survey Nonresponse*. New York: Wiley, pp.71-86.

Bednall, D.H.B., Shaw, M., 2003. Changing response rates in Australian market research. *Australian Journal of Market Research* 11(3), 31-41.

Brennan, M., 1992. The effect of a monetary incentive on mail survey response rates: New data. *Journal of the Market Research Society*, April, 34 (2), 173-177.

Brennan, M., 2008. A Test of Two Procedures for Increasing Mail Survey Response Rates: A Chocolate Incentive and a Replacement Questionnaire. Paper presented at the Australia and New Zealand Marketing Academy Conference, Sydney, Australia, December 1 to 3, 2008.

Brennan, M., Charbonneau, J., 2009. Improving mail survey response rates using chocolate and replacement questionnaires. *Public Opinion Quarterly*, 73 (2), 368-378.

Brennan, M., Xu X., 2009. The effect of questionnaire colour, a chocolate incentive and a replacement return envelope on mail survey response rates. *Australasian Journal of Market and Social Research*, 17(1), June, 29-35.

CMOR., 2003. *Respondent Cooperation and Industry Image Study*. The Council for Marketing and Opinion Research, Port Jefferson, New York. Available online at http://www.cmor.org/resp_coop_studies.cfm (accessed December 5, 2006).

Cohen, J., 1987. *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Erlbaum.

Curtin, R., Presser, S., Singer, E., 2005. Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly* 69, 87-98.

de Leeuw, E., de Heer, W. 2002. Trends in household survey nonresponse: A longitudinal and international comparison." In Groves, R.M, Dillman, D.A., Eltinge, J.E., Little, R.J.A. (Eds.), *Survey Nonresponse*. New York: Wiley, pp. 41-54.

de Leeuw, E., 2001. Reducing missing data in surveys: An overview of methods. *Quality and Quantity* 35, 147-60.

Finn, D.W., Wang, C-K., Lamb, C.W., 1983. An examination of the effects of sample composition bias in a mail survey. *Journal of the Market Research Society* 25, 331-338.

Furse, D.H., Stewart, D.W., 1982. Monetary incentives versus promised contribution to charity: New evidence on mail survey response. *Journal of Marketing Research* XIX, 375-80.

Gendall, P., Leong, M. & Healey, B,. 2005. *The effect of pre-paid non-monetary incentives in mail surveys*. In proceedings of ANZMAC 2005 Conference, December 5-7, Freemantle, Australia. (CDROM) ISBN 0-646-45546-X.

Goritz, A. 2004. The impact of material incentives on response quantity, response quality, sample composition, survey outcome, and cost in online access panels. *International Journal of Market Research*, 46 (3), 327-345.

Groves, R.M., 2006a. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly* 70, 646-75.

Groves, R.M., Couper, M.P., Presser, S., Singer, E., Tourangeau, R., Acosta, G.P., Nelson, L., 2006b. Experiments in producing nonresponse bias. *Public Opinion Quarterly* 70, 720-36.

Groves, R.M., Couper, M., 1998. *Nonresponse in Household Interview Surveys*. John Wiley & Sons, New York.

Hansen, R.A., 1980. A self-perception interpretation of the effect of monetary and nonmonetary incentives on mail survey respondent behaviour. *Journal of Marketing Research* XVII, 77-83.

James, J.M., Bolstein, R., 1990. The effect of monetary incentives and follow-up mailings on response rate and response quality in mail surveys. *Public Opinion Quarterly* 54, 346-361.

McDaniel, S.W., Rao, C.P., 1980. The effect of monetary inducement on mailed questionnaire response quality. *Journal of Marketing Research* XVII, 265-68.

Mizes, S.J., Fleece, E.L., Roos, C., 1984. Incentives for increasing return rates: Magnitude levels, return bias and format. *Public Opinion Quarterly* 48, 794-00.

Nederhof, A.J., 1983. The effects of material incentives in mail surveys: Two studies. *Public Opinion Quarterly* 47, 103-11.

Olson, K., 2006a. Survey participation, nonresponse bias, measurement error bias, and total bias. *Public Opinion Quarterly* 70, 737-58.

Olson,K., Kennedy, C,. 2006b. *Examination of the relationship between nonresponse and measurement error in a validation study of alumni*. Paper presented at the 61st Annual Conference of the American Association for Public Opinion Research, Montreal, May 18 - 21. ASA Section on Survey Research Methods, 4181-4188.

Robertson, D.H., Bellenger, D.N., 1978. A new method of increasing mail survey responses: contributions to charity. *Journal of Marketing Research* 15, 632-33.

Ryu, E., Couper, M.P, Marans, R.W., 2006. Survey incentives: Cash vs in-kind; face-to-face vs mail; response rate vs nonresponse error." *International Journal of Public Opinion Research* 18, 89-106.

Welkowitz, J., Cohen, B.H., & Ewen, R.B. 2006. *Introductory Statistics for the Behavioural Sciences (6th edition)*, New Jersey: John Wiley & Sons.

Whitmore, W.J., 1976. Mail survey premiums and response bias. *Journal of Marketing Research* XII (February), 46-50.

Willimack, D., Schuman, H., Pennell, B.E., Lepkowski, J., 1995. Effects of a pre-paid non-monetary incentive on response rates and response quality in a face-to-face survey. *Public Opinion Quarterly* 59, 78-92.

Wotruba, T.R., 1966. Monetary inducements and mail survey response. *Journal of Marketing Research* 3, 393-400.

Yan, T., Jans, M., Curtin, R., 2006. *Changes in nonresponse to income questions*. Paper presented at the 61st Annual Conference of the American Association for Public Opinion Research, Montreal, May 18 - 21.

Address correspondence to :

Mike Brennan
Department of Communication,
Journalism & Marketing
Massey University
Private Bag 11222,
Palmerston North, New Zealand.
Ph +66 6 356 5576
Email: m.brennan@massey.ac.nz
(corresponding author)