The Canadian Journal of Program Evaluation Vol. 12 No. 2 Pages 81–104
ISSN 0834-1516

ETHICAL AND METHODOLOGICAL IMPROVEMENTS TO MAIL SURVEY RESEARCH: AN ALTERNATIVE FOLLOW-UP METHOD

P. D. Sully
P. R. Grant
Department of Psychology
University of Saskatchewan

Abstract: Follow-up contacts increase response rates to mail surveys substantially. However, for individuals who do not wish to participate such repeated contacts are intrusive, creating an ethical concern. An alternative follow-up procedure was developed that allows contacted individuals to choose not to participate by returning a "Reasons for Not Responding" (RNR) form. This study was designed to compare the RNR method with a traditional follow-up procedure, the Total Design method (TDM). Specifically, a two (follow-up method) by two (early versus late return) between-subjects design was used to examine satisfaction with and quality of responses to a survey mailed to 300 randomly selected households. The results showed that respondents in the RNR conditions were more satisfied with the research than respondents in the TDM conditions, particularly after repeated contacts. Further, although response quality was equivalent for the two methods in the early returns, it was lower in the TDM group, but not in the RNR group, after repeated contacts. These results suggest that mail survey researchers should seriously consider using the RNR method, especially because (a) the overall response rate to the two methods was very similar, (b) the nonrespondents frequently indicated on their RNR forms that they had made an informed decision not to participate, and (c) the RNR form allows researchers to collect information on the demographic characteristics of nonrespondents.

Résumé: Les contacts de suivi améliorent considérablement les taux de réponse aux enquêtes postales. Cependant, les individus qui ne veulent pas y participer trouvent ces communications répétées intrusives, ce qui soulève une question d'éthique. C'est pourquoi un processus de rechange permet à ces individus de décliner l'invitation à répondre en remplissant un formulaire de non-participation (RNR). Cette étude compare la méthode du

RNR avec le processus traditionnel de suivi, la méthode de conception totale (TDM). Plus précisément, une conception 2/2 entre sujets (méthode du suivi et réponse précoce par rapport à réponse tardive) a permis d'examiner la satisfaction des répondants et la qualité des réponses à un sondage envoyé par courrier à 300 ménages choisis au hasard. Les résultats montrent que les répondants ayant rempli le RNR étaient plus satisfaits de l'enquête que ceux auxquels la TDM avait été appliquée, particulièrement après des contacts répétés. En outre, bien que la qualité des réponses se valait pour les deux méthodes dans les réponses précoces, elle a baissé dans le groupe de la TDM, et pas dans le groupe RNR, après plusieurs contacts. Ces résultats suggèrent que les administrateurs d'enquêtes postales devraient envisager sérieusement d'utiliser la méthode du RNR, surtout parce que a) les taux généraux de réponse pour les deux méthodes étaient très semblables, b) les non-répondants ont fréquemment indiqué sur le RNR qu'ils avaient pesé leur décision, et c) le RNR permet aux chercheurs de recueillir des renseignements sur les caractéristiques démographiques des non-répondants.

Mail surveys have many qualities that are of value for data collection. Some of the most important of these are their low cost of administration, their ability to reach a target population scattered across a large geographic area, and the increased validity created when mail survey respondents consult personal documents, family members, friends, or co-workers to ensure the accuracy of their responses (Faria, Dickenson, & Filipic, 1990; Kanuk & Berenson, 1975; Linsky, 1975). However, these advantages are offset or diminished by a low response rate, a result that is often obtained when this survey method is used. Researchers have therefore devoted much effort to developing effective methods that will increase response rates to mailed surveys. This increasing refinement of mail survey methodology has yielded response rates from general population samples that rival those from personal interviews (Dillman, 1978, 1991; Faria, Dickenson, & Filipic, 1990), with the single most important improvement being the use of follow-up contacts (Dillman, 1991; Lockhart, 1984). Indeed, Dillman (1978) claims that the inclusion of follow-up mailings can double response rates to mail surveys.

The ability of follow-up contacts to increase response rates is supported by a substantial body of research. Specifically, Heberlein and Baumgartner (1978) conducted a quantitative analysis on a large sample of the published research investigating factors that affect

response rates to mailed surveys. The authors coded 71 factors reported in 98 methodological studies of mailed surveys and examined their effect on response rates. The results of this study show that 42% of the variance in final response rates can be attributed to the number of follow-up contacts, with each contact increasing the predicted response rate by approximately 12%. Numerous studies have replicated these results with more recently published research, and found the same strong effects of follow-up contacts (Eichner & Habermehl, 1981; Goyder, 1982; Yu & Cooper, 1983).

Although traditional follow-up contacts may increase response rates and thereby decrease the effects of nonresponse bias, these methods may introduce ethical and methodological problems. In particular, traditional follow-up methods do not distinguish between individuals who do not wish to complete the survey (nonconsenters) and individuals who have misplaced or forgotten to return their surveys. As a result, follow-up contacts are mailed to people who have chosen not to participate. These second, third, and fourth mailings may result in nonconsenters feeling pressured to complete and return a survey instrument, or to hastily and inaccurately return a survey instrument merely to prevent further follow-up contacts. Many researchers (Nevin & Ford, 1976; Ray & Still, 1987; Sudman, 1985) have observed that conducting follow-up contacts in this manner may have negative consequences for participants and may decrease the response quality of mail surveys. However, without follow-up contacts researchers are faced with the problem of low response rates.

Given this dilemma, an alternative follow-up contact method that allows researchers to differentiate nonconsenters from consenters who have misplaced or forgotten to return their surveys may be of value to many researchers who are dissatisfied with traditional follow-up contact methods. However, this is true only if the traditional method does in fact have a negative impact on participants and does decrease the quality of their responses. Therefore, the current research investigated the claim that traditional follow-up contacts do have a negative impact on survey participants and examined the impact of an alternative follow-up contacts method. That is, the research was designed to provide information relevant to the ethical and methodological soundness of a new, alternative follow-up contact method in comparison with a traditional follow-up contact method.

POTENTIAL ETHICAL PROBLEMS WITH TRADITIONAL FOLLOW-UP CONTACTS

Two aspects of traditional follow-up contacts may have a negative impact on potential respondents. The first is that the follow-up letters often use persuasive appeals on the assumption that nonrespondents are not making well-thought-out and informed decisions. Because researchers using traditional follow-up methods are unaware of the reasons for nonresponse, this assumption may be inaccurate. Therefore, the persuasive appeals may place unnecessary pressure on initial nonconsenters who have appropriately expressed their right to decline participation.

The second aspect is the fact that individuals who have decided not to participate and thus have discarded their survey instruments often receive two, three, or four follow-up contacts. These persistent follow-up mailings combined with strong persuasive appeals from a perceived high-status source may indicate to nonconsenters that they are not really free to decline participation, or that their choice to decline participation was not respected. If this is true, such follow-up procedures would be in violation of most ethical codes of conduct followed by applied researchers. For example, the first principle of the Canadian Psychological Association's (1991) code of ethics states that researchers must "take all reasonable steps to ensure that consent is not given under conditions of coercion or undue pressure" (principle I.20) and must "respect the right of individuals to discontinue participation or service at any time, and be responsive to nonverbal indications of a desire to discontinue if the individual has difficulty with verbally communicating such a desire" (principle I.23).

Although there is no empirical evidence that researchers do, in fact, apply "unduly persuasive techniques," there are a number of examples in the literature that suggest this is indeed the case. For instance, in an early attempt to decrease so-called "errors due to low return rates," Eckland examined "the effects of prodding to increase mail-back returns" (1965, p. 165). According to Eckland, potential participants who had not responded to initial follow-up contacts are "resistant subjects" who require "intensive prodding" in order to get them to return a survey instrument. Eckland advises that so-called resistant subjects should be prodded with three regular follow-up mailings, one certified follow-up mailing, and a follow-up telephone contact. However, because traditional follow-up contacts do not obtain information regarding the reasons for nonresponse, Eckland's

suggestion that the nonrespondents in mail survey research are "resistant subjects" is not based on any information. In fact, it is more in keeping with ethical guidelines to assume that such individuals have made a well-informed decision not to participate.

These problems with traditional follow-up contacts appear to exist in the contemporary research literature as well. An examination of the follow-up procedures described in the Total Design method of mail survey research (Dillman, 1978) reveals that current methods and attitudes are similar to Eckland's (1965). The objective of the three TDM follow-up contacts is to persuade initial nonrespondents to complete a survey instrument. As Dillman states, "each follow-up mailing differs somewhat from the one that preceded it, as attempts are made to invoke new more persuasive appeals … The appeals are designed to crescendo, with later follow-ups being stronger attempts at persuasion than the preceding ones" (1978, p. 182).

Dillman (1978), like Eckland (1965), suggests that the numerous follow-ups and strong attempts at persuasion are appropriate because nonrespondents are "inappropriately resistant" to research. Dillman states that those who require follow-up contacts have "ignored … previous mailings and therefore might be classified as 'hard core holdouts'" (p. 189). Once again, this perception of nonrespondents as "ignoring" and "holding out on the researcher is not based on any information about their actual motives. Indeed, individuals who have appropriately expressed their desire not to participate may be unduly pressured to respond. Nevin and Ford (1976) agree, suggesting that follow-up contacts such as Dillman's increase responding because they involve a coercive element. They claim that "[Dillman's] follow-up letter contain[s] a veiled threat, that is, if you do not respond to this questionnaire we will send you another questionnaire and another until you take the time to respond" (p.116). Indeed, after using follow-up techniques that adhered closely to the procedures advocated by Dillman, a number of researchers (Heberlein & Baumgartner, 1978; Sudman, 1985; Wellman, Hawk, Roggenbuck & Buhyoff, 1980) have come to similar conclusions.

To conclude, the purpose of traditional follow-up contacts is to extend another chance to those nonrespondents who have forgotten to return or misplaced their survey instruments and to those who are "inappropriately resistant" to research. The problem is that researchers using traditional follow-up contacts are unaware of the reasons for nonresponse and cannot identify nonconsenters. Indeed, it seems

likely that persuasive follow-up contacts are being received by nonconsenters who had initially made a well-informed decision not to participate and who are being inappropriately pressured to respond. However, the literature on the negative effects of follow-up contacts is anecdotal. Therefore, research is needed to investigate the impact of traditional follow-up contacts on initial nonrespondents.

THE EFFECT OF TRADITIONAL FOLLOW-UP CONTACTS ON RESPONSE QUALITY

Not only may persistent follow-up mailings combined with persuasive inducement letters place undue pressure on initial nonconsenters to participate, it may also result in surveys being responded to carelessly or even vindictively merely to prevent further follow-up mailings. That is, traditional follow-up methods may result in a decline in the quality of mail survey data.

Both psychological theory and empirical evidence suggest that traditional follow-up contacts cause a decrease in response quality. For example, reactance theory states that when an individual is urged or pressured to comply her or his freedom is threatened, and therefore the individual feels compelled to reassert her or his freedom (Brehm, 1966). The more pressure used, the more the individual feels threatened, and the more reactance is aroused. Therefore, traditional follow-up contacts may result in decreased validity of responding because the pressure applied by follow-up contacts may result in some form of reactance.

Many researchers using follow-up contacts have made observations that are consistent with reactance theory (Dillman, 1991; Hawkins, 1979; Ray & Still, 1987). Wellman et al. summarized these researchers' observations when they stated that "if late respondents are less motivated to complete the survey, but are acting under the pressure of repeated follow-ups, they might be expected to put less effort into their answers. Thus, it might be the case that gains in sample representativeness would be offset by losses in data quality and attendant declines in explanatory power and validity" (1980, p. 171).

There are two studies that provide some empirical support for the claim that reactance is created by pressuring potential participants with follow-up contacts and that this reactance takes the form of acquiescent or invalid responding. Doob and Zabrack (1971) found

that no matter how intense or freedom-threatening the follow-up contact used, there was no difference in response rates. An unexpected finding, however, indicated that the reactance caused by the increased pressure to respond was not in the form of decreased responding, but decreased quality of responding. Unfortunately, Doob and Zabrack's study was not designed to examine the effect of reactance on response quality; thus, this conclusion should be interpreted with caution.

More recently, Ray and Still (1987) directly examined the degree to which follow-up techniques caused reluctant responders to provide meaningless data due to acquiescence. The results of this study show that the use of follow-up contacts may cause a "serious acquiescence problem" (p. 572). The authors concluded that although follow-up contacts are successful at improving the overall response rate, this is accompanied by a much greater amount of acquiescence response bias in the data.

To conclude, psychological theory and empirical evidence suggest that traditional follow-up procedures may result in poor response quality, due, for example, to reactant responding. Therefore, the development of an alternative follow-up method may be of great value to researchers using mail survey methods.

AN ALTERNATIVE FOLLOW-UP CONTACT METHOD

This research compares a new alternative follow-up method, the Reasons for Not Responding (RNR) method, with an often-used traditional follow-up method developed by Dillman (1978), the Total Design method (TDM). The alternative method was designed to allow researchers to distinguish true nonconsenters from potential participants who have misplaced or forgotten to return their survey, so that only the latter receive further follow-up. The RNR method was further designed to allow the researcher to determine nonconsenters' reasons for not wishing to take part in the research and their demographic characteristics.

The key design element of the alternative follow-up method is the inclusion of a Reason(s) for Not Responding form in the initial mail-out package. The RNR form has three purposes: (1) to identify those people in the sample who do not wish to participate, (2) to ask nonconsenters why they have chosen not to participate, and (3) to ask nonconsenters to provide some demographic information.

The recipient of the initial mail-out package is asked in a cover letter to fill out the RNR form only if they do not wish to participate in the research project. They are told that the information on the RNR form is important both because it ensures they will not receive further follow-up mailings and because the researcher can use this information to improve current and future research. Each RNR form contains an identifying code number so that the researcher can remove nonparticipants' addresses from the initial mail-out list when the form is returned.

Individuals who have not returned a survey or an RNR form within three weeks after the initial mailing have most likely chosen to participate but have forgotten to mail back or have misplaced their surveys. These individuals are sent a follow-up letter, a second survey, and another RNR form. The follow-up letter states that the researcher only intends to contact those individuals who have chosen to participate but did not receive, misplaced, or forgot to return the initial survey. It goes on to say that if the individual has chosen not to participate but did not return the RNR form, they should not feel pressured to participate, but should fill out the second RNR form so that no further follow-up mailings are sent to them.

Seven weeks after the initial mailing a second follow-up contact is carried out in the same manner as the first.

Conducting follow-up procedures in this manner is believed to result in very few nonconsenters feeling pressured by the researcher to return a survey instrument. It is assumed that decreasing the pressure to respond will make mail survey procedures less invasive and will improve response quality by decreasing the amount of reactant and acquiescent responding. These assumptions underlie the hypotheses tested in this study. Hypothesis 1 is that repeated follow-up contacts result in a decline in respondents' satisfaction with the research and the researcher when TDM is used, but not when the RNR method is used. Hypothesis 2 is that repeated follow-up contacts result in a decline in the quality of the participants' responses when TDM is used, but not when the RNR method is used.

METHOD

Respondents

The respondents were obtained from 300 residences randomly selected from the telephone directory of a small midwestern Cana-

dian city. In the cover letter, one adult individual from each residence was asked to participate.

Procedure

The research project was "piggy-backed" onto a mail survey.[1] The residences selected to receive the survey were randomly assigned to one of two different methods of follow-up. The first group (n = 150) received a traditional follow-up procedure that adhered closely to the Total Design method created by Dillman (1978). The TDM procedure involves four separate mailings. The initial mailing includes a cover letter, a survey, and a stamped return envelope. The first follow-up mailing is a postcard reminder sent to everyone one week later. The second follow-up mailing is sent only to nonrespondents three weeks after the initial mail-out date. This mailing includes a cover letter reminding nonrespondents of the research and a replacement questionnaire. The third follow-up mailing is sent to nonrespondents seven weeks after the initial mail-out date. This mailing also includes a letter reminding participants of the research and a replacement questionnaire. Dillman (1978) suggests that the third follow-up mailing should be sent by certified mail. However, this procedure is very costly; it was therefore sent by regular mail in this study.[2]

The second group (n = 150) received the new RNR follow-up method. This method is the same as the TDM with three exceptions: (1) the original mail-out package and the second and third follow-up mailings include a "Reasons for Not Responding" form; (2) the second and third follow-up mailings are sent only to those households that do not return an RNR form or a completed survey; and (3) a filter question is included in all the cover letters. The filter question states: "Having read a description of this research project do you wish to participate in this research?" If the recipient of the letter circles "yes," he or she is instructed to discard the Reasons for Not Responding form and to complete the survey. If the recipient of the letter circles "no," he or she is instructed to complete the RNR form, which begins with the following paragraphs:

> We would like to take this opportunity to thank you for considering this research and to state that we respect your decision not to participate. If you have the time, please fill out this brief "Reason[s] for Not Responding" form and return it in the postage paid return envelope.

> The information on the "Reason[s] for Not Responding" form is important for three reasons: (1) you will not receive follow-up reminder letters, (2) it will help us understand why you have chosen not to respond, and (3) it will help us determine to which segment of the population the results of this research apply.
>
> If you should choose not to fill out this "Reason[s] for Not Responding" form, please mail back the blank form in the postage paid return envelope so that we do not disturb you with follow-up letters designed to remind those who have chosen to participate but have forgotten to return their surveys. Thank you once again for your time and consideration.

Nonrepondents then respond to five questions on the RNR form. First, they give the reason(s) why they have decided not to participate in the research by circling one or more of the following options: "do not have the time," "not in the mood," "do not like doing research projects," "do not feel that my answers will remain confidential," "not interested in the research topic," "feel that research invades my privacy," and "do not like the way the research is being conducted." They are also allowed to give other reasons for their decision. Then nonrespondents indicate how many times they have been contacted, their age, their sex, and their education level.

Measures

Participant satisfaction. Satisfaction with the research was measured by six questions in a "Participant Satisfaction" scale included at the end of the survey. Respondents used seven-point scales to rate their willingness to participate in future surveys, their motivation to respond, the importance of the research topic, and their satisfaction with the manner and the professionalism with which the survey was explained and conducted (see Appendix A). Two small pilot studies were used to ensure that the satisfaction scale was understandable and reliable (in this survey, Cronbach's alpha = .86, N = 159).

Response quality. Clearly, respondents who feel pressure to complete a survey might do so hastily, resulting in the omission of items. Therefore, our main measure of response quality was the number of items in the survey that were left unanswered, a standard measure that has been used by many researchers (Childers & Skinner, 1985;

Faria, Dickenson, & Filipic, 1990; Stone, Stone, & Gueutal, 1990). James and Bolstein (1990) have also shown that the number of words written in response to an open-ended question is a sensitive measure of the effort exerted by a respondent in completing a survey. Therefore, the number of words provided to an open-ended question was included as a second index of response quality.

Design and Analysis. Follow-up method (RNR vs. TDM group) and time (early vs. late return) were the factors in a between-subjects two-by-two factorial design. The data from those who responded to the initial mailing and the postcard reminder (first follow-up) were collapsed into an "early return" group, and the data from those who responded to the second and third follow-up were collapsed into a "late return" group. The small number of participants (n = 4) returning surveys after receiving the third follow-up necessitated collapsing the responses to the second and third follow-up.

Analysis of variance was used to analyze this factorial design. Support for the hypotheses is shown by a decline in respondent satisfaction and response quality in the TDM group in comparison to the RNR group. Therefore, we expected to obtain an interaction between group and time of return showing this differential decline. If this interaction was obtained, we used simple main effects analysis to test whether the form of the interaction was as predicted. Because all respondents received the initial mailing and the follow-up reminder postcard, we expected there would only be a difference between the groups at the late return stage. That is, we expected simple main effects analyses comparing the RNR with the TDM respondents to show significant differences between the late but not the early return groups. In addition, the RNR follow-up method should prevent late responders from becoming disgruntled at being repeatedly contacted. Therefore, we used a simple main effects analysis to compare early with late responders within the RNR group to test the extent to which this was achieved.

RESULTS

Response Rate

Of the 300 survey packages mailed, 18 were returned undelivered. Six of the undelivered surveys were from the alternative (RNR) follow-up group, and twelve were from the traditional (TDM) follow-up group, resulting in an overall sample size of 282, with 144 in the alternative group and 138 in the traditional group.[3]

The response rate for the two groups was almost identical (RNR group: 57.6%, n = 83; TDM group: 60.9%, n = 84). In addition, 18.1% (26) of the RNR group returned a Reasons for Not Responding form.

Respondents

Information regarding the participants' age, sex, and education was obtained. These demographic variables were also obtained from some of the nonparticipants in the alternative follow-up group who returned their RNR forms.

Both the RNR group and the TDM group were very similar in terms of sex, age, and education (Table 1). Approximately three quarters of the sample were women; 77.1% of the RNR group, 71.4% of the TDM group, χ^2 (1, N = 165) < 1, ns. The mean age range for the sample was 36 to 45 years; $M_{\text{rnr}} = 3.02$, $M_{\text{tdm}} = 3.09$; $t(165) < 1$, n.s., where "3" means 36 to 45 years. Even though the survey was carried out in a university city, a surprisingly large number of respondents (67, or 40.1%) had a university education. Importantly, this high mean education level was not significantly different in the two groups ($M_{\text{rnr}} = 3.95$, $M_{\text{tdm}} = 3.86$; t (163) <1, n.s., where "4" means trade or business college). There were no significant differences in age or education between the male and female respondents.

Nonrespondents

Of the 144 residences that received mail-out packages in the RNR group, 61 (42.4%) did not return a survey. Demographic information is available from 18 (29.5%) of these respondents, who completed the RNR form (Table 2). Although the small sample size precluded statistical tests, it appears as though the male nonrespondents were older (71.4% were over 55 years old) and had a lower level of education than male respondents. In contrast, female nonrespondents were younger than the female respondents (54.5% were between the ages of 26 and 35 years old) but had a similar level of education.

Respondent Satisfaction

Hypothesis 1 states that repeated follow-up contacts result in a decline in the respondents' satisfaction with the research and the researcher when the TDM is used, but not when the RNR method is

Table 1
Demographics of the Survey Respondents by Group

	RNR Method		TD Method		Total	
	N	%	*N*	%	*N*	%
Sex						
Female	64	77.1	60	71.4	124	74.3
Male	19	22.9	22	28.6	41	24.5
Age						
18–25	15	18.0	13	15.5	28	16.8
26–35	18	21.7	19	22.6	37	22.1
36–45	18	21.7	21	25.0	39	23.3
46–55	14	16.9	9	10.7	23	13.8
56+	18	21.7	22	26.2	40	24.0
Education						
Grade 8 or less	2	2.4	1	1.2	3	1.8
Some high school	10	12.0	11	13.1	21	12.6
Completed high school	12	14.5	20	23.8	32	19.2
Trade/business college	24	29.0	18	21.4	42	25.1
University	34	40.9	33	39.3	67	40.1

Note. Two respondents did not give their level of education and another two did not give their gender.

Table 2
Demographics of the Nonrespondents Completing the RNR Form

	Women		Men		Total	
	N	%	*N*	%	*N*	%
Age						
18–25	2	18.2	0	00.0	2	11.1
26–35	6	54.5	0	00.0	6	33.3
36–45	0	00.0	1	14.3	1	5.6
46–55	0	00.0	1	14.3	1	5.6
56+	3	27.3	5	71.4	8	44.4
Education						
Some high school	2	18.2	1	14.3	3	16.7
Completed high school	1	9.1	2	28.6	3	16.7
Trade/business college	4	36.4	2	28.6	6	33.3
University	4	36.4	2	28.6	6	33.3

used. This hypothesis was tested using a two (RNR vs. TDM) by two (early vs. late return) between-subjects analysis of variance. This analysis revealed a significant main effect for follow-up method: the RNR group reported greater satisfaction (M = 5.33) than did the

Table 3
Satisfaction and Response Quality as a Function of Follow-Up Procedure and Time of Return: Means, Standard Deviations, and Summary Statistics

	Follow-Up Procedure				Interaction		
	RNR		TDM				
	Early	Late	Early	Late	*F*	df	MSE
Satisfaction	5.38	5.07	5.02	3.96	2.85+	1,161	1.14
Standard deviation	1.04	1.06	1.11	1.04			
No. of respondents	(70)	(12)	(67)	(16)			
Item omission[a]	0.45	0.17	0.43	1.47	5.94*	1,163	1.72
Standard deviation	1.16	0.39	1.02	2.70			
No. of respondents	(71)	(12)	(67)	(17)			
Respondents omitting at least one item	21.13% (71)	16.67% (12)	22.39% (67)	35.29% (17)			
Open-ended question	26.03	39.58	27.33	27.00	3.99*	1,162	276.22
Standard deviation	15.17	23.83	16.18	18.48			
No. of respondents	(71)	(12)	(67)	(16)			

+ $p < .10$; * $p < .05$

Note. High numbers indicate greater satisfaction (range 1 to 7), more items not answered in the survey, and more words in response to the open-ended question. The percentage of respondents who did not respond to at least one item in the survey is also shown.

[a] Two respondents in the TDM—late returns group did not respond to a large number of the survey items (one missed 22 and the other 66 out of a possible 67 items), with the rest of the sample missing 7 or fewer items. Therefore, the number of items omitted for these two respondents was recoded as "8" when this dependent variable was analyzed. The adjusted mean is shown in the table; the raw mean is 5.47.

TDM group ($M = 4.82$; $F(1,161) = 10.81$, $p < .001$), and a significant main effect for number of contacts: the early return respondents reported greater satisfaction ($M = 5.20$) than the late return respondents ($M = 4.43$); $F(1,161) = 8.34$, $p < .01$.

The top row of Table 3 gives the mean satisfaction ratings for the four groups in this design and shows that these main effects were qualified by a marginally significant interaction. In support of hypothesis 1, simple main effects showed that the TDM—late return group was significantly less satisfied than the RNR—late return group; $F(1,161) = 7.39$, $p < .01$. As well, the RNR respondents did not show a significant decline in satisfaction between those who responded early and those who responded later; $F < 1$. Unexpectedly,

the TDM—early return group was slightly but significantly less satisfied than the RNR—early return group ($F(1,161) = 3.89$, $p < .05$), indicating that the two groups were not equivalent in terms of satisfaction, even at the early return stage.

Response Quality

Hypothesis 2 states that repeated follow-up contacts result in a decline in the quality of participants' responses when the TDM is used, but not when the RNR method is used. In this study, the main measure of response quality was the number of items left unanswered in the survey. A two (RNR vs. TDM) by two (early return vs. late return) between-subjects analysis of variance was used to analyze this measure. No significant main effects were obtained, but there was a significant interaction (second row of Table 3). The pattern of means shows considerable support for the hypothesis. Specifically, simple main effects showed that there was no significant difference between the TDM group and the RNR group for the early return participants ($F < 1$); however, for late return participants, the TDM group left a significantly larger number of items unanswered than the RNR group; $F(1,163) = 6.92$, $p < .01$. Also consistent with hypothesis 2, the RNR group did not show a significant difference in the number of item omissions between the early return group and the late returns group; $F < 1$. That is, the RNR follow-up method did not result in a decline in response quality as indexed by this measure. The third row of Table 3 shows the percentage of respondents who neglected to answer at least one survey item. Again, these results support the hypothesis, with 35.3% of the respondents in the TDM—late return group neglecting to answer at least one survey item, in comparison with an average of 21.3% of the respondents in the other three groups combined.

The results analysis of the item omission measure may be questionable because, as Table 3 shows, the scores within each cell of the design are positively skewed and because the within-cell variances are not homogeneous. That is, two assumptions underlying analysis of variance were violated in this instance. However, Table 3 also shows that larger variances are not associated with smaller cell sizes, the circumstance under which an inflated F value is obtained (Tabachnick & Fidell, 1995). Indeed, the smallest variability was obtained in the RNR—late returns condition, where 10 of the 12 respondents answered all the questions in the survey, whereas the largest variability was obtained in the TDM—late returns condi-

tion, where we expected the respondents to be most careless in their responses. That is, the heterogeneity of variance in the item omission variable appears to be caused by the hypothesized differential reactions of the individuals in the RNR as compared to the TDM conditions who eventually responded to the survey after receiving follow-up mailings (the late return conditions).[4]

A two (TDM vs. RNR) by two (early vs. late return) between-subjects design was also used to analyze the mean number of words written in response to an open-ended question, a second measure of response quality. Again, the interaction was the only significant effect that was obtained (fourth row of Table 3). Simple main effects showed that the late return participants from the RNR group provided significantly longer written responses to the open-ended question than the TDM—late return group ($F[1,162] = 3.93$, $p < .05$); and there was no significant difference between the TDM group and the RNR group ($F < 1$) for the early return participants. That is, hypothesis 2 was not supported because the interaction was due to the RNR—late return group writing more in response to the open-ended question than the RNR—early return group ($F[1,162] = 6.82$, $p < .01$), rather than the TDM—late return group writing less in response to the open-ended question than the TDM—early return group.

Nonparticipant Use of the Reasons for Not Responding Form

An important aspect of this research is the extent to which the nonconsenting participants used the RNR forms, and the information gathered from these forms. As stated earlier, 26 of the 61 nonrespondents in the alternative follow-up group returned an RNR form (18.1% of this group; 34.4% of the nonrespondents). Of those participants returning an RNR form, a large majority (88.5%, $N = 23$) returned a completed or partially completed form even though they were encouraged to return a blank form if they desired. We also found that those individuals returning RNR forms did so early in the mail-out process. Specifically, 55.6% of the RNR forms were returned prior to the first follow-up contact, and 83.3% of the RNR forms were returned prior to the second follow-up contact.

The reasons nonparticipants gave for not returning a survey are shown in Table 4. This table reveals that many nonparticipants disliked research projects (34.8%) or felt that social research invades their privacy (39.1%). In addition, a quarter (26.1%) of the nonparticipants indicated they did not have time to complete the survey.

Table 4
Reasons for Not Responding Given on the RNR Form

Reason for Nonresponse	*N*	%
The research invades my privacy	9	39.1
Don't like research projects	9	34.8
No time	6	26.1
Not interested	3	13.0
Not in the mood	3	13.0
Worried about confidentiality	2	7.7
Other	6	26.1

Note. **Twenty-three people gave their reasons for not responding to the survey. The percentage column does not add to 100 because nonrespondents could give more than one reason for their refusal to participate.**

DISCUSSION

The results of this research indicate that the alternative follow-up method has three advantages over the traditional method. First, the alternative method appears more ethical. Specifically, in support of hypothesis 1, it was found that the alternative follow-up group was more satisfied than the traditional follow-up group with the research topic and with the manner in which the research was conducted, at both the early and late return stages. Indeed, the fact that the alternative follow-up group reported greater satisfaction at the early return stage suggests that merely providing participants the opportunity to express a desire not to participate increases satisfaction with the research. To elaborate, at the early return stage both groups were treated in exactly the same way, except that the initial mail-out packages sent to the alternative follow-up group included the RNR forms and reference to the RNR forms in the cover letters. Therefore, the difference in satisfaction between the two methods at this stage is likely to be a result of the opportunity to decline participation.

Greater satisfaction with the way the research was conducted also supports Sudman's (1985) claim that offering people the chance not to respond to a mail survey creates trust in and empathy with the researcher. That is, if participants perceive that the researcher is making an effort to decrease any unnecessary disturbance or pressure to participate, they may feel that he or she has their best interest in mind.

More importantly, hypothesis 1 was supported by the finding that the traditional follow-up group showed a significant decline in satisfaction with the research and those conducting the research from the early to late return stages, whereas the alternative follow-up group maintained a constant level of satisfaction at these two stages. The only difference between the groups at the late return stage was that in the alternative group follow-up mailings were not sent to participants who had returned their RNR forms, and the RNR forms were included in the third and fourth mailings. Therefore, by having nonparticipants identify themselves, the RNR follow-up method appears to have reduced the pressure associated with later follow-up contacts.

The ethical importance of identifying nonrespondents is enhanced by the fact that the majority of nonrespondents returning RNR forms claimed that they had chosen not to respond to the survey because they did not wish to be involved in a social research project, and not merely due to ambivalence or lack of interest. To be more precise, 39% of the nonrespondents returning RNR forms felt that survey research is an invasion of privacy, and 34.8% reported that they did not like doing research projects. It can be inferred that if these participants had not been removed from the follow-up mailing list (as would have been the case with traditional follow-up methods), they would have been quite upset at receiving one or two more mailings requesting participation, potentially adding to their distrust and dislike for social research. In other words, these findings suggest that tradition follow-up methods used when conducting a mail survey may exacerbate and validate the negative beliefs about and feelings toward social research held by a subset of nonresponders. Further, given the results for the RNR group, this serious ethical problem is easily avoidable.

The second advantage of the alternative follow-up method is that it seems to improve response quality over the traditional follow-up method. Specifically, the traditional follow-up group showed a significant increase in the number of items left unanswered at the late return stage of follow-up mail contacts compared to the alternative follow-up group, which showed a consistently low number of item omissions at both the early and late follow-up contact stage. This finding is important because it suggests that the use of the RNR form not only answers ethical concerns regarding the negative impact of repeated contacts on participants, but also improves the quality of the data collected.

Response quality was also measured by length of response to an open-ended question. The assumption was that participants who were providing a better quality of response would provide a lengthier response. Surprisingly, given the findings for the item omission measure, the results showed that the traditional group did not demonstrate the predicted decline in the number of words written from the early to the late return stages in the traditional group. Further, the alternative follow-up group showed an increase in the number of words written in response to the open-ended question compared to the traditional follow-up group.

We offer two speculative reasons to explain these unexpected findings, although there are undoubtedly others. Perhaps most straightforwardly, the number of words used to answer an open-ended question is a measure of response quantity, not response quality. Alternatively, assuming that this criterion does measure response quality, these findings can be explained by Sudman's (1985) observation, mentioned earlier, that providing participants in mail research the opportunity to decline participation increases trust in and empathy with the researcher. Remember that participants in the RNR—late return group were encouraged three or four times to inform the researcher of any desire not to participate. Perhaps, then, this repeated show of respect increased these participants' trust in the researchers. This increased trust may then have been expressed as an increased willingness to provide detailed personal information in answer to the open-ended question. That is, the results from this measure may indicate that the alternative follow-up procedure increased response quality rather than that the traditional follow-up procedure decreased response quality.

An important finding in this study was that the alternative follow-up group had a final response rate similar to the traditional group. As Sudman stated, "the obvious danger of [encouraging nonconsenters to identify themselves] is that it might encourage persons who would otherwise cooperate to refuse" (1985, p. 354). However, this did not happen. Indeed, not only did the RNR method maintain the same response rate as the TDM method, it also generated a substantial amount of information that would otherwise not have been obtained. Specifically, 34.4% of the nonrespondents in the alternative follow-up group returned an RNR form, and of those returned 88.5% were either entirely or partially completed. Therefore, a considerable amount of information was gathered on why nonconsenters chose not to participate in the survey and on their

demographic characteristics. The latter information can be extremely useful to researchers when considering the limits to the generalizability of their findings. For example, in this research the male nonrespondents were generally older than the male respondents. In addition, many nonrespondents were willing to state why they did not wish to respond and to criticise the research. Thus, the RNR method can determine the extent to which nonresponse was a result of methodological or procedural problems, a social climate resistant to research, or a target population's disliking a particular research topic.

Although this study appears to suggest that for little extra effort researchers and participants may benefit from following the RNR follow-up procedure, there are areas where this issue should be examined further. For example, a limitation of this research is that what has been described as a "traditional follow-up" method is only a loose description of the procedures followed by survey researchers. In reality, follow-up procedures vary according to each researcher's beliefs about the best way to maximize response rates from a specific target population as well as for other, more practical reasons such as cost. For example, in this research certified mail was not used in the final mailing even though this is part of the TDM procedure (Dillman, 1978). Researchers also often vary the degree of the response-inducement techniques, such as the amount of persuasion used, the timing of mail-outs, the salience of the survey used, or the length of survey. Therefore, further work is necessary to examine whether the RNR follow-up method is advantageous when compared with a variety of other follow-up procedures.

Initially all the dependent variables were examined for skewness and kurtosis in order to ensure that they were normally distributed within each cell of the design. Then the homogeneity of variance assumption underlying analysis of variance was carefully examined, a procedure that was particularly important because the necessarily smaller sample sizes in the late return conditions may result in serious violation of this assumption. In general, these preliminary analyses indicated that these potential problems did not materialize. In particular, responses to the satisfaction measure are distributed normally and have very similar variability in all conditions of the design. However, as Table 3 shows, the scores on the item omission measure, the main measure of response quality, were positively skewed and the within-cell variance differed considerably. Nevertheless, these violations did not result in conditions in which the

fewest respondents have greater variability, the circumstance under which an inflated *F* value occurs. Therefore, we feel confident that the significant interaction that we obtained for this measure was not artifactual, but represented a real effect: individuals who responded to the survey after repeated contacts answered the survey more completely in the RNR condition than in the TDM condition. Nevertheless, a replication of this research with a much larger sample would be of value in order to demonstrate the robustness of these results.

Replicating this study with a larger sample size would be valuable for two further reasons. First, more individuals would respond to the fourth follow-up mailing, thereby increasing the power to detect a difference at this stage. This increased power would allow for more specific and detailed information on how several follow-up contacts influence potential respondents. Second, replicating the findings of this research with a larger and more varied random sample is important because, unexpectedly, a disproportionate number of responses to this survey came from women and well-educated individuals (Table 1). Therefore, the extent to which these findings can be generalized is unclear.

To conclude, despite the limitations of this research and the need for further investigation, the results of this study suggest that many advantages can be gained by using the RNR follow-up method in mail survey research. Simply providing participants in mail surveys with the opportunity to decline participation provides both ethical and methodological benefits with no negative consequences for response rates. Further, the RNR method furnishes important information on the reasons for nonresponse and on the demographic characteristics of people who choose not to respond to the survey.

ACKNOWLEDGEMENT

The research described in this paper was P.D. Sully's master's thesis in the Applied Social Psychology graduate program under P.R. Grant's supervision. Thanks are extended to Louise Alexitch, David Hay, Carl von Baeyer, and Louis Webb for their constructive comments.

NOTES

1 This survey examined emotional empathy and cognitive perspective-taking in relation to varying levels of "social psychopathy."

2 This is the usual practice followed by faculty and graduate students in the Applied Social Psychology graduate program at the University of Saskatchewan when conducting needs assessments and consumer satisfaction surveys in the local community. It is their experience that the response rate to the third follow-up is not adversely affected, and the savings are significant.

3 Three participants from the traditional follow-up group returned blank surveys and were classified as nonrespondents.

4 An analysis of variance on a logarithmic transformation of the item omission measure yielded very similar results.

REFERENCES

Brehm, J.W. (1966). *A theory of psychological reactance*. New York: Academic Press.

Canadian Psychological Association. (1991). *A Canadian code of ethics for psychologists: Revised*. Ottawa: CPA Publications.

Childers, T., & Skinner, S. (1985). Theoretical and empirical issues in the identification of survey respondents. *Journal of the Market Research Society, 27*, 39–53.

Dillman, D. (1978). *Mail and telephone surveys: The Total Design method*. New York: John Wiley & Sons.

Dillman, D. (1991). The design and administration of mail surveys. *Annual Review of Sociology, 17*, 225–249.

Doob, A.N., & Zabrack, M. (1971). The effect of freedom-threatening instructions and monetary inducement on compliance. *Canadian Journal of Psychology, 3*, 408–412.

Eckland, B.K. (1965). The effects of prodding to increase mail-back returns. *Journal of Applied Psychology, 49*, 165–169.

Eichner, K., & Habermehl, W. (1981). Predicting response rates to mailed questionnaires. *American Sociological Review, 46*, 361–363.

Faria, A.J., Dickenson, J.R., & Filipic, T.V. (1990). The effect of telephone versus letter prenotification on mail survey response rate, speed, quality, and cost. *Journal of Market Research Society, 32*, 551–568.

Goyder, J. (1982). Further evidence on factors affecting response rates to mailed questionnaires. *American Sociological Review, 47*, 550–553.

Hawkins, D. (1979). The impact of sponsor identification and direct disclosure of respondents' rights on the quantity and quality of mail survey data. *Journal of Business, 52*, 577–590.

Heberlein, T.A., & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review, 43*, 447–462.

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.

Kanuck, L., & Berenson, C. (1975). Mail survey response rates: A literature review. *Journal of Marketing Research, 12*, 440–453.

Linsky, A.S. (1975). Stimulating responses to mailed questionnaires: A review. *Public Opinion Quarterly, 38*, 82–101.

Lockhart, D.C. (1984). The stages of mailed questionnaire returning behaviour. In D. Lockhart (Ed.), *Making effective use of mailed questionnaires* (pp. 89–99). San Francisco: Jossey-Bass.

Nevin, J.R., & Ford, N.M. (1976). Effects of a deadline and a veiled threat on mail survey responses. *Journal of Applied Psychology, 61*, 116–118.

Ray, J.J., & Still, L.V. (1987). Maximizing the response rate in surveys may be a mistake. *Personality and Individual Differences, 8*, 571–573.

Stone, E.F., Stone, D.L., & Gueutal, H.G. (1990). Influence of cognitive ability on responses to questionnaire measures: Measurement precision and missing response problems. *Journal of Applied Psychology, 75*, 418–427.

Sudman, S. (1985). Mail survey of reluctant professionals. *Evaluation Review, 9*, 349–360.

Tabachnick, B.G., & Fidell, L. S. (1995). *Using multivariate statistics*. New York: HarperCollins.

Wellman, J., Hawk, E.G., Roggenbuck, J.W., & Buhyoff, G.J. (1980). Mailed questionnaire surveys and the reluctant respondent: An empirical examination of differences between early and late respondents. *Journal of Leisure Research, 12*, 164–172.

Yu, J., & Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Market Research, 20*, 36–34.

Appendix A
Participant Satisfaction Scale

1. In your opinion, how important is the research topic addressed by this survey?
   not at all important 1 2 3 4 5 6 7 very important
2. To what extent do you feel that this survey was conducted in a professional manner?
   not at all professional 1 2 3 4 5 6 7 very professional
3. How motivated were you to take part in this research?
   not at all motivated 1 2 3 4 5 6 7 very motivated
4. In the future, if you were asked to take part in a similar type of mail survey study, how willing would you be to participate?
   not at all willing 1 2 3 4 5 6 7 very willing
5. How satisfied were you with the way in which this research was explained to you?
   not at all satisfied 1 2 3 4 5 6 7 very satisfied
6. How satisfied were you with the way in which this mail survey was conducted?
   not at all satisfied 1 2 3 4 5 6 7 very satisfied