

The Effects of Response-Stimulating Factors on Response Rates and Data Quality in Mail Surveys

A Test of Dillman's Total Design Method

Edith D. De Leeuw¹ and Joop J. Hox²

Abstract: Dillman's Total Design Method (TDM) for mail surveys has proved effectual in attaining high response rates in the U.S.A. The TDM, however, has two drawbacks: its relatively high costs of a mail survey and its susceptibility to response errors.

We tested the ability of the TDM techniques to increase response rates. Furthermore, the influence of the TDM on the response error was investigated.

In this experiment, the effects of personalization of the cover letter and a final reminder by certified mail were tested on a sample of the Dutch population. If the TDM is followed completely the response rate in the Netherlands is comparable to the response rate in the U.S.A (70%). Omission of personalizing

the cover letter or sending a reminder by certified mail results in a significantly lower response rate. Using the TDM in toto does not lead to responses of inferior quality. We find that a reminder by certified mail in combination with a personalized cover letter does not result in more socially desirable answers, more item non-response, or a more deviating response pattern. However, the use of only a personalized cover letter, i.e., without a certified mailing, did result in more socially desirable answers.

Key words: Mail survey; Total Design Method, TDM; personalization of cover letter; reminder by certified mail; non-response error; response error.

¹ Department of Research Methods, Free University of Amsterdam, Koningslaan 22-24, 1075 AD Amsterdam, The Netherlands. EARN: V70UEDIT at HASARA11.

² Department of Education, University of Amsterdam, Prinsengracht 227, 1015 DT Amsterdam, The Netherlands. EARN: A716HOX at HASARA11.

This article was written while E.D. De Leeuw was a Visiting Fulbright Scholar at the Social and Economic Sciences Research Center, Washington State University. She is indebted to the Nuts Seminarium of the University of Amsterdam, which kindly gave permission for this sabbatical. Both authors wish to thank A.W.M. Duijx for his assistance, and D.A. Dillman for his comments.

1. Introduction

Using mail surveys, response rates of 70% for the general population are attainable (Dillman (1972, 1978), Hippler and Seidel (1985)). The potentially high response rate in mail surveys together with their low costs make mail surveys an attractive and popular survey method (Houston and Ford (1976), Kanuk and Berenson (1975)). Mail surveys have no interviewer bias (Hippler (1988)) and are geographically flexible, i.e., they reach a widely dispersed or inaccessible sample (Kanuk and Berenson (1975)).

The feasibility of high response rates for mail surveys is the result of much methodological research over the past three decades. A large number of empirical studies on response-increasing techniques for mail surveys has been summarized in several comprehensive reviews (Baumgartner and Heberlein (1984), Goyder (1982), Heberlein and Baumgartner (1978), Kanuk and Berenson (1975), Linsky (1975), Yu and Cooper (1983)). Of the four factors found to be most effective in increasing response rates – saliency, sponsorship, follow-up contacts, and incentives – only the last two can be readily controlled by an individual survey researcher.

Further research has indicated that a special kind of reminder, for instance by certified mail, would be more effective than an ordinary follow-up (Heberlein and Baumgartner (1978), House, Gerber, and McMichael (1977)). A possible explanation is that a follow-up sent by certified mail underscores the importance of the study and that respondents feel rewarded by the attention they receive (cf. Dillman and Moore (1983)).

Personalization is a fifth factor that also influences response rates but it is difficult to interpret the influence of personalization. Personalization, however, seems to give mixed results. For instance, an increase in response rates is reported by Carpenter (1974) and by Dillman and Frey (1974); a decrease

in response rates is reported by Andreasen (1970) and by Houston and Jefferson (1975). It appears that personalization interacts with the other characteristics of the survey. Personal attention can be rewarding for the respondent, but, on the other hand, it can threaten the respondent's anonymity (Houston and Jefferson (1975), Wiseman (1976)). To understand these effects fully, it is necessary to develop theories of mail questionnaire response (Baumgartner and Heberlein (1984)).

A theoretical framework for mail surveys is given by Dillman (1978) who, guided by social exchange theory, has integrated several well-known procedures for increasing response rates. This set of procedures is called the Total Design Method (TDM). In the TDM, the completion and return of a mail questionnaire is viewed as a social exchange between the researcher and the respondent. According to social exchange theory (Homans (1973)), a potential respondent will complete and return a questionnaire if the (material and psychological) costs of responding are low in comparison with the benefits. Consequently, the researcher needs to minimize the expected costs and maximize the expected benefits of participation.

Important features of Dillman's TDM are: a personalized cover letter, a simple and attractive questionnaire, and a follow-up mailing. One week after the initial mailing, the entire sample (respondents and non-respondents) receives a postcard serving as a "thank you" or reminder. Three weeks after the initial mailing all non-respondents receive a new questionnaire and a cover letter. Seven weeks after the initial mailing this procedure is repeated, but this time the questionnaire is sent by certified mail. According to Dillman (1978, pp. 7–27), none of these procedures should be omitted. The TDM should be seen as a package of procedures which interact to produce a maximum response rate. *

The TDM has the advantage of achieving a high response rate, but this is done at considerable cost (Kulka, Shirey, Moore, and Woodbury (1981)). The cost can be reduced by not using the personalized cover letter or by not sending the reminder by certified mail. However, the omission of a reminder by certified mail can considerably lower the response rate (Heberlein and Baumgartner (1978)).

The quality of the data is also an important factor for consideration. Response errors have received comparatively little attention in methodological studies on mail surveys (Houston and Ford (1976)). This is unfortunate as it is conceivable that the same procedures (e.g., personalization, use of certified mail) that increase the response rate decrease the data quality (Jones and Lang (1982), Nevin and Ford (1976), Sigelman (1982)).

Three well-known sources of error are (1) item non-response (Bradburn, Sudman, Blair, and Stocking (1978), Houston and Ford (1976)), (2) psychometrically deviant response patterns or non-scalability of responses (Turner and Fiske (1968), Tatsuoka and Tatsuoka (1983)), and (3) social desirability effects (Bradburn et al. (1978), Sudman and Bradburn (1974)). Attempts to increase the response rate by personalizing cover letters and sending reminders by certified mail will increase the pressure on the respondent (De Leeuw and Hox (1987)). Respondents can react by completing the questionnaire as fast as possible without much thought. This may produce erratic individual responses which will show up as item non-response and psychometrically deviant response patterns. Furthermore, personalization of the cover letter can act as a threat to the respondent's anonymity and privacy. Respondents may react to this threat by giving more socially desirable answers.

This study is a controlled field experiment

to test the efficacy of Dillman's Total Design Method in the Netherlands. In this experiment, the effects of the two most expensive factors of the TDM, personalization of cover letter and a final reminder by certified mail, were tested on a sample of the Dutch population. The effects investigated were response rate, item non-response, non-scalability of responses, and social desirability.

2. Implementation of the Experiment

The sample consisted of 500 persons aged 18 and older. Everyone in the sample was sent a mail questionnaire on his/her education and upbringing. With the exception of the experimental variations, all of Dillman's recommendations concerning the questionnaire, the cover letters, and the implementation of the survey were followed closely. All 500 potential respondents received a cover letter, a questionnaire, and a postage-paid reply envelope. One week later, all received a "thank you" reminder postcard. Three weeks after the initial mailing, those who had not yet responded received a differently worded cover letter, another questionnaire, and a postage-paid reply envelope. Two experimental factors were manipulated: personalization of the cover letter and a third reminder by certified mail sent seven weeks after the initial mailing.

The two experimental factors were implemented in the following way.

2.1. Personalization of cover letter

Cover letters explaining the survey and asking for cooperation were printed on letterhead stationary. Each letter was individually signed with a blue ballpoint pen. In the personalized treatment, the name and the full address were typed on each letter, using the same type style as the rest of the letter. In the non-personalized treatment, the name and address were omitted.

2.2. Final reminder by certified mail

Seven weeks after the initial mailing, a third and final reminder was sent. This reminder consisted of a new, differently worded cover letter and a copy of the questionnaire. In the certified treatment, these reminders were sent by certified mail; in the not certified treatment, they were sent by first class mail.

These experimental treatments were crossed in a randomized 2×2 design, resulting in four experimental groups. In the treatment “personalized and certified,” the TDM is fol-

lowed completely which makes it possible to compare the response rate obtained with the range of response rates Dillman reports for the TDM in the U.S.A.

A fifth experimental treatment was added to this design. In this treatment we tried to lower the costs as much as possible. As a result, the questionnaires were mailed as printed matter; the cover letters were not signed by hand, but were reproduced with a signature. Nor was the cover letter personalized (i.e., the letter did not include the name and address of the respondent).

Table 1. Number of completed questionnaires, non-response rate, ineligible and response rate per treatment after the third mailing

	Treatments				
	1	2	3	4	5
Completed	46	55	51	55	65
Non-response	48	40	45	35	26
Ineligible	6	5	4	10	9
Sent out	100	100	100	100	100
Response Rate ^(a)	49%	58%	53%	61%	71%

Treatments: 1 = least expensive; 2 = not personalized, not certified; 3 = personalized, not certified; 4 = not personalized, certified; 5 = personalized, certified (complete TDM).

(a) $p = 0.02$

3. Results

3.1. Response rate

After the first mailing, 141 questionnaires of the 500 sent out were completed and returned. After the second mailing, a total of 251 questionnaires had been received. After the first and second mailings there were no statis-

tically significant differences between the experimental treatments. After the third mailing, differences in response rates between the treatments did occur. The response rates after the third mailing are summarized in Table 1. The total number of questionnaires received was 272.

The independent variable "mailing treatment" is a nominal variable with five categories. These categories can be ordered according to costs (the least expensive (1) to the most expensive (5)) as follows.

Treatment 1 was not personalized, had no hand written signature in blue ink, and was sent by third class mail (printed matter).

Treatment 2 was not personalized, and was not sent by certified mail.

Treatment 3 was personalized, but was not sent by certified mail.

Treatment 4 was not personalized, but was sent by certified mail.

Treatment 5 was both personalized and sent by certified mail, i.e., the complete TDM.

This ordinal scale can be modeled by coding the independent variable as four dummy variables using reverse Helmert coding (Cohen (1968), Kenny (1979)). These dummy variables were used in a logistic regression to predict the response rate (McCullagh and Nelder (1983)). This procedure makes it possible to compare different mailing treatments and to test whether a certain treatment increases the response rate significantly. Following Kviz (1977), the response rate was defined as the number of completed questionnaires calculated as a percentage of the total number mailed minus the number of ineligible. A strict definition of ineligibility was used. A respondent was ineligible if the post office returned the questionnaire because the address was unknown, the addressee had moved, or the addressee was deceased.

The logistic regression shows only one regression parameter which differs significantly from zero. This parameter represents

the dummy variable contrasting the combination of personalized cover letter and certified mail, i.e., the complete TDM procedure. The other treatments had a p -value less than 0.01. For the complete TDM treatment the response rate is 71%. This response rate is comparable with the figures reported by Dillman for the U.S.A. where the average response rate for 48 TDM surveys was 74 % and none of the 48 surveys obtained a response rate under 50 %.

The other regression parameters were not significant. The next smallest p -value was 0.09 for the regression weight contrasting the least expensive condition (cover letter not personalized, no signature in blue ink, all mailings by third class mail) with the next less expensive (cover letter individually signed with a blue ballpoint pen, but not personalized, all mailings by first class mail).

3.2. Data quality

We analyzed the influence of the response-increasing factors of the TDM (i.e., personalization of cover letter and reminder by certified mail) on item non-response, on scalability of individual response patterns, and on social desirability.

The proportion of item non-response for the twelve items that measure intellectual stimulation in youth was computed (Wolf (1965), Zaal and Van der Flier (1980)). This did not differ significantly between experimental treatments (see also Table 2).

Non-scalability of individual response patterns was defined as a large inter-item variance for the items on the intellectual stimulation scale. The inter-item variance, i.e., the variance between items within respondents, was computed omitting respondents with a missing value on one or more of these items. Again, there were no statistically significant differences between experimental treatments (see Table 2). Apparently, the response-stimulating factors used in the

Table 2. Threats to data quality: item non-response and inter-item variance per experimental treatment (not significant)

Experimental treatment	Mean proportion item non-response ^(a)	Mean inter-item variance ^(b)
1	.004	1.23
2	.002	1.42
3	.021	1.36
4	.017	1.30
5	.021	1.39

Analysis of variance: (a) $p=0.34$, (b) $p=0.37$

TDM do not threaten the data quality as indicated by the obtained item non-response and inter-item variance. Our hypothesis – that respondents react to the increased pressure of the TDM by completing the questionnaire as fast as possible without much thought – was not confirmed. The increased pressure to complete the questionnaire did not produce erratic individual responses (i.e., more item non-response and a higher inter-item variance).

To assess the social desirability bias, 25 persons rated the intellectual stimulation items according to social desirability. An item was said to be socially desirable if the response to that item was highly influenced by societal expectations (cf. De Maio (1984)). The assessors were not chosen at random, rather, they comprised a heterogeneous group with regard to age, sex, and education. The agreement among assessors was substantial, as shown by the high consistency index for the combined ratings (Cronbach's alpha, an index of psychometric consistency, was 0.78).

The effect of the experimental treatments on social desirability was analyzed by multiple regression, again using reverse Helmert coding to contrast the successive treatments. There was a significant effect of personalization on social desirability. As expected, respondents who received personalized

cover letters gave more socially desirable answers than respondents who did not receive personalized cover letters ($\beta = 0.23$, $p<0.01$). Sending the reminder by certified mail did not influence the number of socially desirable answers ($p=0.45$). Surprisingly, the combination of personalized cover letter and final reminder by certified mail did not influence the social desirability of the answers ($p=0.65$). Probably, personalization in combination with certified mail is perceived as a business-like approach which does not threaten the privacy of the respondent (cf. Kerin and Peterson (1977)).

4. Summary and Discussion

In our study, Dillman's Total Design Method (TDM) resulted in a non-response rate that is comparable with the average TDM results obtained in the U.S.A. Only the combination of personalized cover letter and final reminder by certified mail, in which the TDM is followed completely, has an effect which is statistically significant. Omission of both the personalized cover letter and the certified mail has a detrimental effect on the number of completed questionnaires. Our finding that a combination of treatments is needed to achieve maximal effect supports Dillman's claim that the procedures of the TDM reinforce each other.

There were 100 persons in each experimental treatment, giving a total sample size of 500 persons. The sample size has consequences for the power of the statistical tests used. In particular, the probability of detecting small effects for response rates is not very high (Cohen (1969)). Some of the effects in Table 1 are large enough to be of interest, even though they are not significant at the conventional 5% level. The treatment "not personalized and not certified" and the treatment "not personalized and certified" seem to raise the response rate compared with the other treatments. However, in the Netherlands, as in most other Western countries, sending a questionnaire by certified mail is expensive. If a high response rate is important enough to warrant this expense, it will certainly warrant the comparatively small additional expense of personalizing the cover letter. Consequently, the treatment "not personalized but certified" is not recommended. On the other hand, the main cost difference between the least expensive treatment and the next is the difference between the costs of third and first class mail. In most cases, the potential gain in response rate will justify this small additional cost.

No treatment effects were found on item non-response and on deviant individual response patterns. However, it was found that the personalization of cover letters did increase the number of socially desirable responses, while a reminder by certified mail did not. Furthermore, the TDM-combination of these two factors did not influence social desirability. Linsky (1975), while discussing response rates, suggests that personalization and an anonymity guarantee – which are both part of the TDM – do not go together well and may lead to an undesirable result. Labrecque (1978) states that personalization is detrimental when it is perceived as artificial. The treatment "personalization only" can be seen by the respondent as arti-

cial and make the anonymity guarantee look insincere. On the other hand, personalization combined with certified mail, as used in the TDM, may look very natural indeed.

In our study, the treatments personalization and certified mail produced good results on both response rates and data quality. However, other interactions may well have had detrimental effects (cf. Linsky (1975), Labrecque (1978)). It is recommended that experiments are designed which permit a systematic investigation of interaction effects (cf. Heberlein and Baumgartner (1978, p. 460)).

Reviews of the literature on mail survey response (such as Heberlein and Baumgartner (1978), Goyder (1982), Baumgartner and Heberlein (1984)) suggest that the present TDM does not use all potentially effective treatments for increasing response rates. Further research on the optimalization of the TDM is needed. For instance, how should incentives be incorporated into the TDM? Is there a better or less expensive alternative for the special final reminder by certified mail?

Nevertheless, when followed completely, the TDM is a highly successful method for conducting mail surveys. It seems to reduce the non-response rate without increasing the response error.

5. References

- Andreasen, A.R. (1970): Personalizing Questionnaire Correspondence. *Public Opinion Quarterly*, 34, pp. 237–277.
- Baumgartner, R.M. and Heberlein, T.A. (1984): Recent Research on Mailed Questionnaire Response Rates. In: *Making Effective Use of Mailed Questionnaires*, edited by D.C. Lockhart. Jossey-Bass, San Francisco.

- Bradburn, N.M., Sudman, S., Blair, E., and Stocking, C. (1978): Question Threat and Response Bias. *Public Opinion Quarterly*, 42, pp. 221–234.
- Carpenter, E.H. (1974): Personalizing Mail Surveys: A Replication and Reassessment. *Public Opinion Quarterly*, 39, pp. 614–620.
- Cohen, J. (1968): Multiple Regression as a General Data Analytic System. *Psychological Bulletin*, 70, pp. 426–443.
- Cohen, J. (1969): *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, New York.
- De Leeuw, E.D. and Hox, J.J. (1988): Artifacts in Mail Surveys; The Influence of Dillman's Total Design Method on the Quality of the Responses. In: *Sociometric Research*, Vol. II, edited by W.E. Saris and I.N. Gallhofer. MacMillan, London.
- De Maio, T.J. (1984): Social Desirability and Survey Measurement: A Review. In: *Surveying Subjective Phenomena*, Vol. II, edited by C.F. Turner and E. Martin. Wiley & Sons, New York.
- Dillman, D.A. (1972): Increasing mail questionnaire response in large samples of the general population. *Public Opinion Quarterly*, 36, pp. 254–257.
- Dillman, D.A. (1978): *Mail and Telephone Surveys: The Total Design Method*. Wiley & Sons, New York.
- Dillman, D.A. and Frey, J.H. (1974): Contribution of Personalization to Mail Questionnaire Response as an Element of a Previously Tested Method. *Journal of Applied Psychology*, 59, pp. 297–301.
- Dillman, D.A. and Moore, D.E. (1983): Improving Response to Mail Surveys: Results From Five Experiments. Unpublished manuscript, Washington State University, Pullman, Washington.
- Goyder, J.C. (1982): Further Evidence on Factors Affecting Response Rates to Mailed Questionnaires. *American Sociological Review*, 47, pp. 550–553.
- Heberlein, T.A. and Baumgartner, R.M. (1978): Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature. *American Sociological Review*, 43, pp. 447–462.
- Hippler, H.J. (1988): *Methodische Aspekte schriftlicher Befragungen: Probleme und Forschungsperspektiven*. (Methodological Aspects of Mail Surveys: Problems and Research Perspectives.) *Planung und Analyse*, 6, pp. 244–248.
- Hippler, H.-J. and Seidel, K. (1985): *Schriftliche Befragung bei allgemeinen Bevölkerungsstichproben: Untersuchungen zur Dillman'schen Total Design Method (Mail Surveys of the General Population: A Test of Dillman's Total Design Method)*. ZUMA-nachrichten, Zentrum für Umfragen und Analysen, Mannheim.
- Homans, C.G. (1973): *Social Behavior*. Routledge and Kegan Paul, London.
- House, J.S., Gerber, W., and McMichael, A.J. (1977): Increasing Mail Questionnaire Response: A Controlled Replication and Extension. *Public Opinion Quarterly*, 41, pp. 95–99.
- Houston, M.J. and Ford, N.M. (1976): Broadening the Scope of Methodological Research on Mail Surveys. *Journal of Marketing Research*, 13, pp. 397–403.
- Houston, M.J. and Jefferson, R.W. (1975): The Negative Effects of Personalization on Response Patterns in Mail Surveys. *Journal of Marketing Research*, 12, pp. 114–117.
- Jones, W.H. and Lang, J.R. (1982): Reliability and Validity Effects Under Mail Survey Conditions. *Journal of Business Research*, 10, pp. 339–353.
- Kanuk, L. and Berenson, C. (1975): Mail Surveys and Response Rates: A Literature Review. *Journal of Marketing Research*, 12, pp. 440–453.

- Kenny, D.A. (1979): *Correlation and Causality*. Wiley & Sons, New York.
- Kerin, R.A. and Peterson, R.A. (1977): Personalization, Respondent Anonymity and Response Distortion in Mail Surveys. *Journal of Applied Psychology*, 62, pp. 86–89.
- Kulka, R.A., Shirey, S.J., Moore, R.P., and Woodbury, N. (1981): A Factorial Experiment on the Responses of Professional Nurses to a National Mail Survey. *American Statistical Association, Proceedings of the Section on Survey Research Methods*.
- Kviz, F.J. (1977): Towards a Standard Definition of Response Rate. *Public Opinion Quarterly*, 41, pp. 265–267.
- Labrecque, D.P. (1978): A Response Rate Experiment Using Mail Questionnaires. *Journal of Marketing*, pp. 82–83.
- Linsky, A.S. (1975): Stimulating Responses to Mailed Questionnaires: A Review. *Public Opinion Quarterly*, 39, pp. 82–101.
- McCullagh, P. and Nelder, J.A. (1983): *Generalized Linear Models*. Chapman and Hall, London.
- Nevin, J.R. and Ford, N.M. (1976): Effects of a Deadline and a Veiled Threat on Mail Survey Responses. *Journal of Applied Psychology*, 61, pp. 116–118.
- Sigelman, L. (1982): The Uncooperative Interviewee. *Quality and Quantity*, 16, pp. 345–353.
- Sudman, S. and Bradburn, N.M. (1974): *Response Effects in Surveys: Review and Synthesis*. Aldine, Chicago.
- Tatsuoka, K.K. and Tatsuoka, M.M. (1983): Spotting Erroneous Rules of Operation by the Individual Consistency Index. *Journal of Educational and Psychological Measurement*, 20, pp. 221–230.
- Turner, C.B. and Fiske, D.W. (1968): Item Quality and Appropriateness of Response Processes. *Educational and Psychological Measurement*, 28, pp. 297–315.
- Wiseman, F. (1976): A Reassessment of the Effects of Personalization on Response Patterns in Mail Surveys. *Journal of Marketing Research*, 13, pp. 110–111.
- Wolf, R. (1965): The Measurement of Environments. *Proceedings of the 1964 Invitational Conference on Testing Problems*. Educational Testing Service, Princeton, N.J.
- Yu, J. and Cooper, H. (1983): A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. *Journal of Marketing Research*, 20, pp. 36–44.
- Zaal, J.N. and Van der Flier, H. (1980): *Sociaal Culturele Verschillen in de Gezinsomgeving (Social and Cultural Differences in the Home Environment)*. *Gedrag*, 4, pp. 197–217.

Received February 1988

Revised October 1988