

The Impact of Incentives and "Multi-Media Approaches" on the Levels of Response Rates and the Quality of Response

Julian Baim Senior Vice President Mediamark Research Inc.

Abstract

This study tested two approaches for improving response rates in MRI's readership survey. Using a predictive model for identifying geographic areas with historically low response rates, MRI compared the impact of higher monetary incentives and "multi-media approaches" with our current interviewing procedures on response rates. We hypothesized that additional monetary incentives and/or intensive efforts to contact the respondent would lead to significantly higher response rates.

The study results did not support the hypothesis. We found that:

- (1) A \$10.00 incentive failed to achieve a higher response rate than the traditional incentive offered by MRI. This test was conducted among 22 pairs of matched interviewing clusters.
- (2) A \$10.00 incentive along with a "multi-media approach" produced significantly lower response rates than the traditional incentive offered by MRI. This test was conducted among 18 pairs of matched interviewing clusters.

Finally, we compared the screen and read levels of the control and test groups. We hypothesized that different incentives and treatments would produce different readership responses. This hypothesis, too, went unsupported.

The decline in survey response rates is undeniable. More and more Americans refuse to cooperate with interviewers for surveys ranging from measuring media to gauging public opinion. Whether this phenomenon has been propelled by the proliferation of telemarketing, the increased desire for privacy or by a myriad of other factors, this simple fact remains: non-participation threatens the accuracy and reliability of survey projections.

The degree to which this issue dominates the research agenda is reflected in the spate of recent articles on the subject. In the two most recent issues of *Public Opinion Quarterly*, five articles, including a reprint of AAPOR's annual Presidential address, are devoted to the subject. Norma Bradburn (1992), President of AAPOR, maintained "We all believe strongly that response rates are declining and have been for some time. Part of the problem is locating respondents, and part of the problem is getting co-operation. Locating respondents is more difficult than it used to be because of life-style changes, especially among women. People are just not at home as much as they used to be. They also appear to be harder to convince that they should participate in a survey."

To understand and even counter the "non-response problem," researchers at MRI worked with Dr Valentine Appel to develop a predictive model for anticipating the likelihood of response rate problems. Since MRI knew in advance the geodemographic descriptors of its upcoming interview clusters, such a model might prove useful in taking corrective measures to increase response rates in "difficult to interview" areas.

That model, using several years of MRI data, was presented to the research and advertising community at the 37th annual ARF conference. Our analysis suggested that "81.5% of the variance in response rates among PRIZM clusters is accounted for by three demographic variables: (1) the extent of urbanization (2) median home value and (3) the proportion of householders who are retired and relatively well off." All three variables were negatively correlated with response rates. Other variables such as household income or educational attainment were important in explaining response rate levels, but their impact was highly correlated with median home value.

Beyond the theoretical value of demonstrating the model's predictive capability, MRI wanted to test whether remedial measures could be taken in clusters where our model predicted high levels of non-response. Thus, MRI conducted two experiments to test alternative measures for increasing the respondent cooperation level.

The Test Designs:

There are a number of possible options for trying to improve response rates, ranging from invoking some authoritative source as the study's sponsor to matching the interviewer's demographic characteristics to the profile of the subsample population. Although a number of alternatives were considered, MRI examined the following two factors: (1) offering of an increased incentive (2) employing a "multi-media approach" for contacting respondents.

This decision was based on feedback from our interviewing staff as well as on the geodemographic characteristics of the target interview clusters. We knew that the people in these clusters were upscale with many living in exclusive, relatively inaccessible areas or residences. We believed that the offered incentive should be commensurate with the survey population's profile; hence, a larger incentive should be proffered to a more affluent population. In addition, because these respondents were more difficult to reach, we believed that diversifying our attempts to alert and contact the households would lead to a heightened awareness of the study and an improved response rate. We then tested the impact of a monetary incentive alone and the effect of an increased incentive in conjunction with a "multimedia approach."

The first test involved comparing response rates for 22 pairs of matched interviewing clusters where the test group received an upfront \$10.00 incentive while the control group received our traditional key chain/Swiss army knife "door opener." The selected PRIZM clusters, based on the predictive model, were Furs and Station Wagons, Pools and Patios and Two More Rungs. (These descriptive nicknames identify various "neighbourhood types" through factor analysis. The "neighbourhood types" represent geographically compact areas characterized by households with similar demographic profiles). Each of these type clusters had demonstrated historically lower than average response rates and each fit the modelled profile of "difficult to interview" areas.

The test was administered under tight controls. Since the experiment involved paired clusters, MRI used the same interviewer for each pair. This procedure controlled for interview effect on the final response rate levels. The interviewers were thoroughly briefed and were later debriefed about the effect of the increased incentive. They were not informed beforehand of the purpose of this study. In addition, all interviewing attempts in the paired clusters began simultaneously to discount the effect of staggered starting times or even seasonal variations on the results. Consistent with MRI's usual procedures, six attempts were made to complete an interview in each household.

The second test compared response rates for 18 pairs of matched clusters but involved a much more elaborate effort to solicit cooperation in the test group. Prompted by Richard Lysaker's presentation at the International Symposium of Readership Research in Hong Kong (1991), MRI devised a "multi-media approach" in an effort to maximize response rates. Lysaker had found that "the value of making additional contacts through an additional medium turns out to be far from trivial. Each medium makes an important contribution to overall response levels."

Consequently MRI devised a systematic scheme for contacting pre-designated households. Each sample household having a listed phone number would initially be contacted by phone, screened for the appropriate respondent and an attempt would be made to set up an appointment to conduct the interview. Like the first test group, the second test group respondents would be offered a \$10.00 incentive to participate. Following a scheduled appointment, a personalized mail-gram would be sent to thank the respondent for their willingness to cooperate and to confirm the appointment time and date. An additional reminder phone call would be made to the respondent to confirm the appointment prior to the interview.

For households where no phone number was available, a generic mail-gram announcing MRI's presence in their area along with the offered \$10.00 incentive was sent. The mailgram also contained a toll-free number which respondents could use to set up an appointment. If the household did not respond to the mail-gram, MRI reverted to its established procedures for contacting each household.

The selected sample for this test included the Urban Gold Coast, Money and Brains, Blue Blood Estates and Bohemian Mix PRIZM clusters.

In contrast to the test group, the control group households were contacted in the usual manner. This meant a pre-alert, mailed notice informing the household of their selection, describing the importance of the study and soliciting their cooperation when our interviewer arrived unannounced at their doorstep. Once again, each respondent would receive the keychain/ Swiss army knife as our token of appreciation. The same strict controls on interviewers and timing were maintained in this study.

Comparability of Control and Test Groups

Before discussing the response rate results, we compared the demographics of the test and control groups for each of the two tests. Since it was extremely important that the paired clusters reflected respondents with similar demographic characteristics, we wanted to validate our design. We found no significant difference (two-tailed paired t-test at the .05 level) between the control and test groups for mean household income and mean individual employment income. In addition, employment levels and mean age were not significantly different between the control and test groups. As a result we were even more confident that the "treatment effect" could be isolated as the defining variable in explaining response rate differences.

Test Results: Pilot Study #1

The final response rate comparisons for the first pilot test are shown below:

TABLE #1
RESPONSE RATE COMPARISON
CONTROL VS. TEST

	CONTROL (SWISS ARMY KNIFE)	TEST (\$10.00)
ELIGIBLE	270	257
COMPLETE	175	162
RESPONSE RATE	64.8%	63.0%

As the table demonstrates, the results were unexpected. Whereas we had hypothesized that the increased incentive would improve the response rate, cooperation was 1.8% worse in the test group where the higher incentive was given. Although the difference was not statistically significant (one-tailed t-test at the .05 level), the data do not support our hypothesis that an increased incentive would stimulate cooperation.

A more intensive analysis of the periodic response rates for both groups and the final disposition of non-respondents was illuminating. Table 2 shows that the periodic recovery rates for the control and test groups were comparable; neither procedure provided an advantage in obtaining cooperation more rapidly.

TABLE #2
PERIODIC RESPONSE RATE COMPARISON
CONTROL VS. TEST

WEEKS IN FIELD	CONTROL RES. RATE %	TEST RES. RATE %
6	48.9	44.3
10	59.4	56.8
FINAL	64.8	63.0

Table 3 presents the disposition distribution for non-response for the two groups. There is very little difference (2.6%) between the two refusal distributions, suggesting strongly that the increased incentive failed to reduce the household/respondent refusal rate. Since this is the area where the additional incentive should prove effective, the table further documents why an increased response rate was not obtained for the test group.

**TABLE #3
RESPONSE & NON-RESPONSE COMPARISONS
CONTROL VS. TEST**

DISPOSITION	CONTROL %	TEST %
COMPLETE	64.8	63.0
CALL BACK/ NOT AT HOME	11.9	10.9
REFUSALS/ TERMINATES	20.7	23.3
OTHER (SICK, LANGUAGE BARRIER, ETC.)	2.6	2.7
TOTAL ELIGIBLE	100.0	100.0

MRI and Chilton Research Services, which conducts our fieldwork, personally debriefed the interviewers in the study to get their insights. One particular interviewer commented "(The monetary incentive) doesn't make any difference - especially in rich areas. A lot are embarrassed to take it." Another felt that the \$10.00 "was more of a door opener for the interview. It might have made way for the product book to be left. No real impact on readership question."

Others, however, felt that the incentive was powerful. One felt that the "upfront money makes people more willing to do it" while another interviewer maintained that "money works better - people more receptive." It was interesting that many of the interviewers believed that the monetary incentive produced results although the data belied this impression. Perhaps the \$10.00 helped build the interviewer's confidence in obtaining cooperation without any material effect on the overall response rate.

To corroborate the findings even further, MRI compared the response rates in these two groups to those obtained in the similarly defined PRIZM clusters over the past two years. The historic MRI response rate in these areas was 63.6%, indicating that the test group's response rate didn't differ markedly from historic patterns.

Test Results - Pilot Study #2

The response rate comparisons for the two groups in the second pilot test also failed to support the original hypothesis. In fact, as Table 4 amply demonstrates, the cooperation level for the test group was significantly lower (one-tailed t-test at the .05 level) than that of the control group.

**TABLE 4
PILOT TEST 2
RESPONSE RATE COMPARISON
CONTROL VS. TEST**

	CONTROL (SWISS ARMY KNIFE)	TEST (MULTI-MEDIA APPROACH)
ELIGIBLE	212	230
COMPLETE	138	113
RESPONSE RATE	65.1%	49.1% *

* Significant at .05 level (one-tailed t-test)

MRI carefully assessed the reasons for these unanticipated results. One striking and compelling explanation for the lower response rate in the test group was the need for more lead time in developing multi-media contacts.

The comparable periodic response rate levels for the control and test groups are shown in Table 5.

TABLE 5
PILOT STUDY #2
PERIODIC RESPONSE RATE COMPARISONS
CONTROL VS. TEST

WEEKS IN FIELD	CONTROL RES. RATE %	TEST RES. RATE %	DIFFERENCE BETWEEN GROUPS %
6	48.0	12.8	35.2
10	60.4	36.6	23.8
FINAL	65.1	49.1	16.0

Although the experiment required simultaneous starting times for both groups, the test group procedures entailed an initial phone call as the first step. As opposed to a cold personal call in the control group following a mailout, we had expected that alerting the respondent to the study, offering an incentive and scheduling an appointment would increase cooperation. Instead, this initial phone contact merely prevented interviewers from promptly getting out into the field. This delay explains why the test group response rate level at six weeks was only half that of the control group and, although the gap eventually narrowed, the time lag proved an insurmountable obstacle in obtaining cooperation. (The logistical problem with the test procedures were also clearly documented by the Chilton interviewers).

We might speculate that gearing up earlier for the test group procedures would overcome this logistical problem but Table 6 below suggests that there are more serious difficulties with the multi-media approach.

Table 6, contrasting the non-response dispositions for the two groups, reveals that refusals and terminations were markedly higher in the test group. This result is quite unlike the similar comparison in the first study where there was no difference in refusals between the two groups. The data indicate that the refusal/terminate rate is 45.8% higher for the test group. This difference cannot be attributed to the later start. Instead, it demonstrates that the multi-media approach enables potential respondents to refuse being interviewed at the phone approach stage, significantly reducing the potential for completing an interview later on. (Although the interviewers were instructed to attempt an interview with these refusals anyway, the respondents perceived any further contact as an intrusion).

The following comments about the telephone approach by the interviewers support this explanation:

"People say no without even listening"

"The telephone prealert was not productive... People tired of approach by telephone (sales calls, etc.)"

"In-person far superior"

"Not effective. People consider approach telemarketing/sales"

"Nothing beats an in-person interview - the only way to go. Telephone easy to put off. One potential respondent irritated by the double approach. Personal is unexpected, difficult to say no."

"Waste of time. People don't remember letters or phone."

"Not necessary - advance lets them refuse."

TABLE 6
 PILOT STUDY #2
 NON-RESPONSE DISPOSITION COMPARISONS
 CONTROL VS. TEST

DISPOSITION	CONTROL %	TEST %
COMPLETE/	65.1	49.1
CALL BACK/	16.5	24.8
REFUSALS/TERMINATES	17.0	24.8*
OTHER	1.4	1.3
TOTAL	100.0	100.0

* Significant at the .10 level (one tailed t-test)

The use of a mail-gram in the test group also proved ineffectual. When interviewers mentioned the mail-gram during their personal contacts, few respondents even recalled receiving such a mail-gram. Below are sample comments from the interviewers:

"Everyone we talked to threw it in garbage (junk mail). Only 1 or 2 had read them."

"A lot of people threw them away (advertisements). Some recall, but not content."

"10% recall. No real difference."

"Most people did not mention them when asked if they remembered them."

"No recall."

Both the data and interviewer comments clearly reinforced the conclusion that a multimedia approach is unlikely to improve response rates for MRI-type surveys.

FAVOURABILITY OF RESPONDENTS

An ancillary aspect of these studies was to assess the effect, if any, different incentives have on response patterns. Previous research has shown that monetary incentives can produce more favourable attitudes of respondents towards the survey sponsor. Although the literature is not conclusive on this matter, we wanted to test whether differential treatment might engender differential response.

In this case, there is no *prima facie* definition of either survey sponsor or favourable disposition. Instead, we used screen-in and readership levels as surrogate measures of favourability. We hypothesized that respondents might reasonably assume survey sponsorship by magazines and they might also reason that higher readership levels would be beneficial to the supposed survey sponsor.

The comparable screen and read levels are shown below:

TABLE 7

	Test 1		Test 2	
	Control	Test	Control	Test
Mean Screens	11.73	12.46	9.73	11.31
Standard Error	(1.05)	(1.203)	(.839)	(.993)
Mean Reads	6.80	6.26	4.73	5.71
Standard Error	(.668)	(.501)	(.448)	(.789)

The results do not support the hypothesis. Both test groups displayed higher mean screen levels than did the respective control groups, but these differences are not statistically significant (one-tailed t-test at .05 level). The mean readership levels are higher in the control group in test 1 but lower for test 2. Neither difference is statistically significant (.05 level). It is possible that the multi-media approach in test 2, which yielded substantially lower response rates for the test group, also over-represents more heavy magazine readers. This may explain why the differences, though not statistically significant, are in the expected direction.

Conclusion:

The test findings were sobering for MRI's efforts to improve response rates. At the same time, we have not exhausted all avenues in pursuing better respondent cooperation. We concentrated on increasing the monetary incentive alone or using this procedure in conjunction with a "multimedia approach". We focused on the "economic exchange" relationship between interviewer and respondent. It might be more productive had we relied on the "social exchange" relationship established between interviewer and respondent. A number of recommended procedures for improving response rates address that relationship.

Still, it might be argued that the problem of obtaining cooperation among certain segments of the population is intractable. The simple fact may be that there exists a hard-core group of non-responders who will not cooperate regardless of the survey approach. Companies such as MRI are limited in their ability to tailor the study's subject matter or to invoke government or university sponsorship as means for increasing response rates. We need to rely more on procedures (e.g. tailoring the introduction process, "maintain contact" in Groves's terms) to heighten response and we constantly try to do so. Our model has suggested there are geodemographic correlates to non-response but that knowing these in advance doesn't imply we can provide suitable remedies. This conclusion may be unduly pessimistic, but it is not inconsistent with the declining trend in overall response rates. It also does not mean that we should quietly accept the trend; rather, we need to direct our ideas and resources to turning back the tide of non-response. It might be more productive had we relied on the "social exchange" relationship established between interviewer and respondent.

References

- Bradburn, Norman M. "Presidential Address: A Response to the Nonresponse Problem." 1992. *Public Opinion Quarterly*. 56(3) : 391-7.
- Brennan, Mike and Janet Hoek. 1992. "The Behaviour of Respondents, Nonrespondents, and Refusers Across Mail Surveys." *Public Opinion Quarterly*. 56(4) : 530-35.
- Groves, R.M., R.B. Cialdini and M.P. Couper. 1992. "Understanding the Decision to Participate in a Survey" *Public Opinion Quarterly*. 56(4) : 475-95
- James, Jeannine M. and Richard Bolstein. 1992. "Large Monetary Incentives and Their Effect on Mail Survey Response Rates." *Public Opinion Quarterly*. 56(4) : 442-53
- Lysaker, Richard L. 1991. *Using Multiple Media to Maximise Response Rates*
Paper presented at the 5th Readership Symposium, Hong Kong. February.
- Pol, Louis G. 1992. "A Method to Increase Response When External Interference and Time Constraints Reduce Interview Quality." *Public Opinion Quarterly*. 56(3) : 391-7