The effects of monetary gratuities on response rates to mail surveys have been considered in a number of studies. This meta-analysis examined: (1) the nature of the population surveyed; (2) the effects of gratuities in relation to the number of follow-ups; (3) whether the gratuity was equally effective across different populations; (4) whether the gratuity was promised or enclosed; and (5) the year of publication of the study. The bulk of the studies was done in the context of market research. Findings indicate that a monetary gratuity can be a powerful tool for decreasing non-response bias in mail surveys. Although follow-up mailings generally increased response rates in themselves, the gratuity typically had more impact than did an additional mailing. The gratuity was almost equally effective for professional and general populations. Enclosing the gratuity was more effective than was promising it upon return of the questionnaire. The year of publication had some effect in that the impact of the gratuity was slightly greater in earlier studies. The additional expense of the gratuity to improve response rates appeared cost-effective for most purposes. These findings have clear implications for survey practice. When differences in response with and without gratuities were found, they tended to show greater response integrity in the gratuity-receiving group. (SLD)
Monetary Gratuities in Survey Research: 
A Meta-Analysis of their Effects on Response Rates*

Kenneth D. Hopkins  
University of Colorado, Boulder

Arlen R. Gullickson  
University of South Dakota

Background

The mailed questionnaire is the most common type of data-gathering procedure employed in survey research. Despite its widespread use in the social and behavioral sciences, the proportion of the surveyed sample who elect to not participate is usually substantial and questions of non-response bias present serious inferential problems. With a low return rate, there is no entirely satisfactory answer to the question, "Is the experimentally accessible population representative of the target population?" Bias checks can be very useful, but there is no substitute for maximizing the proportion of respondents.

Several studies over the past half-century have found substantial increases in response rate (and consequent external validity) from the use of monetary gratuities (Kanuk & Berenson, 1975; Yu & Cooper, 1983). Most of these studies, however, have been done in the context of market research where general populations have been surveyed. The extent to which these findings are generalizable to content domains in which the respondents have a greater intrinsic interest is uncertain. Previous reviews have either ignored or not fully utilized the power of meta-analysis.

The current study considers the nature of population surveyed, the effects in relation to the number of follow-ups (if effects are minimized or disappear when follow-ups are used), whether the gratuity is equally effective across differing populations, and if the gratuity is promised or enclosed. The year of publication is also considered to see if the findings are consistent over the past several decades.

The reviews of the effects of monetary incentives have not been entirely consistent (e.g., Heberlein & Baumgartner, 1978; Yu & Cooper, 1983; Fox, Crask & Kim, 1988). Also, questions of cost-effectiveness and differential effects depending on the amount of the gratuity, and whether follow-up mailings are utilized continue to persist. The differing conclusions may be due in part to three shortcomings in current reviews. The earlier reviews based their conclusions upon a small number of studies from limited data bases (e.g., the Heberlein et al. review used only 18 studies; the Fox et al. review was based on only 15 studies). In addition, monetary incentives have often been treated as a single variable, regardless of whether the incentive was given as a gratuity prior to the return of the questionnaire or whether it was promised after the questionnaire was completed and returned. If the two "modes of delivery" have differential effects, the failure to distinguish between the two strategies confounds the effects on return rates.

Previous reviews have also failed to account for the changing value of the dollar. The psychological and purchasing-power value of a quarter incentive in 1931 when the first study was published is obviously quite different from the same coin today. If the amount of the incentive at the time the study is not calibrated to a common metric, the magnitude of the effects on the return rate is possibly confounded by the changing value of the dollar.

the effects for various amounts of gratuities will be distorted. In this study the U.S. Consumer Price Index was used to calibrate the value of the monetary incentives for the time-frame of each study to 1987 dollars.

Method

The review of literature began with an on-line computer search of data bases in education, sociology, and business. Reference lists in the studies captured in these searches were then examined to identify additional pertinent studies. More than thirty studies from several different disciplines were located which included both an experimental (E or gratuity) group and a (usually randomly assigned) control group (C or no gratuity). These studies provided more than fifty effects for the meta-analysis from which the influence of a gratuity on response rate could be examined.

The coding for the meta-analysis considered several factors: (1) the amount of the gratuity, calibrated to 1987 dollars, (2) whether the gratuity was enclosed or promised, (3) whether the population surveyed was general or professional, (4) whether or not follow-ups were included, and if so how many, (5) the increase in return rate for each follow-up, (6) the year the study was conducted (or published if information was not given), and (7) sample size. Two effect-size outcome measures for the gratuity were calculated from the response-rate data: (1) the absolute increase (A) in the percent of the sample returning the survey instrument (the difference in percent of response for the respective experimental and control groups), and (2) the relative increase (R) in response rate (the ratio of the experimental group's response rate to the response rate of the control group). Because of its lack of ambiguity, the findings will focus on the absolute difference except when the pattern of effects is different for A and R. Of special interest is whether the effect of the monetary gratuity depends on the response rate under control conditions, e.g., is the effect of the gratuity influenced by the willingness of the C group to respond without a gratuity?

Analyses were conducted on each of the coded factors individually. In addition, the collective impact that the respective independent variables on the response rate was examined via multiple regression, which optimized the prediction of the return rate using the various coded variables as predictors.

Results

Taken as a group, the mean response rate for the initial mailing in the C groups (n = 41) was only 32.7% (Mdn = 33.3%); the variability, as expected, was substantial: s = 13.2%, with a range from 10% to 59%. The fact that the bulk of the studies were done in the context of market research (where the average response rate is about 40% (Heberlein & Baumgartner, 1978) probably accounts for the generally low return rates.

Descriptively [see box-and-whisker plots in Figure 1] the distribution was slightly positively skewed (γ1 = .27 and platykurtic (γ2 = -.76) (viewed inferentially, these values are not sufficient to reject the normality hypothesis, z1 = .73; z2 = 1.39). On the other hand, the response rate for the corresponding E groups was much higher (t = 12.6, p < .0001), with a mean of 51.2% (Mdn = 50.0%). The variability in return rates within the E group

---

1Initially this factor had three levels, but since the return rates for the middle and low groups were virtually identical, the middle and low levels were combined.
2Since the outcome measure is not arbitrary, and is directly interpretable, it was judged to be a superior measure to the commonly used effect size (standard deviation) metric.
3When a study did not provide separate return rates for the first and second mailings, in this instance, the total return rate was used for mailing #1; this causes a conservative estimate of the effect of the gratuity.
was somewhat greater: $s = 16.6 (t = 3.53, p < .001; \text{Glass} \& \text{Hopkins, 1984, p. 268})$, the range extended from 18\% to 86\%; greater variability in the gratuity group is not surprising since the amount of the gratuity, and whether it was enclosed or promised, also varied within the group. The distribution for the 54 $E$ observations differed little from a normal distribution ($\gamma_1 = .08, \hat{z}_1 = .23; \gamma_2 = -.43, \hat{z}_2 = .67$). Not surprisingly, there was a high correlation between the $E$ and $C$ response rates ($r = .774$; see Figure 1).

Figure 1. Response Rates with (upper curve) and without a Monetary Gratuity (lower curve), with corresponding Box-and-Whisker Plots for $E$ and $C$ Groups, and for the Absolute Differences ($A$) in Response Rates.
**Relationship between Absolute (A) and Relative (R) Return Rates.** In order to determine whether the focus of the analysis could be on the absolute, rather than the relative effect of the gratuity, the relationship between absolute (A), and relative (R) measures was studied to determine whether the two variables yielded the same pattern of findings. The linear correlation between A and R was substantial, .689. When the logarithm (log) transformation of R was used, the relationships with A became essentially linear, and the linear correlation increased to .760.

The correlations of the absolute and relative effects with the other coded variables were similar with two logical exceptions: A was positively correlated with the E return rate (.62), and R was substantially negatively correlated (-.54) with the C return rate.

The absolute effect of the gratuity (A) did not depend the C return rate (r = -.01); thus, the absolute increase in return rate was stable regardless of the return rate under "normal" conditions. In other words, the effect of the gratuity was to increase the absolute response rate (A) substantially—by an average of 18.0% (Med = 17.0%), regardless of whether the base return rate was low or high. The distribution of A [see Figure 1] was quite variable ($s = 11\%$) and skewed positively ($\gamma_1 = .58$). The .95 confidence interval for the mean effect on the absolute response rate extends from 15.1% to 20.9%.

A more complete graphic display of the ...ings is given in Figure 2. The lower curve gives the response rates for each of the control groups (rank-ordered from low to high); the upper curve gives the response rates for the associated E groups. Notice that in 53 of the 54 instances, the return rates associated with the E groups were higher. [The single exception is from Denton (1988), a study with a very small sample size (36)].

**The Relationship between Ancillary Factors and Effect**

*Promised vs. Enclosed.* Enclosing a monetary gratuity appears to have a much more powerful effect than when a gratuity is promised upon the return of the questionnaire. Enclosing a gratuity increased the absolute response rate by almost three times as much as did the promised gratuity, 20.0 (n = 45) vs. 7.7% (n = 8)![n = 8]! This occurred in spite of the fact that the promised gratuities tended to be larger. Indeed, when ANCOVA was used to equate the two groups on amount of gratuity (log10 of 1987 dollars) and type of population, the adjusted return rate for the promised gratuity differed very slightly from that without a gratuity. Perhaps the promised gratuity is perceived more in the "pay" mode (in which case it is virtually always inadequate and could be viewed as an insult), rather than as a gratuity. Note also that since the overall absolute effect described above was an aggregate of both enclosed and promised gratuity, the effect of enclosed gratuities *per se* is underestimated; an absolute increase of 20% is a better estimate based only on those studies with enclosed monetary gratuities.

**Amount of Gratuity.** Using the value of the gratuity (calibrated to 1987 dollars), four levels of the gratuity factor were formed: (1) $$.50 or less (n = 17; mean gratuity = $.29), (2) $.51-$1.99 (n = 13; mean = $.75), (3) $1-$1.99 (n = 10; mean = $1.30), $1-$1.99 (n = 10; mean = $1.30).  

---

4All statistical inferences in this analysis assume that the observations (effects) are statistically independent. Since several studies contributed more than one effect (e.g., had two or more E groups with different gratuity amounts, each being compared to a common control group), the effects are not purely independent. The studies contributing only a single effect were compared with studies having more than one effect, in both means (t) and variance (F) with respect to return rates for C, E, and A and R. In none of the eight inferential tests was significance approached, the smallest p-value being .15 for A (where the mean for the studies with a single effect was somewhat higher—21.2% vs. 16.6%). Consequently, it is judged that the consequence of imperfect independence does not seriously distort the inferential statements.
and (4) \$2 or more \((n = 14; \text{mean } \$3.84)\). Since the groups differed with respect to promised-vs.-enclosed gratuity, it was necessary to covary on this factor. The ANCOVA allowed the null hypothesis to be rejected for the absolute increase in response rate \(F_{3,46} = 7.93, p < .001\). The effect of the gratuity in the lowest group (adjusted mean = 5\%) was much below that of the three groups with larger gratuities.

**Type of Population.** Under usual (C) conditions, the return rate for the professional populations (38\%, and coded "1") was slightly greater \((r = .09)\) than for the general (61\%, and coded "0") populations. Of special importance is whether the gratuity effect interacts with type of population; the effect of the gratuity was almost as strong within professional groups as for the general population--the correlation between population and absolute increase in return rates was only -.08. This finding is consistent with that reported by Goodstadt *et al.* (1977).

The type of population factor is probably related to the salience of the content of the survey, which has been shown to be highly related to return rates (Heberlein & Baumgartner, 1978), thus the issue is more complex than just the definition of the target population.

**Effects and Follow-ups.** Only nine studies included more than one mailing. Follow-ups were less common in the earlier studies; they were also more common with professional populations. Only three of the studies included two follow-ups.

It is well known that follow-up mailings substantially increase return rates. What is of particular interest in the present investigation is whether the effect of gratuity is reduced or eliminated by additional mailings. In the nine studies having more than one mailing, the response rate for the control group increased from 35.5\% to 54.3\% (18.8\%) after the second mailing--very similar to the figure of 19.9\% reported in Heberlein & Baumgartner's (1978) review. In the experimental group the initial return rate of 58.7\% increased to 68.0\% (9.7\%) after the second mailing. The absolute difference in response rates for the E and C groups in these nine studies decreased from 23.2\% after one mailing to 13.7\% after two mailings; the gratuity effect remained substantial even after two mailings. Note that the return rate for group C after the second mailing remained below that of the E group for only one mailing.

For the three studies having three mailings, an absolute difference in return rates of 10.4\% favoring the E group remained even after the third mailing. Note that the return rate in the E group after two mailings (82.4\%) remained higher than that for the C group after three mailings (78.7\%). Thus, it can be concluded that the effect of the gratuity on return rates remains large enough to be of practical significance even when three mailings are involved.

**Year of Publication.** The 54 effects came from studies having a mean year of publication of 1970 (\(Mdn = 1972\)), with a range from 1931 to 1986. Year-of-publication did not correlate with adjusted amount of gratuity \((r = .01)\) or the promised-vs.-enclosed dichotomy \((r = -.00)\), but did have a slight correlation with the return rates \((r = .19)\), perhaps suggesting some methodological progress in return rates over the last half-century (the correlations remained essentially unchanged when the type-of-population was partialled out). The impact of the gratuity appears to have been slightly greater in earlier studies; the correlation of Year with the absolute difference was -.22 \((p = .06)\).
Predicting Return Rates.

To assess how accurately the total (percent) response rate ($\hat{Y}$) could be predicted, and to assess the relative contributions of each of four independent variables in the prediction, a multiple regression equation was developed based only on those studies with enclosed gratuities using the following predictors:

- $G$: Gratuity group ($0 = \text{None}, 1 = \$.50, 2 = \$.51-.99, 3 = \$1-1.99, 4 = \$2$)
- $M$: The number of mailings (1, 2, or 3)
- $P$: Type of Population surveyed ($1 = \text{General}, 2 = \text{Professional}$)
- $D$: Decade of the study ($4 = \text{before 1950}, 5 = 1950s, ..., 8 = 1980s$)

A multiple correlation of .627 resulted (.600 after correction for error-fitting), using the raw-score regressions equation below.) Beta weights are given below each predictor to help assess the strength of each predictor.

$$\hat{Y} = 7.47G + 12.23M + 3.68P + .89D + 13.2$$

For example, the predicted rate of return for a mail survey of a professional population with a $1 gratuity enclosed and two mailings in the 1980s is:

$$\hat{Y} = 7.47(3) + 12.23(2) + 3.68(1) + .89(8) + 13.2 = 70.9\%.$$  

The same study with two mailings, but without any gratuity would have a predicted return rate of only 48.5%; the increase in return rate from the $1 gratuity is predicted to be 22.4%. The standard error of estimate for these prediction is quite large (14.9%); the predictions are far from precise. As would be expected, there are many other factors that influence return rates than those explored here.

Cost/Benefit Considerations

Using the predicted return rates from the regression equation above, and assuming a professional population survey by mail under non-gratuity conditions, the return rate for the initial mailing (in the 1980s) is estimated to be only 36.4%; as indicated above, this would be predicted to increase to 48.5% after a second mailing. Assuming first class mailings of one ounce or less, with enclosed stamped return envelopes, the cost per returned questionnaire would average approximately $1.37 for a single mailing, and $1.68 for two mailings.

The same population surveyed with a $1 gratuity, would have predicted return rates of 58.6% and 70.9% after one and two mailings, respectively. The mailing costs would average $2.56 per returned questionnaire after one mailing and $2.41 for two mailings (no gratuity enclosed in second mailing). If a 25c gratuity were enclosed, the return rates for one and two mailings would be predicted to be 43.7% and 55.9% respectively. The cost per returned questionnaire would average $1.72 for a single mailing, and $1.85 for two mailings--the latter value being less than 10% above that for the non-gratuity group.

Of course these cost comparisons do not give the benefit perspective. The critical question is, "does the improved external validity of the statistical information justify the additional cost?" With two mailings without a gratuity, a return rate of only 48.5% is predicted, thus more than one-half of the target population is not being represented. With a $1 gratuity and two mailings, the predicted return rate of 70.9% is a great improvement. Related to typical manpower costs, the additional expense to shore up a chief weakness of
mail surveys would appear to be highly cost-effective for most purposes. The $400 addition cost of the prototypical survey $n = 400$ with a $1$ gratuity appears nominal.

Even in a fixed, minimal cost context, the gratuity strategy should be considered. A smaller sample with a gratuity $71\%$ rate of return will be more credible than a larger sample with a $49\%$ return rate. Note that without the gratuity, it would take 206 mailings to yield 100 returns, whereas with the gratuity it would take only 141. Note also that the study with a larger sample also has additional costs other than mailing costs--more envelopes, more copies of the questionnaires, more labor costs. All things considered, the survey of 300 with a $1$ gratuity will cost less money than a survey of 600 without a gratuity. But the important methodological question remains, "which study will provide a more valid representation of the target population"? Stated differently, which study would yield a more believable confidence interval? Most statisticians would prefer findings based on data from a return rate of $71\%$ for a sample of 300, than those based on data from sample of 800 with a $49\%$ return rate because the larger sample has a $51\%$ non-response rate vs. a $29\%$ for the smaller sample. (Of course bias checks, e.g., comparing the responses from the first with the second mailings, remain critical even with the higher response rates.)

Another promising option in terms of cost-efficiency that appears to have been essentially unexplored is the use of a gratuity only in the second mailing to those who did not respond initially. The only study (Huck & Gleason, 1974) on this variation found the gratuity effect in the initial mailing of $28\%$ was reduced to only $7\%$ when a gratuity was included in the second mailing to the $C$ (but not in the second mailing to the $E$ group).

**Discussion**

Survey research findings based on a low response rate is likely to serve poorly as a basis for policy decisions and practice. Because the mailed survey is so frequently incorporated into the research/decision making process by policy makers in government, education, and society, it is particularly important that the quality of the research data be as good as possible.

The findings of this meta-analysis indicate that a monetary gratuity can be a powerful tool for decreasing non-response bias in mail surveys. The results also indicate that the use of additional follow-up mailings reduce the difference between the $E$ and $C$ return rates, additional mailings, although important, do not eliminate the practical benefit of the gratuity. The gratuity typically has more impact than an additional mailing. A monetary gratuity appears to be almost equally effective for professional (who typically would be assumed to have more intrinsic motivation to respond) and general populations. These findings appear to have clear and strong implications for survey practice, especially in light of the fact that comparative studies have found no or negligible difference in the response as a consequence of the gratuity (Hancock, 1940; McDaniel, 1980; Maloney, 1954; Mizcs et al., 1984; Shuttleworth, 1931; Wotruba, 1966; Zusman & Duby, 1984). When differences have been found, the have tended to show greater response integrity in the gratuity group.
References


Denton, J. J., Chiou-Yueh Tsai (1988). Two Investigations into the Influence of Incentives and Subject characteristics on mail survey responses in teacher education. Unpublished study from Instructional Research. Laboratory, College of Education Texas A & M University, College Station TX.


