International Journal of Internet Science
2006, 1 (1), 58-70

International Journal of Internet Science

IJIS.NET

Incentives in Web Studies: Methodological Issues and a Review

Anja S. Göritz

University of Erlangen-Nürnberg, Germany

Abstract: Two meta-analyses review the effectiveness of incentives in Web studies. The first analysis summarizes 32 experiments on the impact of material incentives on response. It is based on a sample of 212,810 people. A significant effect (odds ratio = 1.19; 95% confidence interval: 1.13—1.25) is revealed, indicating that incentives motivate people to start a Web survey. The second analysis comprises 26 experiments on the impact of incentives on retention and is based on a sample of 7,073 people. It yields a significant effect (odds ratio = 1.27; 95% confidence interval: 1.12—1.44), indicating that once people have accessed a survey for whatever reasons, they are more likely to finish if an incentive is offered. The established incentive effects are stable across various study characteristics. Conclusions did not appear to be due to publication bias.

Keywords: Incentive, meta-analysis, experiment, response, retention

Introduction

As stand-alone Web surveys (Cook, Heath, & Thompson, 2000) and online panels are widespread (Göritz, Reinhold, & Batinic, 2002), economical and methodological success go hand in hand with acquiring knowledge about the successful solicitation of participants. A traditional way of motivating people to take a survey is to offer incentives. Meanwhile a considerable number of experiments have evaluated the effectiveness of incentives to increase participation in Web-based studies. This evidence, however, is scattered, and researchers and practitioners may find it difficult to arrive at a global picture. This paper summarizes the available data to find out whether material incentive increase participation in Web studies, and if yes, how large effects are. The findings of this summary furnish a basis for deciding whether to use material incentives in a given study. Moreover, several study characteristics are examined as to whether they influence the impact of incentives on participation. The results of these analyses enable researchers and practitioners to make full use of incentives by tailoring them to the situation at hand.

This paper starts with a comparative overview of the effectiveness of incentives in offline studies. Next, the logistic and methodological particularities of employing incentives in online studies are discussed. Finally, it is meta-analyzed if and under what conditions incentives are effective in Web studies.

Anja Göritz can be contacted at the following address: Department of Psychology, esp. Organizational and Social Psychology, University of Erlangen-Nürnberg, Lange Gasse 20, 90403 Nürnberg, Germany. Phone: +49-(0)911-5302373, Fax: +49-(0)911-5302243, anja.goeritz@wiso.uni-erlangen.de, http://www.goeritz.net.
Acknowledgments: This work has been supported by a University of Erlangen-Nürnberg post-doctoral scholarship (HWP). I thank Klaus Moser for helpful comments on an earlier draft.

Numerous reviews discuss the impact of incentives on *offline* response rates. A meta-analysis of mail-survey response rates found an overall incentive effect of d = 0.24 (Church, 1993). This means that, on average, in the incentive condition the response rate is roughly one quarter of the pooled standard deviation higher than in the no-incentive control condition. Another meta-analysis encompassing face-to-face and telephone surveys established a third of a percentage point average difference between incentive and zero-incentive conditions per dollar of incentive paid (Singer et al., 1999). Yu and Cooper (1983) obtained overall effects of monetary and nonmonetary incentives on response in mail, telephone, and face-to-face surveys of d = 0.30 and d = 0.18, respectively. Moreover, several experiments have indicated that response increases with rising amounts of incentives (cf. Church, 1993; Singer et al., 1999; Yu & Cooper, 1983). This relationship, however, is not always linear (Armstrong, 1975; Fox, Crask, & Kim, 1988; Linsky, 1975; Mizes, Fleece, & Roos, 1984).

On account of medium-related differences, it is questionable whether findings from offline studies can be generalized to the online realm. For the offline realm, researchers widely agree that if incentives are to increase response, they need to be given in advance instead of being made contingent on the return of the questionnaire (e.g., Armstrong, 1975; Church, 1993; James & Bolstein, 1992; Linsky, 1975). Provided this also holds with Web studies, because anonymous participants cannot easily be sent a dollar bill or a physical token of appreciation (Couper, 2000), researchers need to find Web-appropriate pre-paid incentives. If respondents are approached during their Web use as in most ad hoc surveys, they can still be awarded something in advance before taking part in a survey. For example, potential participants can be given a code to access or download material for free, such as stock prices, payable news, e-books, games, or software. As an additional benefit, multiple participations in order to get more incentives would be discouraged because informational incentives can be copied at negligible cost anyway. However, these prepaid incentives might call forth a bias toward more computer-literate users and it is doubtful that they are equally interesting to everybody. In list-based web surveys, however, people are personally addressable, which makes it easier to deliver a prepaid incentive. For example, one might include electronic gift certificates from online-shops or redeemable loyalty points in the invitation e-mail. Even electronic money can be sent by using products such as *Paypal.* However, e-money is not "money in the hand" and collecting it is more or less cumbersome (depending on a user's Web literacy and willingness to register with a transaction party). Therefore, prepaid electronic money is not (yet) comparable to prepaid rewards offline. For list-based samples whose postal address is known, there always exists the possibility to switch back to letter or parcel in order to deliver a prepaid incentive.

Besides these logistic challenges, there are differences between Web and offline surveys with regard to the respondents and the participation process. Unless respondents have unlimited online access (e.g., for a fixed monthly fee or at university/work) they have to pay for their online session (e.g., provider fee, dial-up charges). By contrast, researchers incur the monetary cost of participation in most cases offline (e.g., by including a stamped envelope, phoning the interviewees, or providing a toll free number). Therefore, when collecting data online, incentives might be necessary to counteract an over-representation of respondents who have a free or flat rate Internet connection. Another difference between online and offline is that respondents in the more anonymous medium Internet might be distrustful that, for example, a lottery to raffle rewards for survey participation actually exists and will be run fairly (Porter & Whitcomb, 2003). Moreover, participation in Web rather than in face-to-face studies is often more comfortable and easy. For example, respondents can participate at a convenient time in their own home while wearing casual clothes. Also, the time effort tends to be less than offline and the questionnaire is often more appealing because multimedia stimuli can be included. Finally, especially Web users who are approached while surfing the Web (i.e., intercepted users) are easily distracted by voluminous amounts of information, which makes them less tolerant toward material that is not of interest to them, such as a banner or newsgroup posting announcing a survey. An attractive incentive might compensate for this (Simsek & Veiga, 2001) or act as an attention grabber.

The situation is somewhat different with online panel surveys. While they share the same medium with ad hoc Web surveys, the style of solicitation is different. Panelists who are invited to a survey have signed up with the panel beforehand (opted in) to occasionally take part in studies. Thus, they have already made up their mind about taking part in principle, and a study invitation does not come unexpectedly. Furthermore, with list-based surveys and therefore also in online panels, potential respondents are invited to a study via E-mail and thereby decide on participation in a context farther away from concurrent distractions, compared to seeing a banner while surfing the Web. In addition, it can also be assumed that with online panelists, a certain degree of trust has already been established. In contrast to the demand to give incentives in advance, online panelists expect to receive an incentive *after* their participation. Furthermore, each panelist's participation history is recorded. As this information can be used to decide who is invited to future studies, panelists might think twice whether to decline an invitation and they also might answer more conscientiously. For one-time respondents this is irrelevant. For all these reasons, incentives might be less effective in list-based than in ad hoc surveys. What speaks against this assumption, however, is the fact that panelists are requested to participate repeatedly and therefore the researcher cannot take advantage of the curiosity-bonus of one-time studies. Moreover, incentives might be more important in opt-in panels with a high frequency of studies (i.e., if panel members receive several survey requests per week, or indeed several per day for some panels) than in less busy opt-in panels and in list-based samples for whom a survey request is still a relatively rare or unexpected phenomenon.

Besides their possible benefits, such as higher response and lower drop-out rates, incentives bring about potential dangers to survey validity. Some drawbacks are the same off- and online: Incentives might attract particular respondents (Enander & Sajti, 1999) and thus alter a sample's composition; incentives can also place a burden on intrinsically motivated respondents (cf. Deci, 1971) with the ultimate effect of driving respondents away, or they might bias the study's outcome. Some downsides, however, are novel or exacerbated when conducting Web studies – especially ad hoc surveys. Firstly, incentives might instigate some people to fill out the survey many times, and it is not always possible to detect skillful fraud. With list-based surveys, the potential for these abuses is reduced because respondents are identifiable and by definition can only complete the survey by invitation, and only once (Comley, 2002). Second, the risk exists that people fill in rubbish data to get to the end of the survey quickly in order to be eligible for an incentive. Without an incentive, they tend to give up when bored and are therefore easily identifiable. Finally, legal aspects of offering incentives might differ across countries.

In order to assess whether incentives in Web surveys are worthwhile despite these dangers, we need to know whether and how strongly incentives increase response rates and decrease dropout rates. To this goal, a comprehensive literature and WWW search was conducted. The aim was to identify every Web-based incentive experiment on response or retention. Response is defined as the number of solicited people who call up the first page of a Web survey in contrast to all people who fail to move to the first survey page. Retention is the number of responding people who stay until the last page of the study, in contrast to those who loaded the first study page but then drop out before reaching the last study page[1]. The hypotheses are as follows:
H1. Response is higher when an incentive is offered than if no incentive is offered.
H2. Retention is higher when an incentive is offered than if no incentive is offered.

Method

Criteria for Inclusion

In order to be included, the studies needed to be incentive experiments and not merely surveys, thus allowing a synthesis of incentive net effects (cf. Fox et al., 1988). All experiments needed to contrast at least one material incentive-condition with a no-incentive control condition. Either the response or retention rate or both needed to be reported. Only experiments that both solicited and surveyed participants online were admitted. That is, in the case of list-based studies, participants needed to be invited by E-mail and in the case of ad hoc surveys, they needed to be recruited on the Internet through links, newsgroup postings, banners, search engine listings, pop-up windows, etc. Mixed-mode surveys were excluded, for example, if potential respondents were solicited on the phone to take a Web-based survey.

Literature search

To locate relevant experiments, the following databases were searched: PsycInfo, Psychology and Behavioral Sciences Collection, Econlit, ERIC, Sociological Abstracts, Dissertational Abstracts, Medline, Ingenta, and Educational Research Abstracts, using a combination of the following key words: response, retention, participation, dropout, effect, effectiveness, lottery, prize draw, gift, gift certificate, present, cash, incentive(s), inducements(s), survey(s), study, experiment, Web, Internet, WWW, and online. In addition, the WWW was scanned with Google search engine using the queries above. Finally, personal contacts were tapped, references from the already located experiments tracked, and conference programs and proceedings perused.

The search brought to light 14 sources (cf. references section), each reporting at least one incentive experiment that met the criteria. One of the initially considered experiments was excluded because upon closer inspection it did not fit the criteria for inclusion[2] while two experiments with minor quality flaws were retained for analysis[3].

---

[1] The retention rate is the inverted dropout rate (i.e., 100 – dropout rate = retention rate). Retention was used instead of dropout because it has the same polarity as response (i.e., high response and retention are the desired outcomes).

[2] In order to recruit participants for a Web survey, Tuten, Bosnjak, and Bandilla (2000) rotated different banner ads on two search engines. One banner read "Your opportunity to win valuable prizes" and the other one "Your opportunity to

Thus, 32 comparisons on the impact of incentives on response and 26 on retention were left. Several experiments (i.e., Bosnjak & Tuten, 2003; Cobanoglu & Cobanoglu, 2003; Göritz, 2005; Porter & Whitcomb, 2003; Tuten, Galesic, & Bosnjak, 2004) entailed two or more experimental interventions, for example, three different amounts of raffled money compared to a single no-incentive group. To avoid statistical dependencies in these cases, the control group – while maintaining the original proportions – was broken up into as many parts as there were treatment groups. Apart from rounding errors, the effects' sizes were unaffected by this procedure[4]. Moreover, several experiments reported retention rates at different stages of the survey, for example at Page 1, at Page 2, and so on. However, only overall retention (i.e., based on the last survey page) was included in the meta-analysis. The sample sizes used across comparisons pertaining to response ranged from 331 to 170,265 with a median of 675 respondents. The sample sizes used across comparisons pertaining to retention ranged from 93 to 791 with a median of 140.

Computation of Effect Size Estimate

There were two outcome variables and therefore two separate meta-analyses were conducted: Response is a measure for how much incentives persuade people to have a look at a questionnaire, whereas retention reflects how much incentives affect respondents' persistence to stay until the end of a study. Dropout can be the result of a participant's deliberate decision to quit or result from disconnection due to technical failure. This uncertainty on the individual level, however, does not pose a problem with between-group comparisons.

As the outcome measures are dichotomous (responded/refused and retained/dropped out), odds ratio (OR) was chosen as effect size measure (Fleiss, 1994; Haddock, Rindskopf, & Shadish, 1998). An OR is the odds of the event occurring in one group (e.g., incentive group) divided by the odds of the event occurring in the other group (e.g., control group). For example, in Frick, Bächtiger, and Reips (2001), of all 789 participants, 394 were assigned to the incentive condition. The ratio of events in the incentive condition (i.e., completed the survey) to non-events (i.e., dropped out prematurely) is 357/37 = 9.65. Of the 395 participants allocated to the control condition, 322 completed the survey, whereas 73 did not, yielding odds of 322/73 = 4.41. The OR then is 9.65/4.41 = 2.19. If an experimental intervention (e.g., offering an incentive) has no effect, the OR is 1. If it reduces the chance of having the event, the OR is less than 1; if it increases the chance of having the event, the OR is bigger than 1. The smallest value an OR can take is zero. Thus, in Frick et al. (2001) with an OR of 2.19, incentives increased the odds of a person staying until the last survey page by 119%.

Each of two meta-analyses comprised the following steps: 1. heterogeneity examination by means of the *Q*-test, 2. pooling of individual effect sizes to derive an overall effect size estimate, 3. exploratory moderator analysis, and 4. consideration of publication bias. Parts of the analyses were done using the meta-analytical freeware *MetaAnalysis 3.0* (Chang, 2001) and the commercial software *Comprehensive Meta-Analysis*. For pooling individual ORs, a random-effects model was chosen (DerSimonian & Laird, 1986) because it is the more general model. If ORs turn out to be homogenous, it reduces to the simpler fixed-effects model.

Moderator Analyses

If available, the following information was collected and coded from each comparison: sample type (there were 44 list-based and 14 ad hoc recruited samples), lottery incentive or not[5] (51 with lottery and 6 no lottery), monetary (4 comparisons) or nonmonetary incentive (e.g., free delivery, gifts, gift certificates, lottery: 54

---

contribute to an important study". Thus, with the first banner, neither type of incentive nor its value nor that people were required to participate in a study were communicated.

[3] Dropout between the third and the last page of O'Neil and Penrod's (2001) survey was confounded with the requirement to disclose personal information. Therefore, overall retention at the penultimate (and not the last) survey page was extracted. Instead of retention, Bosnjak and Tuten (2003) report an "incompletion" rate. It combines genuine dropouts, lurkers, and item-nonresponders. Because dropping out, lurking, and item omission all reflect the seriousness of participation, the experiment was retained for analysis.

[4] Consider the example of a study where two incentive conditions were compared to a single no-incentive control group: If in the control group 200 people responded to the study and 100 did not, for each of the two comparisons the size of the control group was adjusted to 100 respondents and 50 non-respondents.

[5] In addition, there was one comparison in Cobanoglu & Cobanoglu (2003) where the incentive consisted of a lottery *and* a gift.

comparisons), timing of the incentive (2 prepaid and 56 on return/promised), offer of a result summary[6] (18 with summary and 40 without), survey sponsorship (52 nonprofit/governmental and 6 commercial), study recency[7] (5 comparisons in 1999, 5 in 2000, 11 in 2001, 3 in 2002, and 34 in 2003), lottery payout[8] (across 51 comparisons the mean lottery payout was 137 USD), and number of survey pages[9] (mean of 10 pages across 53 comparisons). These study characteristics were examined as to whether they moderated the incentive effects. The dichotomous study characteristics were subgroup-analyzed according to a random effects model[10]. The influence of the continuous covariates *study recency*, *number of survey pages*, and *lottery payout* on the log incentive effect was examined using unrestricted maximum likelihood mixed-effects meta-regression.

Results

Response

Among the comparisons that pertained to response, there was significant heterogeneity, $Q = 76.35$, $df = 31$, $p < .001$, indicating that these comparisons were probably not representing a common phenomenon or that outliers were present. To identify outliers, the radial plot was inspected (Galbraith, 1988). Eyeballing the radial plot in Figure 1 shows reasonable homogeneity, with the exception of two outliers (i.e., Rager, 2001 and one comparison in Tuten et al., 2004). This judgment is confirmed by a *z*-test using a *p*-value ≤ .01 as indicating an outlier (Greenland, 1987). According to the *z*-test, the outlying comparisons had *z*-values of 5.62, $p < .001$ and 3.91, $p < .001$, respectively. After exclusion of the two comparisons, the remaining comparisons are homogenous, $Q = 22.61$, $df = 29$, $p = .79$

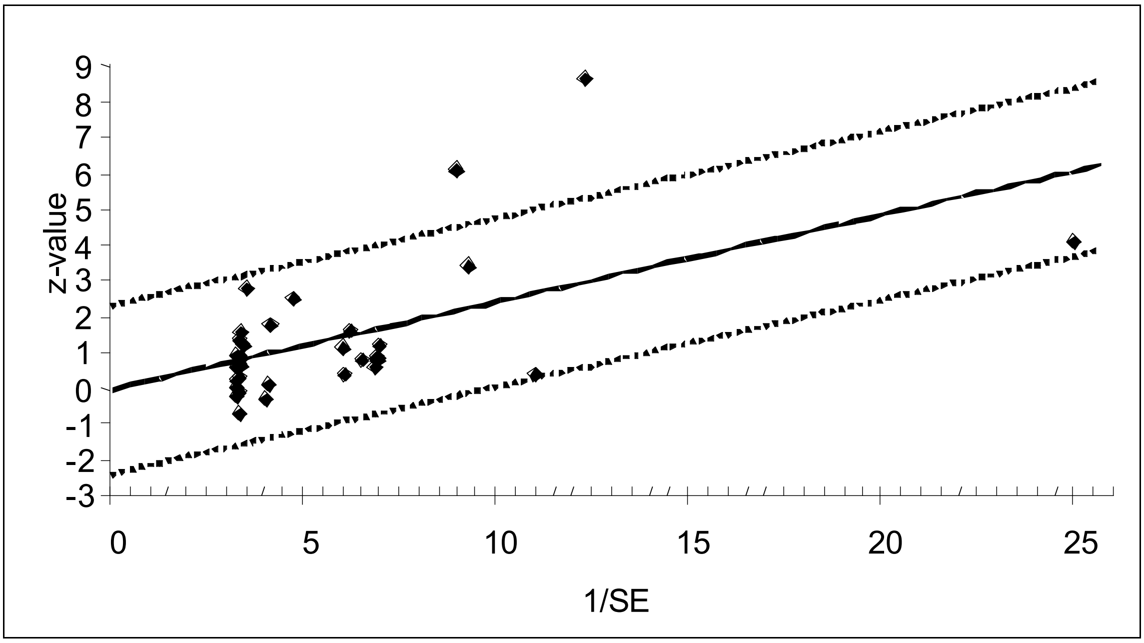

Figure 1. Radial Plot of the Effect of Material Incentives on Response.[11]

[6] For an experiment to be included, the summary had to be offered to both incentive and control group, so that the groups only differed in the reception of the material incentive.

[7] If reported, the year the survey was conducted was extracted; otherwise the year of the publication/presentation was taken.

[8] Only comparisons with lotteries were taken into account for the moderator analysis pertaining to lottery payout. Because the number of invitees was often not communicated to potential participants, lottery payouts cannot be compared with individual payments. If several prizes were raffled, they were totaled. All non-US currencies were converted into USD using the exchange rate from the time the survey was conducted (http://www.oanda.com/convert/fxhistory).

[9] If only the number of questions was reported, six questions were counted as one survey page.

[10] Random effects within subgroups was computed assuming a common among-study variance across subgroups (τ is computed within subgroups, and pooled across subgroups).

[11] Radial plot of the effect of material incentives on response. The point estimate for each comparison (z-value) is plotted against its precision (1/SE). Comparisons lying outside the 99% confidence interval of the regression line with the constant set to zero (dotted lines) are contributing significantly to between-comparison heterogeneity.

The overall effect for the impact of material incentives on response is OR = 1.19 with a 95% confidence interval (CI) ranging from 1.13 to 1.25. The effect is highly significant, $z = 6.47$, $p < .001$. Thus, material incentives increase the odds of a person responding by 19% over the odds without incentives (cf. Figure 2, left panel). Transforming this OR into the standardized mean difference using the method recommended by Hasselblad and Hedges (1995) yields $d = 0.10$. With the two outlier comparisons taken into account, the overall effect would have been OR = 1.28 (95% CI = 1.16 - 1314), $z = 5.07$, $p < .001$, $d = 0.14$.

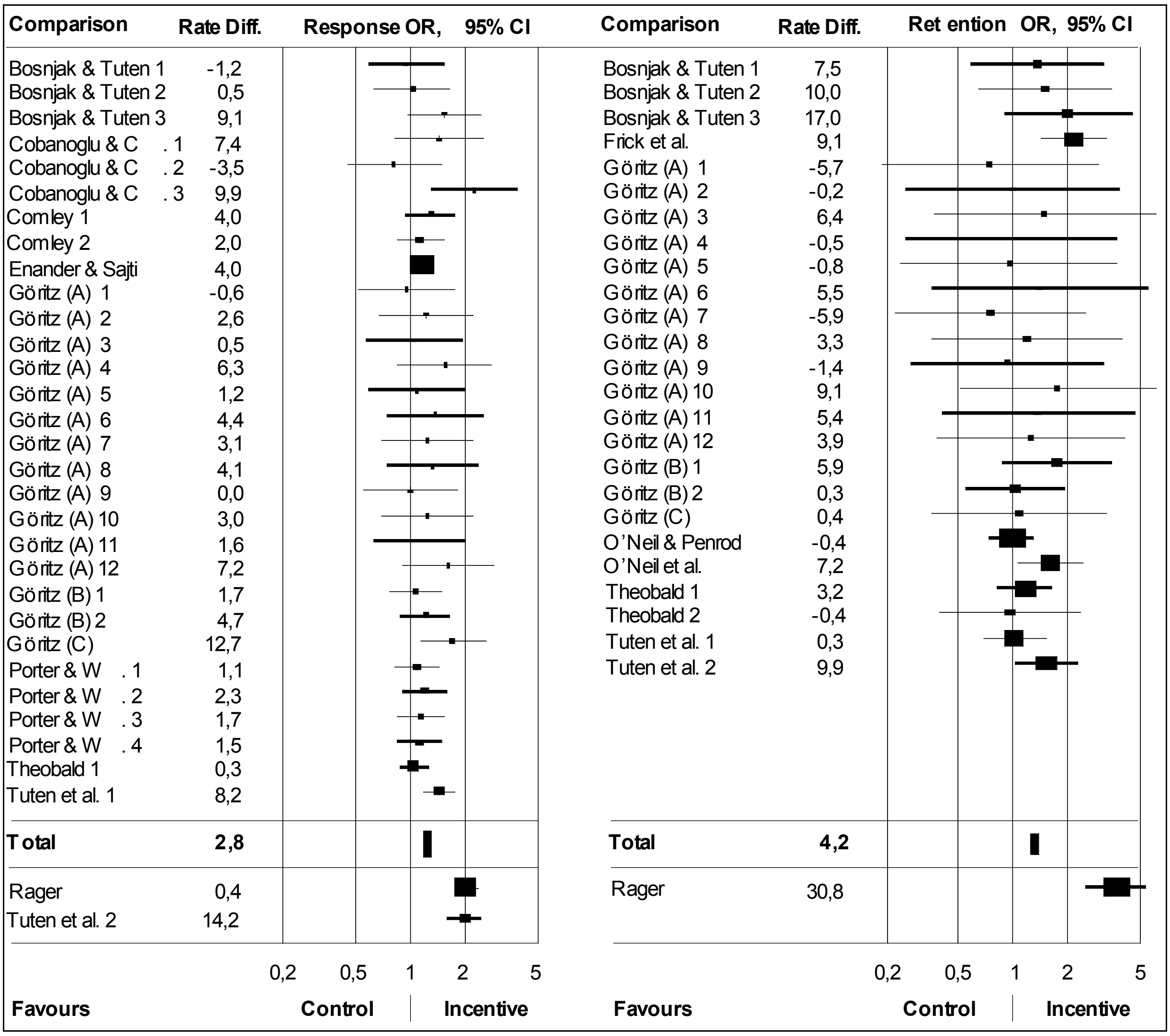

Figure 2. Individual and Overall Effect Sizes of the Impact of Material Incentives on Response and Retention in Web Surveys.[12]

The moderator analyses, which were conducted without the outliers, show that none of the study characteristics significantly correlate with the incentive effect (cf. Table 1, middle column). This outcome is not surprising considering the comparisons' homogeneity after exclusion of the outliers.

[12] The square on each line of the two forest plots represents the odds ratio (OR) for this comparison. Its size indicates the relative weight of that comparison toward the combined result. In addition, the difference in the response/retention rate due to employing an incentive is listed for each comparison. Positive values indicate an increase in the response or retention rate due to using an incentive, whereas negative values indicate that the response or retention rate was lower in the incentive condition than in the control condition. The line "Total" indicates the weighted average rate difference across all comparisons and the overall OR of combining the data from all comparisons (excluding outliers). Outlying comparisons are listed underneath the pooled results. Letters behind a comparison stand for different experiments reported in the same source, numbers behind a study denote multiple comparisons within the same experiment.

Table 1
Results of the Moderator Analyses

Characteristic	Response	Retention
List-based sample Not list-based	OR = 1.19, CI = 0.12 - 1.26, *n* = 26 OR = 1.19, CI = 1.06 - 1.34, *n* = 4 $Q < 0.01$, $p = .94$	OR = 1.28, CI = 1.00 - 1.63, *n* = 18 OR = 1.27, CI = 1.09 - 1.47, *n* = 7 $Q < 0.01$, $p = .97$
Lottery incentive No lottery	OR = 1.19, CI = 0.11 - 1.28, *n* = 25 OR = 1.17, CI = 1.09 - 1.27, *n* = 4 $Q = 0.09$, $p = .77$	OR = 1.26, CI = 1.11 - 1.44, *n* = 23 OR = 1.42, CI = 0.79 - 2.58, *n* = 2 $Q = 0.15$, $p = .70$
Monetary incentive Nonmonetary	OR = 0.98, CI = 0.70 - 1.38, *n* = 2 OR = 1.19, CI = 1.13 - 1.26, *n* = 28 $Q = 1.21$, $p = .27$	OR = 1.26, CI = 1.11 - 1.44, *n* = 23 OR = 1.42, CI = 0.79 - 2.58, *n* = 2 $Q = 0.15$, $p = .70$
Incentive prepaid Incentive on return	OR = 0.94, CI = 0.60 - 1.53, *n* = 1 OR = 1.19, CI = 1.13 - 1.26, *n* = 29 $Q = 1.21$, $p = .27$	OR = 1.35, CI = 0.58 - 3.15, *n* = 1 OR = 1.27, CI = 1.11 - 1.44, *n* = 24 $Q = 0.02$, $p = .88$
Result summary No summary	OR = 1.23, CI = 1.05 - 1.44, *n* = 9 OR = 1.18, CI = 1.12 - 1.25, *n* = 21 $Q = 0.18$, $p = .67$	OR = 1.19, CI = 0.84 - 1.68, *n* = 9 OR = 1.28, CI = 1.12 - 1.47, *n* = 16 $Q = 0.17$, $p = .68$
Nonprofit study Commercial study	OR = 1.20, CI = 1.11 - 1.29, *n* = 27 OR = 1.18, CI = 1.10 - 1.27, *n* = 3 $Q = 0.06$, $p = .81$	OR = 1.35, CI = 1.17 - 1.56, *n* = 24 OR = 0.98, CI = 0.73 - 1.31, *n* = 1 $Q = 3.74$, $p = .053$
No. of study pages	$\beta = .004$, $n = 25$ $Q = 0.74$, $p = .39$	$\beta = -.007$, $n = 25$ $Q = 0.64$, $p = .43$
Year of study	$\beta = .027$, $n = 30$ $Q = 2.64$, $p = .10$	$\beta = -.028$, $n = 25$ $Q = 0.37$, $p = .54$
Payout in lottery	$\beta < -.001$, $n = 25$ $Q = 0.03$, $p = .86$	$\beta < -.001$, $n = 23$ $Q = 0.26$, $p = .61$

Because in incentive experiments all possible outcomes are equally interesting, researchers and publishers should have no reason to submit or publish experiments with non-significant or unexpected results less frequently or less quickly. The Rank Correlation test (Begg & Mazumdar, 1994) between the standardized effect and the inverse of its standard error corroborates this reasoning, *Kendall's* $t = -0.03$, $p = .80$. The Rank Correlation test is combined with the Linear Regression test proposed by Egger, Smith, Schneider, and Minder (1997). The intercept of the regression line ($\alpha = .22$) is not significantly different from zero, $t = 0.82$, $p = .42$. With the outlier comparisons included, the outcome would have been the same: The Rank Correlation test is not significant, *Kendall's* $t = -0.03$, $p = .83$; and the intercept of the regression line ($\alpha = -0.03$) is not significantly different from zero, $t = 0.05$, $p = .96$. It is concluded that no publication or availability bias occurred. However, readers are invited to e-mail study results that were not included in the present review to the author or to fill out the form at http://www.goeritz.net/incentives.htm. Submitted data will be included in a future update and will help in ruling out publication bias as an explanation for the results.

Retention

There was significant heterogeneity among the comparisons that pertained to retention, $Q = 45.58$, $df = 25$, $p = .007$. Inspection of the radial plot in Figure 3 indicates one outlier (i.e., Rager, 2001), which is confirmed by a *z*-value of 4.92, $p < .001$. After exclusion of this outlier, the remaining comparisons are homogenous, $Q = 18.21$, $df = 24$, $p = .79$.

The overall effect for the impact of material incentives on retention is OR = 1.27 (95% CI = 1.12 - 1.44), $d = 0.13$. It is highly significant, $z = 3.66$, $p < .001$. Thus, material incentives increase the odds of a person completing a Web survey by 27% of the odds without incentives (cf. Figure 2, right panel). With the two outlier studies included, the overall effect would have been OR = 1.39 with a 95% CI = 1.15 - 1.68, $z = 3.39$, $p = .001$, $d = 0.18$.

Table 1, right column, shows that none of the study characteristics significantly moderates the incentive effect. In tendency, the incentive effect is larger if a study is commercial rather than if it is a nonprofit study. However, this effect should be taken with caution for two reasons. First, it is based on only one commercial study.

Second, because overall as many as 18 moderator tests were performed, roughly one effect is expected to be significant by chance alone.

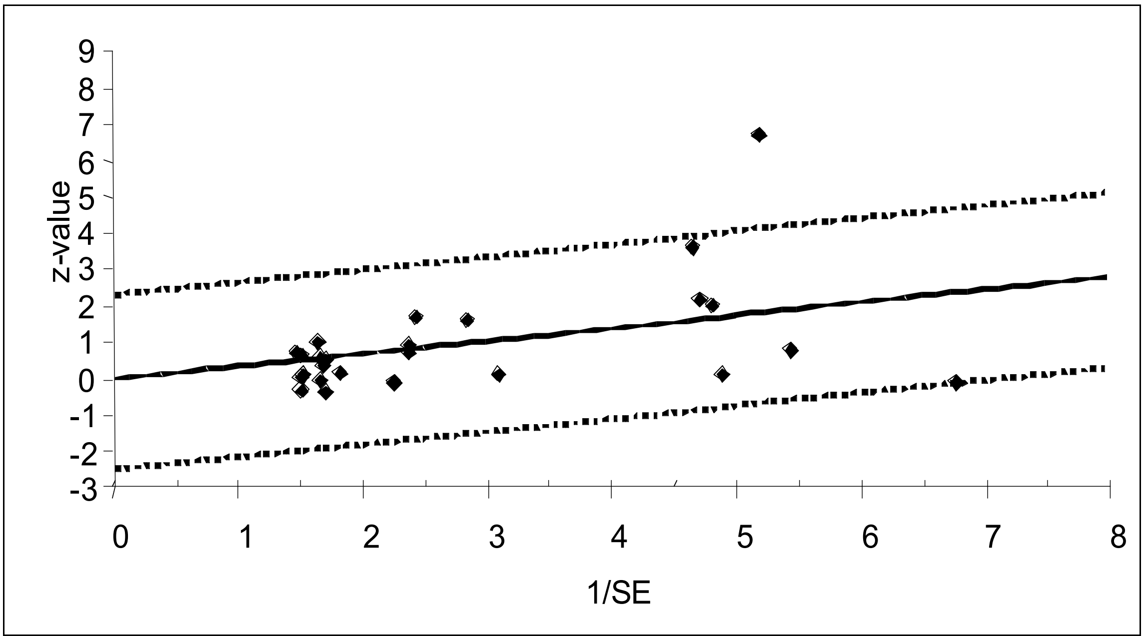

Figure 3. Radial plot of the effect of material incentives on retention.[13]

As regards possible publication bias, the Rank Correlation test provides no hints that a significant number of experiments has been left out, *Kendall's t* = -0.03, $p = .82$. The Linear Regression test ($\alpha = -0.05$) points to the same conclusion, $t = 0.13$, $p = .90$. With the outlier included, the outcome would have been the same: The Rank Correlation test is not significant, *Kendall's t* = -0.03, $p = .81$; and the intercept of the regression line ($\alpha = -0.39$) is not significantly different from zero, $t = 0.74$, $p = .47$.

Discussion

The two meta-analyses reveal that material incentives promote response and retention in Web surveys. Thus, both hypotheses H1 and H2 are supported. The incentive effect on response is OR = 1.19. This means that material incentives increase the odds of a person responding by 19% over the odds without incentives. An OR of 1.19 corresponds to a standardized mean difference between the incentive condition and the no-incentive condition of $d = 0.10$. In terms of absolute percentage differences, in the available studies an incentive increased the response rate on average by 2.8% over no incentive. The incentive effect on retention is OR = 1.27, which corresponds to $d = 0.13$. In the summarized studies, an incentive increased retention by 4.2% on average. In absolute terms, the established effects on both response and retention are small. It took these meta-analyses to reveal them, as most of the individual studies were underpowered to detect the effects (note that most CIs in Figures 1 and 2 include "1").

The obtained ORs of 1.19 and 1.27 allow making predictions about the increase in the completion rate (i.e., the share of a contacted sample who stay until the end of a survey) if incentives are employed. Table 2 lists the *combined* effect of incentives on response and retention for different baseline response and retention rates. As a reading example, imagine a sample of 10,000 people being invited to take a Web survey. From a previous study it can be assumed that the baseline response rate (i.e., the response rate without an incentive) in the new survey would be 40% (i.e., 4,000 people call up the first survey page) and the baseline retention rate 65%. This means that of the contacted sample 2,600 people or 26% would stay until the end of the survey. Table 2 indicates what increase in completion rate can be expected if an incentive rather than no incentive is employed: The cell

[13] Comparisons lying outside the 99% confidence interval of the regression line (dotted lines) are contributing significantly to between-comparison heterogeneity.

pertaining to a baseline response rate of 40% and a baseline retention rate of 65% reads "5.07". Thus in the new survey, 31.07% (26% + 5.07%) of the contacted sample or 3,107 people can be expected to stay until the last survey page if an incentive is offered, compared to 26% or 2,600 people if no incentive is offered. With this figure in mind, one can weigh whether offering an incentive is worthwhile or not by contrasting the gain in respondents against the increase in cost incurred by the incentive itself and its distribution in terms of money and manpower.

Table 2
Differences in completion rates when using an incentive over no incentive listed for different baseline response and retention rates.

Baseline Retention Rate without Incentive \ Baseline Response Rate without Incentive	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
5	0.12	0.23	0.34	0.44	0.53	0.62	0.70	0.77	0.84	0.90	0.96	1.02	1.06	1.11	1.14	1.18	1.21	1.23	1.25
10	0.23	0.44	0.65	0.84	1.01	1.18	1.33	1.47	1.60	1.72	1.83	1.93	2.01	2.09	2.16	2.22	2.27	2.31	2.34
15	0.33	0.64	0.93	1.20	1.45	1.68	1.90	2.10	2.28	2.45	2.60	2.73	2.85	2.96	3.05	3.13	3.19	3.25	3.28
20	0.42	0.81	1.18	1.53	1.84	2.14	2.41	2.66	2.89	3.09	3.28	3.45	3.59	3.72	3.83	3.91	3.99	4.04	4.08
25	0.50	0.97	1.41	1.82	2.20	2.55	2.87	3.16	3.42	3.66	3.87	4.06	4.23	4.37	4.48	4.58	4.65	4.70	4.73
30	0.58	1.12	1.62	2.08	2.51	2.90	3.26	3.59	3.89	4.15	4.39	4.59	4.77	4.91	5.03	5.13	5.19	5.24	5.25
35	0.64	1.24	1.80	2.31	2.78	3.22	3.61	3.97	4.28	4.57	4.82	5.03	5.21	5.36	5.48	5.56	5.62	5.64	5.64
40	0.70	1.35	1.96	2.51	3.02	3.49	3.91	4.28	4.62	4.91	5.17	5.39	5.57	5.71	5.82	5.89	5.93	5.93	5.91
45	0.75	1.45	2.09	2.68	3.22	3.71	4.15	4.54	4.89	5.19	5.45	5.66	5.83	5.97	6.06	6.11	6.13	6.11	6.05
50	0.80	1.53	2.21	2.83	3.39	3.90	4.35	4.75	5.10	5.40	5.65	5.86	6.02	6.13	6.20	6.23	6.22	6.17	6.08
55	0.83	1.60	2.31	2.94	3.52	4.04	4.50	4.90	5.25	5.55	5.79	5.98	6.12	6.21	6.26	6.26	6.21	6.12	5.99
60	0.87	1.66	2.38	3.04	3.62	4.15	4.61	5.01	5.35	5.63	5.86	6.03	6.15	6.21	6.23	6.19	6.11	5.98	5.80
65	0.89	1.70	2.44	3.10	3.69	4.22	4.68	5.07	5.39	5.66	5.86	6.01	6.10	6.13	6.11	6.03	5.91	5.73	5.50
70	0.91	1.73	2.48	3.14	3.74	4.25	4.70	5.08	5.38	5.63	5.81	5.92	5.98	5.97	5.91	5.79	5.61	5.38	5.10
75	0.92	1.75	2.50	3.16	3.75	4.25	4.68	5.04	5.33	5.54	5.69	5.77	5.78	5.74	5.63	5.46	5.23	4.95	4.61
80	0.92	1.76	2.50	3.16	3.73	4.22	4.63	4.96	5.22	5.40	5.51	5.55	5.52	5.43	5.27	5.05	4.76	4.42	4.01
85	0.92	1.75	2.49	3.13	3.69	4.15	4.54	4.84	5.06	5.21	5.28	5.27	5.20	5.05	4.84	4.56	4.21	3.80	3.33
90	0.92	1.74	2.46	3.08	3.62	4.06	4.41	4.68	4.86	4.97	4.99	4.94	4.81	4.61	4.33	3.99	3.58	3.10	2.56
95	0.91	1.71	2.41	3.02	3.52	3.93	4.25	4.48	4.62	4.68	4.65	4.54	4.36	4.10	3.76	3.35	2.87	2.32	1.70

Judging from the small size of the incentive effects, at first glance, incentives seem to be less effective in Web than in offline surveys. For traditional offline surveys (be it CATI, CAPI, face-to-face, mail), definitions for calculating nonresponse are different than those used for online surveys (cf. The American Association for Public Opinion Research, 2000), and therefore it is difficult to compare online and offline surveys. In offline surveys, especially in mail surveys, response and retention are confounded to a large part: People who start filling out a mail survey but do not return it are counted as unit-nonresponders and not as break-offs, as they would be counted in online surveys. In Web surveys, by contrast, it is recorded when people call up a survey page. Therefore a sharper distinction exists among refusing, responding, and retained people. However, an approximate comparison between online and offline studies can be made by combining the incentive effects on response and retention in Web surveys and by contrasting this joint effect with the incentive effect on response in offline studies.

As most studies in the present meta-analyses rely on promised nonmonetary incentives, to compare the online and offline mode as fairly as possible, the effectiveness of only these types of incentives is to be contrasted: For the offline mode, Yu and Cooper (1983) obtained an overall effect of nonmonetary incentives of $d = 0.18$[14] while Church (1993) found an overall effect of promised nonmonetary incentives of $d = 0.02$. In the present meta-analyses, the overall effect on response in the 28 valid comparisons with a promised nonmonetary incentive is OR = 1.19 ($d = 0.10$); and the effect on retention in the 23 relevant comparisons is OR = 1.26 ($d = 0.12$). Combining these effects yields an overall effect of promised nonmonetary incentives in Web surveys OR_{total} = 1.50 (1.19 * 1.26) or $d_{total} = 0.22$ (0.10 + 0.12). Thus, promised nonmonetary incentives seem to work better in online than in offline studies. A possible reason might be that promised nonmonetary incentives have been the most common type of incentive online and Internet users might have come to expect them (Bosnjak & Tuten, 2003). However, the tendency that promised nonmonetary incentives are more effective in online than in offline studies shall only be taken as a starting point for experimental study. On the basis of these meta-analyses, one-to-one mode comparisons are not possible because of simultaneous differences in underlying definitions, procedures, and samples.

These meta-analyses have demonstrated that material incentives increase response and decrease dropout. Therefore, it is generally recommended to use material incentives in Web surveys. However, dependent upon the researcher's aims (e.g., minimizing the cost per participant or reducing nonresponse error) as well as the characteristics of a given survey and its target group, it should be well considered whether incentives are to be employed at all, and if so which are the most economic ones to stimulate desired returns in a particular situation. Especially expensive incentives might not be worth the extra costs because even the combined effect of incentives on response and retention is still small.

With regard to the circumstances under which incentives prove particularly useful, the present meta-analyses have not revealed any significant impact of study characteristics on the effectiveness of incentives. Instead of retaining considerable between-study heterogeneity on which to model moderator effects, with these meta-analyses the heterogeneity was already greatly reduced after exclusion of a few outliers. The nonsignificant moderator tests do not warrant the conclusion that the effects of incentives on response and retention are independent of the examined study characteristics. The comparatively small number of summarized studies might have hindered possible effects to be detected. For the same reason, only a certain number of possible moderators could be taken into account. Other study characteristics not examined here, such as whether a reminder was sent, whether a study was cross-sectional or longitudinal, or the field time of a study, might well have an impact on the magnitude of the incentive effects. Finally, with the continuous moderators *study recency*, *number of study pages*, and *lottery payout* only a limited range of possible values were instantiated in the primary studies. For example, the number of study pages varied between 4 and 25 and the lottery payouts between 48 and 218 USD. While there were only negligible differences *within* these ranges, it might well be that more extreme values of these study characteristics would render their impact on the incentive effect more visible. While one cannot interpret the nonsignificant moderator effects as nonexistent moderator effects, it can however be assumed that the examined study characteristics within their studied ranges do not exert any *large* influence on the incentive effects. This absence of substantial effects has enormous implications for research practice. For example, all other things being equal, a lottery is similarly effective whether 50 or 200 USD are raffled. Moreover, a given incentive is similarly useful if one's questionnaire is 5 or 25 pages long.

Using material incentives is only one option to influence data quality and quantity (cf. Dillman, 1978). We should not forget about other possibly response-enhancing techniques such as personalization, prenotification, deadlines, reminders, offering result summaries or altruistic appeals (cf. Dillman, 2000). Using such measures might be worthwhile because they do not cost much and might alleviate the dangers to data validity inherent in the use of material incentives. While little is known about these measures' effectiveness and their interplay with material incentives in Web-based studies, we know from these meta-analyses that *material incentives* increase response and decrease dropout. Because online surveys are still quite new and the field is changing rapidly with different segments of the population reaching the Internet and technological innovations coming up, the results of the present meta-analyses are not final. A future update will accommodate ongoing developments.

[14] With nonmonetary incentives, they did not differentiate between prepaid and promised incentives.

References[15]

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

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*, 1088-1101.

*Bosnjak, M., & Tuten, T. L. (2003). Prepaid and promised incentives in Web surveys - An experiment. *Social Science Computer Review, 21*, 208-217.

Chang, A. (2001). *Meta-analytical freeware MetaAnalysis 3.0. [WWW page].* Available: http://department.obg.cuhk.edu.hk/researchsupport/download/downloads.asp

Church, A. H. (1993). Estimating the Effect of Incentives on mail survey rates. A Meta-Analysis. *Public Opinion Quarterly, 57*, 62-79.

*Cobanoglu, C., & Cobanoglu, N. (2003). The effect of incentives in Web surveys: Application and ethical considerations. *International Journal of Market Research, 45 (4)*, 1-14.

*Comley, P. (2002). Online survey techniques: Current issues and future trends. *Interactive Marketing, 4*, 156-169.

Cook, C., Heath, F., & Thompson, R. L. (2000). A Meta-Analysis of response rates in Web- or Internet-based surveys. *Educational and Psychological Measurement, 60*, 821-836.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*, 464-494.

Deci, E. L. (1971). The effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*, 105-115.

DerSimonian, R., & Laird, N. M. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*, 177-188.

Dillman, D. A. (1978). Mail and telephone surveys. The total design method. New York: Wiley.

Dillman, D. A. (2000). Mail and Internet surveys. The tailored design method. New York: Wiley.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629-634.

*Enander, J., & Sajti, A. (1999). Online survey of online customers. Value-added market research through data collection on the internet. In *Proceedings of the ESOMAR Worldwide Internet Conference Net Effects 2,* pp. 35-51. London, February 21- 23, 1999.

Fleiss, J. L. (1994). Measures of effect size for categorical data. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis,* pp. 245-60. New York: Russell Sage Foundation.

*Frick, A., Bächtiger, M.-T., & Reips, U.-D. (2001). Financial incentives, personal information and drop-out in online studies. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science,* pp. 209-219. Lengerich: Pabst.

Fox, R. J., Crask, M. R., & Kim, J. (1988). Mail survey response rate - a meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly, 52*, 467-491.

[15] Experiments taken into account in the meta-analyses are designated by an asterisk.

Galbraith, R. F. (1988). A note on graphical presentation of estimated odds ratios from several clinical trials. *Statistics in Medicine, 7*, 889-894.

*Göritz, A. S. (2005). Unpublished experiments on material incentives in list-based Web studies. University of Erlangen-Nürnberg, Germany.

Göritz, A. S., Reinhold, N., & Batinic, B. (2002). Online Panels. In B. Batinic, U.-D. Reips, M. Bosnjak, & A. Werner (Eds.), *Online Social Sciences*, pp. 27-47. Seattle: Hogrefe.

Greenland, S. (1987). Quantitative methods in the review of epidemiological literature. *Epidemiologic Reviews, 9*, 1-30.

Haddock, C. K., Rindskopf, D., & Shadish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods, 3*, 339-353.

Hasselblad, V., & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin, 117*, 167-178.

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

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

Mizes, S. J., Fleece, L E., & Roos, C. (1984). Incentives for increasing return rates: Magnitude levels, response bias, and format. *Public Opinion Quarterly, 48*, 794-800.

*O'Neil, K. M., & Penrod, S. D. (2001). Methodological variables in web-based research that may affect results: Sample type, monetary incentives, and personal information. *Behavior Research Methods, Instruments, & Computers, 33*, 226-233.

*O'Neil, K. M., Penrod, S. D., & Bornstein, B. H. (2003). Webbased Research: Methodological variables' effects on dropout and sample characteristics. *Behavior Research Methods, Instruments, & Computers, 35*, 217-236.

*Porter, S. R., & Whitcomb, M. E. (2003). The impact of lottery incentives on student survey response rates. *Research in Higher Education, 44*, 389-407.

*Rager, M. (2001). Sozialforschung im Internet: Fragebogenuntersuchungen im World Wide Web [Social Research via the Internet: Surveys using the WWW]. Unpublished Diplomarbeit [final thesis], University of Salzburg, Austria.

Simsek, Z., & Veiga J. F. (2001). A primer on internet organizational surveys. *Organizational Research Methods, 4*, 218-235.

Singer, E., Van Hoewyk, J., Gebler, N., Raghunathan, T., & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics, 15*, 217-230.

The American Association for Public Opinion Research (2000). *Standard definitions: Final dispositions of case codes and outcome rates for surveys*. Ann Arbor, Michigan: AAPOR.

*Theobald, A. (2000a). Determinanten des Online Research [Determinants of online research]. *Planung und Analyse, 11*, 72-76.

*Theobald, A. (2000b). Das World Wide Web als Befragungsinstrument. [The WWW as an instrument for surveying people.] Dissertation. Wiesbaden: Deutscher Universitätsverlag.

*Tuten, T. L., Bosnjak, M., & Bandilla, W. (2000). Banner-advertised web surveys. *Marketing Research, 11*, 17-21.

*Tuten, T. L., Galesic, M., & Bosnjak, M. (2004). Effects of immediate versus delayed notification of prize draw results on response behavior in Web surveys – An experiment. *Social Science Computer Review, 22 (3)*, 377-384.

Yu, J., & Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research, 20*, 36-44.