Translational Issues in Psychological Science
2016, Vol. 2, No. 2, 163–173

2332-2136/16/$12.00 http://dx.doi.org/10.1037/tps0000071

The Longitudinal Effects of Incentives on Response Quantity in Online Panels

Anja S. Göritz
University of Freiburg

Benedikt P. Neumann
University of Würzburg

With the widespread use of online questionnaires, motivating people to participate in web-based studies has become crucial to research success. Much of web-based research is conducted using online panels. This experiment examined the long-term effect of 6 rewards on response quantity in an online panel. All 3,200 newly recruited members of a noncommercial online panel were invited to 7 study waves. One half of the participants were offered a result summary in each wave; the other half was not. In addition, 1 of 6 rewards was promised for participation in each wave: no reward, money via bank transfer, money via PayPal, redeemable loyalty points, donation to a predetermined charity, or donation to a selectable charity. Measures of response quantity were attrition, response, and retention. Latent growth curve models revealed incentive effects on the intercept (i.e., baseline in Wave 1) and on the slope (i.e., change over the waves). Neither incentive influenced attrition and retention. PayPal and donations lowered baseline response. Result summary reduced the slope of response over the 7 waves. To maximize response quantity, researchers and practitioners should avoid using nonmonetary incentives such as donations and result summaries. When considering incentive costs, the findings call into question the usefulness of offering any incentive to boost response quantity, at least in noncommercial online panels. The present research has covered much ground in that it contrasted 6 rewards over time and crossed the rewards with the (non)offer of a summary. Explaining these differences remains to future experiments that trade breadth for depth.

Keywords: incentives, result summary, web-based data collection, longitudinal experiment, response quantity

Supplemental materials: http://dx.doi.org/10.1037/tps0000071.supp

Motivating people to respond to a study request and motivating respondents to stay in the study until its end are key to the quality of resulting data. Data quality is higher in the absence of both nonresponse error and measurement error (Dillman & Bowker, 2001; Groves, 1987). Failure at motivating people to start or to finish a study potentially entails nonresponse error because the loss of sample units at different stages may bias estimates of population parameters, particularly if the loss of sample units occurs systematically and is not controlled for in analysis (Groves & Peytcheva, 2008).

Preventing nonresponse is hard work for researchers and survey practitioners. In recent decades, this task has become even more difficult because of the abundance of academic and commercial surveys, a change in cultural norms of cooperation (Porter, 2004, p. 5), and the fact that studies are increasingly conducted via the web rather than in traditional survey modes such as telephone, mail, or face-to-face. Thus, although web-based data collection bears many

Anja S. Göritz, Occupational and Consumer Psychology, University of Freiburg; Benedikt P. Neumann, Work and Organizational Psychology, University of Würzburg.

This work was supported by DFG Grant GO 1107/4-1 to Anja S. Göritz. We thank Olaf Lahl, who helped in devising and programming the questionnaire in Waves 1 and 2; Peter Caprariello, who helped in devising the questionnaire in Wave 3; Edith de Leeuw, who helped in devising the questionnaire in Wave 4; and Jochen Gebauer, who devised and programmed the questionnaire in Waves 5, 6, and 7.

Correspondence concerning this article should be addressed to Anja S. Göritz, Occupational and Consumer Psychology, University of Freiburg, Engelbergerstr. 41, 79085 Freiburg, Germany. E-mail: goeritz@psychologie.uni-freiburg.de

economic and methodological advantages for conducting research (Göritz, Wolff, & Goldstein, 2008), it is not immune to nonresponse. The danger of nonresponse looms even larger on the web, perhaps because of the greater ease of refusing a survey request that is delivered by a computer rather than a human being. Manfreda, Bošnjak, Berzelak, Haas, and Vehovar (2008) meta-analyzed 45 experimental comparisons between web and other survey modes and found that, on average, web surveys yield an 11% lower response rate compared to other modes.

Web-based data collection often takes place in online panels (Batinic & Moser, 2005; Callegaro et al., 2014). Online panels are pools of people who have agreed to occasionally take part in web-based studies. Online panels are distinct from crowdsources such as Mechanical Turk, which comprise people who have agreed to occasionally carry out all kinds of work. Crowdsources are beset with problems when used to conduct studies (Peer, Vosgerau, & Acquisti, 2014; details are beyond the scope of this article). Göritz and Wolff (2007) discussed several benefits of online panels: Participants must be recruited only once and subsequently are readily available for individual studies. Panelists' profile data and historical data are available and need not be collected repeatedly. Consequently, time and money are saved, respondent burden is reduced, and a wide range of information is accessible for substantive analyses and for quality assurance.

Whether panelists respond to a study request and how long they stay in the study depends on their motivation (Krosnick, 1991). Incentives attempt to foster motivation by enhancing participants' benefit of participation (Göritz, 2010, 2014b). For practical reasons, in web-based data collection, incentives are usually issued after study completion (Göritz, 2014b), that is, they are postpaid. There are monetary and nonmonetary postpaid incentives: On the web, monetary rewards can be implemented in at least three ways: (a) An intermediary such as PayPal can be used to effect a payment (Bošnjak & Tuten, 2003; Göritz et al., 2008), (b) money can be transferred into a respondent's bank account, and (c) loyalty points—an artificial currency that can be redeemed for money after having collected a particular number of points—may be offered to respondents (Batinic & Moser, 2005; Göritz, 2008). With regard to the effectiveness of monetary rewards in offline studies, in a review of 10 meta-analyses, Göritz (2014b) found that postpaid monetary incentives slightly increase the response rate, namely by six to eight percentage points. In online studies, we found the following effects:

1. Payments via PayPal had no effect on response and retention in a cross-sectional study (Bošnjak & Tuten, 2003). Göritz and colleagues (2008) found that use of PayPal diminished response in Wave 1 of a longitudinal experiment; there was a response-enhancing effect in Wave 2 but no influence in Wave 3. However, the null effect in Wave 3 should be interpreted with caution because participants had not received their payment for Wave 2 when they were invited to Wave 3. To summarize existing research on payments via PayPal, when PayPal was used for the first time, it tended to have a neutral or a diminishing influence on response, whereas in later waves, it tended to have a neutral or enhancing influence on response and was generally neutral to or enhancing retention.
2. Some web-based studies have examined loyalty points. In a review of 65 uncontrolled online studies, Batinic and Moser (2005) found a response-enhancing effect compared to no incentive. However, this analysis was not based on experimental studies, and it did not account for longitudinal effects. In a cross-sectional experiment (Göritz, 2004), loyalty points increased response and retention more than did a cash lottery, but there was no no-incentive condition. In a longitudinal experiment (Göritz, 2008), panelists who were offered loyalty points responded to the first wave as frequently as panelists who were offered a cash lottery. However, the attractiveness of loyalty points increased compared with the cash lottery over the five waves of the study. Because there was no no-incentive condition, the absolute effect of loyalty points is unknown.
3. We did not find any study that examined payments via bank transfer.

Researchers and practitioners can provide nonmonetary incentives rather than monetary rewards. The present study examined two types of nonmonetary incentives. First, the panel offered to donate to charity in exchange for panelists' study participation. This type of reward can be implemented in two ways: With a fixed charity donation, the panel preselects a charity, and with a selectable charity donation, panelists are given a choice among several charities. Few studies have investigated the effectiveness of donations in offline studies. In eight studies, donations to a charity yielded a lower response rate compared to 31 studies that used monetary rewards (Engel et al., 2004). We did not find any published study on the effectiveness of donations in a web-based study.

Second, panelists were offered the results of the study in which they participate. This type of nonmonetary incentive is economical for the panel operator because a summary can be composed quickly, and sending it via e-mail does not incur noteworthy costs. Examining the effectiveness of result summaries in offline studies (Edwards et al., 2009) shows that they somewhat decrease response (odds ratio [OR] = 0.90). With regard to online panels, in a review of uncontrolled studies (Batinic & Moser, 2005), studies that offered results elicited lower response than did studies without a result summary. In a review of 204 studies that were conducted in an online panel, 71 studies that offered a result summary did not obtain a different response and retention rate from 133 studies without a result summary (Göritz, 2014a). With regard to experimental evidence, offering a result summary decreased response in a commercial online panel (Göritz & Luthe, 2013b). By contrast, in Göritz and Luthe's (2013a) Experiment 2, which was conducted in a nonprofit panel, a result summary had no effect. Likewise, in Scherpenzeel and Toepoel (2014), various implementations of study results had no or inconsistent effects on participation behavior in a probabilistic online panel. Unless study results are tailored to the respondent (Marcus, Bošnjak, Lindner, Pilischenko, & Schütz, 2007) and the study topic is salient (Tuten, Galešić, & Bošnjak, 2004), study results seem to have no effect or a deleterious effect. Last, a result summary can be combined with other rewards. No research that has tested the combination of a result summary with donations, loyalty points, or money was found.

Because incentives are costly and high-quality data are essential for successful research, it is important to maximize incentive effects. The present study aimed to extend the knowledge about incentive effects in online panels in six directions: First, we conducted an experiment with high internal validity. Second, we conducted this experiment in a natural environment (i.e., in the field), thus ensuring high external validity of findings. Third, as many as six different rewards were examined; some of them were tested for the first time in the context of web-based data collection. We examined the following: no reward, payment via bank transfer, payment via PayPal, loyalty points, donation to a charity that was predetermined by the panel, and donation to a charity that respondents could choose from among three charities. Fourth, the six rewards were systematically crossed with the offer or nonoffer of a result summary, thus revealing potential interaction effects besides main effects. Fifth, the effect of these incentives was assessed in a comprehensive manner, specifically on three measures of response quantity. Dependent measures were attrition, response, and retention. Sixth, this research was implemented as a seven-wave longitudinal experiment to shed light on immediate and long-term incentive effects.

Method

Participants

A total of 3,200 panelists were newly recruited into a noncommercial, university-based opt-in online panel. Various methods of recruitment such as banners, links on other websites, postings in newsgroups and mailing lists, and word-of-mouth were employed. At panelists' sign-up, no specific expectations were roused concerning what type and amount of incentives would be offered for participation in the studies run in this panel. The sample was 60.8% women. Mean age was 41.5 years (SD = 14.9). Most participants were German (97.8%). In terms of education, 29.7% held a university degree. Half of the participants were employed (54.2%).

Measurement and Design

The study was based on a 7 × 2 × 6 factorial design, with the first factor (7) being the number of waves. The second factor (2) was the (non)offer of a result summary: Half of the participants were offered a summary in each wave, and half of the participants were not offered a summary. The third factor (6) was type of reward: Participants received one of the following: no reward, money via bank transfer, money via PayPal, loyalty points that were redeemable against money, the possibility to donate to a charity that had been preselected by the panel (i.e., Amnesty International), or the possibility to donate to a charity that panelists could select from among three renowned charities (i.e., Amnesty International, World Wide Fund for Nature, or Cancer Aid). Panelists were randomly assigned to one of the 12 (2 × 6) incentive conditions and remained in this condition throughout all seven waves. The reward amount varied between 1 and 2 euros (US$1.30 and $2.60) among waves according to survey length but was the same for all types of rewards within any given wave. PayPal is a well-known commercial intermediary for web-based money transfers. For private individuals, there is no cost to open a PayPal account to receive small amounts of money such as were paid within this study. Loyalty points could be exchanged for real money (5 euros [US$6.50]) but only after having collected 50 points. Introducing a threshold for redeeming loyalty points that was not yet reached after a participant had taken part in the first wave made loyalty points differ from the other rewards more than was absolutely necessary. However, we implemented such a threshold in the interests of ecological validity because most panels that rely on loyalty points have a threshold.

There were three dependent variables: (1) *Attrition* refers to the number of panelists who left the panel for good between any study and a subsequent study divided by the number of panelists present in the panel at the end of the previous study. (2) *Response* refers to the number of panelists who accessed the first page of a study divided by the number of panelists who were invited to this study. (3) *Retention* is the number of panelists who reached the final page of a study divided by the number of panelists who accessed the first page of that study. In the following sections, an invitee who loaded at least the first page of a study is called a *responder,* whereas a responder who stayed until a study's final page is called a *retainee*.

Procedure

In each wave, an invitation e-mail that outlined length and topic of the current study was sent to the panelists. Apart from a variable text segment about the incentive, the e-mail was identical for all participants within each wave. The wording of the incentive information was kept constant across waves except for altering the amount of the reward. For instance, the text segment in the PayPal condition read: "In appreciation of your efforts, we will pay you a reward of 1 euro (US$1.30) via the online payment service PayPal for your participation in this study." In the summary condition, the following sentence was added: "In case you are interested, you will receive a summary of the findings of this study as soon as we have analyzed the data." Thus, participants were told prior to their potential participation what and how much they would earn. Note that the panelists were not told that they would participate in a longitudinal study. Participants who were not retained in any of the waves were free to participate again in later waves.

The questionnaire in each wave started with a welcome screen that repeated the topic and length of the study. Content, number of pages, and number of questions varied across waves. The announced completion times ranged from 7.5 to 20 min. According to expected completion time as well as tediousness, the amount of the promised reward varied from 1 to 2 euro (US$1.30–2.60) across waves but was the same within a wave. Thus, we paid roughly 6 euros (US$7.80) per hour, which was slightly lower than the minimum wage, but it was tax free and participation was not dangerous, dirty, or physically demanding and could be done at participants' own discretion with regard to time and place. The interval between any two waves varied from 1 to 3 months. The field time of each wave varied from 1 to 2 weeks (see Table 1). It is typical that in online panels the between-studies interval and the studies' topic, length, amount of reward, and field time vary. The instances of these characteristics as realized in

Table 1
Overview of Waves: Topic, Length, Amount of Reward, Field Time

Wave	Topic	Length (min)	Amount in euros (US$)	Field time (days)
1	Occupational attitudes	20	2.00 ($2.60)	14
2	Appraisal of situations at work	15	1.50 ($1.95)	13
3	Buying behavior and recreational activities	15	1.00 ($1.30)	10
4	Internet use	12.5	1.00 ($1.30)	8
5	Self–other perception and well-being	10	1.50 ($1.95)	10
6	Self–other perception and well-being	7.5	1.50 ($1.95)	8
7	Self–other perception and well-being	10	2.00 ($2.60)	11

this research were typical for online panels as well.

Incentives were paid out within 2 weeks after each wave had been closed. Payments via PayPal and loyalty points as well as donations were made within 3 days after each wave, whereas bank transfers and result summaries took a few days longer. Payments via PayPal, bank transfer and loyalty points were effected without any further notification by the panel. In the case of PayPal payments, participants received an e-mail from PayPal that they had received the given amount in their PayPal account. In the case of bank transfers, participants could see by their bank statement that the money had been transferred into their account. In the case of loyalty points, participants could log into their panel profile and see their credit, and once their credit exceeded the threshold of 5 euros (US$6.50), they could redeem the credit for money that would be transferred into their bank account. Participants in the donation conditions received an e-mail reporting that the money had been donated to the charity to which they had contributed. A link to a web page that showed a proof of donation issued by the charity was included. Participants in the summary condition who in the questionnaire had requested to be informed about the study's results received an e-mail that contained a link to a web page that contained a general summary of the results of the study (i.e., it was not tailored to a participant's individual answers in the study but was an overview of aggregated results) that was one to three pages long.

Analyses

Data were analyzed using latent growth curve models (McArdle, 1986; Meredith & Tisak, 1990; Stoel, van den Wittenboer, & Hox, 2003). In the study at hand, the pattern of change in the dependent variables over time (i.e., trajectory) was of interest (e.g., linear growth, quadratic growth), as were whether the independent variables (i.e., the incentives) would affect the initial level of the dependent variables (i.e., intercept) and whether the incentives would affect change in the dependent variables over time (i.e., slope). Latent growth curve models are appropriate to analyze these effects and take the dependencies in longitudinal designs into account (Duncan & Duncan, 2004).

In latent growth curve modeling, the trajectory of each dependent variable over the seven waves is represented by two latent random factors: (1) The random intercept factor captures the initial level of the dependent variable, whereas (2) the random slope factor captures the change of the dependent variable over the waves. To link these two latent factors to the manifest dependent variables of the seven waves, the factor loadings of the seven dependent variables on the intercept factor were constrained to 1, and the factor loadings of the dependent variables on the slope factor were constrained according to the growth function in question. For instance, when testing for a linear slope, the factor loadings (time scores) were constrained to equally spaced integers ranging from 0 to 6. For a quadratic slope, these factor loadings would be squared. To test for incentive effects, the independent variables were introduced as time-invariant covariates in the latent growth curve model. That is, the two latent factors—intercept and slope—were regressed on the independent variables.

Because the independent variables were discrete, they were dummy-coded to include them

as predictors (Eid, Gollwitzer, & Schmitt, 2011, p. 648 ff.). That is, for each level of the variable reward except for the no-reward condition, a new variable was created and coded "1" if the participant received this type of reward and "0" if not. Similarly, a new variable was created for each interaction between each level of reward and result summary. For instance, the variable Bank Transfer × Result Summary was coded "1" if the participant received both incentives and "0" if not. Because the result summary variable was binary in the first place, it was not recoded. In total, 11 binary predictor (independent) variables resulted. The control condition (no reward and no result summary) had a value of zero on each of the dummy variables. Thus, the control condition acted as a neutral and consistent standard of comparison against which the different incentive conditions were tested.

Mplus Version 6 (Muthén & Muthén, 2007) was used for all analyses. Models were estimated with the weighted least squares mean and variance adjusted (WLSMV) estimator and a probit link function. Missing values were deleted pairwise. For model evaluation, we used the chi-square test statistic, the comparative fit index (CFI; Bentler & Dudgeon, 1996), the root-mean-error of approximation (RMSEA; Hu & Bentler, 1999), and the weighted root-mean-square residual (WRMR; Muthén & Muthén, 2007). Good model fit is indicated by chi-square $p > .05$, CFI $> .96$ with CFI, RMSEA $< .05$, and WRMR $< .95$ (Yu, 2002).

We additionally looked at, as a sensitivity analysis, the results in the final wave relative to randomization groups as inspired by intent-to-treat analysis (Little & Yau, 1996). That is, cell frequencies observed in Wave 7 were compared to corresponding frequencies after participants' assignment to the cells of the experiment prior to Wave 1.

Results

Attrition

Attrition was low across the seven waves of the study. Overall, 117 panelists (3.7%) quit the panel. After seven waves, the largest cumulative attrition had occurred with selectable charity donation (5.6%) and the least attrition with PayPal (2.8%; see the online supplemental materials).

The latent growth curve model with a linear slope fit the data well, $\chi^2(55, N = 3{,}200) = 72.23$, $p = .0594$, CFI = 0.995, RMSEA = .010, WRMR = .313. Note that the attrition model comprises only six time points because attrition can occur only between waves. In exploring the research questions, there was no effect of any independent variable on the intercept of attrition: The baseline, that is, attrition between Wave 1 and Wave 2, was the same with all incentives. Furthermore, none of the independent variables had an effect on the slope. There were marked differences between certain incentives (e.g., attrition after all seven waves in the selectable charity donation condition was twice as large as in the PayPal condition). However, because these differences were not expected, we abstained from performing contrast analyses.

Response

Overall, response dropped from 63.7% in Wave 1 to 46.5% in Wave 7. Testing the fit of the latent growth curve model using a linear slope and introducing all independent variables as time-independent covariates revealed a significant chi-square coefficient, $\chi^2(70, N = 3{,}200) = 106.32$, $p = .003$. However, because the chi-square test is sensitive to large samples, which was the case in the study at hand, and because the other fit indices indicated good model fit (CFI = 0.995, RMSEA = .013, WRMR = 0.872), we considered the model fit acceptable.

With regard to the latent intercept, offering payment via PayPal lowered baseline response ($Z = -0.42$, $p = .026$). Bank transfer and loyalty points did not affect baseline response. By contrast, fixed charity donation ($Z = -0.69$, $p < .001$) and selectable charity donation ($Z = -0.66$, $p = .001$) reduced baseline response. Result summary lowered baseline response, but this effect did not reach a conventional level of significance. There was no interaction effect between result summary and bank transfer, PayPal, loyalty points, or fixed charity donation on response in Wave 1. However, a significant interaction effect occurred between selectable charity donation and result summary on baseline response ($Z = 0.53$, $p = .041$). This indi-

cates that the response-diminishing effect of the combination of these two incentives is weaker than the sum of the response-diminishing effects of the two incentives.

With regard to the slope of response, we found nonsignificant effects of all rewards and their interactions with result summary (see Figure 1). The only incentive that significantly affected the slope was result summary ($Z = -0.24$, $p = .033$); result summary reduced response in the long term.

Retention

Retention was high in all waves of the study (mean = 87.2%). It was lowest in Wave 3 (75.2%) and highest in Wave 6 (91.8%). The latent growth curve model using a linear slope fit the data well, $\chi^2(72, N = 2{,}810) = 90.17$, $p = .073$, CFI = 0.957, RMSEA = .009, WRMR = .917. Neither the incentives nor their interaction affected retention at baseline and over time.

Sensitivity Analysis

We analyzed how many participants of the original sample ($N = 3{,}200$) were no longer in the panel by Wave 7 as a function of reward, summary, and the Reward × Summary interaction by means of stepwise logistic regression analysis (forward: likelihood ratio). Neither reward nor summary nor their interaction influenced attrition.

Next, we analyzed how the original sample had participated in Wave 7. There was no Reward × Summary interaction and no main effect of summary on response, so these terms were removed from the model. However, reward was significant. To locate the source(s) of the effect, we calculated indicator contrasts with *no reward* as a reference category (41.0% response in Wave 7 relative to invitees in Wave 1). Bank transfer (51.0%, OR = 1.50), PayPal (49.9%, OR = 1.44), and loyalty points (50.0%, OR = 1.44) increased response over no reward.

We abstained from analyzing how many participants of the original sample had been retained in Wave 7, because response and retention in Wave 7 correlated at $r = .90$. In a wave-by-wave analysis as with latent growth curve modeling, response and retention were conceptually distinct; however, in this summative analysis the number of people who finished Wave 7 was almost the same as the number of people who responded to Wave 7 compared to all invitees; there was conceptual overlap because finishing a study presupposes having started the study.

This summative way of analyzing the data, which does not distinguish the dependent variables' baseline and trajectory over the waves (i.e., as captured in the latent growth curves)

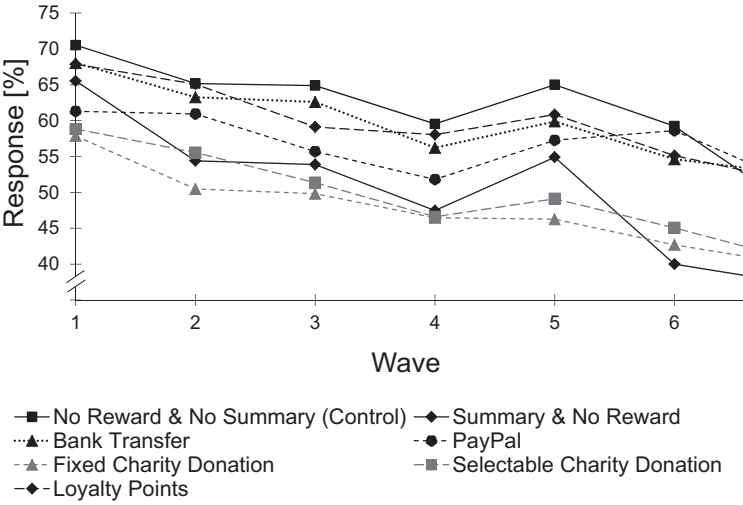

Figure 1. Response by reward and summary.

reveals results that are similar to those for more complex analytical methods, and thanks to its higher power, it brings out more clearly a small participation-enhancing effect of postpaid monetary rewards.

Discussion

This longitudinal experiment examined the effectiveness of three monetary and two nonmonetary rewards and the offer of a result summary to enhance response quantity in online panels as indicated by attrition, response, and retention.

According to the outcomes of fine-grained analytic methods (i.e., latent growth curve modeling), postpaid monetary rewards (i.e., bank transfer, PayPal, and loyalty points) did not reliably foster panelists' willingness to remain a member of the panel, to respond to study invitations, and to finish the studies, compared with those who did not receive any reward. However, the summative picture in the final wave relative to participants' initial randomization showed a mild participation-enhancing effect of the postpaid monetary rewards: Compared with a person who had been assigned to the no reward group, a person who had been assigned to receive monetary rewards for study participation was one and a half times more likely to take part in the seventh study that was run in this panel.

Monetary rewards were somewhat more attractive than no reward in the final wave (i.e., Wave 7), perhaps because panelists were now familiar with them, thus eclipsing initial disadvantages such as (a) the process of receiving the rewards was unfamiliar, (b) the threshold for redeeming the loyalty points was reached after Wave 4 at the earliest, (c) receiving bank transfers requires disclosing one's bank account details, and (d) receiving a payment via PayPal requires opening a PayPal account or at least disclosing one's existing PayPal account. To the extent that previous longitudinal studies have examined long-term effects of monetary rewards, at least in part we found corroborating evidence (Göritz, 2008; Göritz et al., 2008).

The magnitude, form, or immediacy of monetary rewards in this study might explain that they were only slightly better than no reward. Monetary incentives that are contingent on study completion are usually perceived as payment and as such invoke an economic exchange (Dillman, 2007; Porter, 2004). To be effective, the monetary value of any postpaid reward should at least equal the perceived burden of answering the questionnaire. In our case, participants were offered 1 euro (US$1.30) for approximately 10 min of their time. Monetary rewards in this study might have failed to substantially increase response quantity because panelists did not perceive the offered reward as a commensurate payment for their perceived burden of participation (Bošnjak & Tuten, 2003; Church, 1993; Dillman, 2007, p. 14; Edwards et al., 2002; Yu & Cooper, 1983). In addition to the moderate amount that was paid, the monetary rewards were not money in hand (Bošnjak & Tuten, 2003) but were paid out electronically. This might alternatively or additionally explain their small effect. Furthermore, assuming that many panelists were intrinsically motivated (e.g., wanting to find out information about themselves, contributing to topics that help others, contributing to science as a whole), the monetary rewards might have crowded out this intrinsic motivation (Deci & Ryan, 1985). The meta-analysis by Cameron, Banko, and Pierce (2001) revealed reduced motivation to perform initially high-interest tasks when being offered a tangible rather than verbal reward that is loosely tied to the level of performance—all of these conditions of an overjustification effect were met in the study at hand. Finally, the monetary rewards were not immediately delivered, such as cash in hand upon participation, but delivered after an unavoidable delay. All rewards were paid out at the end of the field period collectively; that is, participants who had taken part at the beginning of the field period were not paid before the study closed (see Table 1). In addition to this delay, which was constant across rewards, there was an extra delay that differed among rewards: Payments via PayPal and loyalty points were effected within 3 days after the study had closed, but bank transfer took a few days longer.

PayPal, unlike the other two monetary rewards, significantly lowered response in Wave 1. The finding that PayPal payments initially reduce response confirms the finding of Göritz and colleagues (2008) and thus can be considered robust. Receiving payments via PayPal requires disclosing an e-mail address. Divulging this private information on the web may be perceived as risky. People might disapprove of

the necessity to open an account with PayPal to collect the reward. Moreover, panelists might resent that the noncommercial panel was endorsing PayPal, a commercial intermediary.

Any interpretations of effects that pertain to attrition and retention are to be taken with caution because attrition was generally low, and retention was generally high; hence, bottom or ceiling effects might have left insufficient room for an effect of the rewards. On a practical note, when considering cost, the results challenge the usefulness of employing monetary rewards to increase response quantity in online panels.

Concerning fixed charity donation and selectable charity donation, in the absence of previous results from web-based studies, we merely explored donations' influence on response quantity. However, to no surprise when looking at offline studies (Engel et al., 2004; Hubbard & Little, 1988; Warriner, Goyder, Gjertsen, Hohner, & McSpurren, 1996), both donation types reduced response in Wave 1 compared to no reward and were not able to make up for this initial difference over time. There are a number of possible explanations for why donations are counterproductive. Compared with being offered no incentive, a donation might suggest to panelists that they ought to receive a payment for their efforts because it becomes obvious that the panel has financial resources and is willing to spend them on enhancing cooperation. Invoking equity considerations, panelists might feel that they themselves should receive the money instead of other beneficiaries because panelists carried the burden. Moreover, donation-type rewards might overstrain panelists' readiness to help others, which being reminded of through the request to donate might arouse an adverse feeling of shame. Furthermore, panelists might fear being asked to donate repeatedly once they have gotten involved (i.e., they give an inch and are afraid of being asked to give a yard). This might result in reactance (Brehm & Brehm, 1981)—a reaction of withdrawal when one's freedom or choice is perceived as threatened. To summarize on donations, not only are they counterproductive but they cost money. Therefore, their use is not recommended.

Examining result summary, we found that its null effects on attrition and retention might at least in part be due to bottom or ceiling effects. With regard to response, however, a result summary somewhat reduced the baseline response, and it significantly reduced long-term response, which corresponds to Batinic and Moser (2005) and Göritz and Luthe (2013b). Moreover, this finding corresponds to results from offline studies (Edwards et al., 2009) and thus is to be considered a stable effect. As to potential mechanisms, first, offering a summary highlights the topic of a study and the importance of participating in it. Thus, panelists who are offered a summary compared to panelists who are not offered a summary might more likely find that they are not interested in the topic or that they do not want to take the responsibility of participating in a serious manner. Second, summaries thwart panelists' desire to quickly complete the interaction with the survey. As long as they must wait for the summary, they cannot mentally put aside the matter, preventing the pleasant feeling of closure (Zeigarnik, 1927). Finally, panelists might dislike the prospect that their inbox is cluttered with even more inessential e-mails. With the exception of selectable charity donation, the result summary did not interact with any of the rewards: If a result summary was offered alongside a donation to a selectable charity, the negative effect of a selectable charity donation on baseline response was mitigated. On a practical note, we do not recommend offering summaries as incentives in online panels.

The work at hand is strong because it is a field experiment with random assignment to many conditions conducted on a large and heterogeneous sample spanning seven waves. These properties increase the likelihood that the findings have high internal and external validity. However, as with any research, many questions remain unanswered. For example, this work is on response *quantity* as reflected in attrition, response, and retention. The effect of incentives on response *quality* is the topic of another article, that is, Göritz and Neumann (2016). Also, whether and how the incentives' effectiveness changes if more than seven waves are conducted remains an open question. Moreover, incentives other than those examined in this experiment should be tested. This might entail simple modifications such as different amounts or result summaries tailored to individual participants rather than general summaries. However, perhaps even after refining some of the incentives, noncommercial online panels might fare best in the long term if they simply

allow panelists' helpfulness and curiosity free rein and otherwise treat panelists impeccably.

References

Batinic, B., & Moser, K. (2005). Determinanten der Rücklaufquote in Online-Panels [Determinants of response rates in online panels]. *Zeitschrift für Medienpsychologie, 17,* 64–74. http://dx.doi.org/10.1026/1617-6383.17.2.64

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual Review of Psychology, 47,* 563–592. http://dx.doi.org/10.1146/annurev.psych.47.1.563

Bošnjak, M., & Tuten, T. L. (2003). Prepaid and promised incentives in web surveys: An experiment. *Social Science Computer Review, 21,* 208–217. http://dx.doi.org/10.1177/0894439303021002006

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York, NY: Academic Press.

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (Eds.). (2014). *Online panel research: A data quality perspective*. Chichester, United Kingdom: Wiley. http://dx.doi.org/10.1002/9781118763520

Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *Behavior Analyst, 24,* 1–44.

Church, A. H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly, 57,* 62–79. http://dx.doi.org/10.1086/269355

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press. http://dx.doi.org/10.1007/978-1-4899-2271-7

Dillman, D. A. (2007). *Mail and Internet surveys: The tailored design method* (2nd ed.). Hoboken, NJ: Wiley.

Dillman, D. A., & Bowker, D. K. (2001). The web questionnaire challenge to survey methodologists. In U. D. Reips & M. Bošnjak (Eds.), *Dimensions of Internet science* (pp. 159–178). Lengerich, Germany: Pabst.

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy, 35,* 333–363. http://dx.doi.org/10.1016/S0005-7894(04)80042-X

Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: Systematic review. *BMJ: British Medical Journal, 324,* 1183–1185. http://dx.doi.org/10.1136/bmj.324.7347.1183

Edwards, P. J., Roberts, I., Clarke, M. J., Diguiseppi, C., Wentz, R., Kwan, I., . . . Pratap, S. (2009). Methods to increase response to postal and electronic questionnaires. *Cochrane Database of Systematic Reviews, 2009* (3): MR000008. http://dx.doi.org/10.1002/14651858.MR000008.pub4

Eid, M., Gollwitzer, M., & Schmitt, M. (2011). *Statistik und Forschungsmethoden* [Statistics and research methods] (2nd ed.). Weinheim, Germany: Beltz.

Engel, U., Poetschke, M., Schnabel, C., & Simonson, J. (2004). *Nonresponse und Stichprobenqualitaet. Ausschoepfung in Umfragen der Markt- und Sozialforschung* [Nonresponse and sample quality. Response rate in market and social research surveys]. Frankfurt am Main, Germany: Deutscher Fachverlag HORIZONT productions.

Göritz, A. S. (2004). The impact of material incentives on response quantity, response quality, sample composition, survey outcome, and cost in online access panels. *International Journal of Market Research, 46,* 327–345.

Göritz, A. S. (2008). The long-term effect of material incentives on participation in online panels. *Field Methods, 20,* 211–225. http://dx.doi.org/10.1177/1525822X08317069

Göritz, A. S. (2010). Using lotteries, loyalty points, and other incentives to increase participant response and completion. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 219–233). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/12076-014

Göritz, A. S. (2014a). Determinants of the starting rate and the completion rate in online panel studies. In M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, & P. J. Lavrakas (Eds.), *Online panel research: A data quality perspective* (pp. 154–170). Chichester, United Kingdom: Wiley.

Göritz, A. S. (2014b). Incentive effects. In U. Engel, B. Jann, P. Lynn, A. Scherpenzeel, & P. Sturgis (Eds.), *Improving survey methods: Lessons from recent research* (pp. 339–350). London, United Kingdom: Taylor & Francis.

Göritz, A. S., & Luthe, S. C. (2013a). How do lotteries and study results influence response behavior in online panels? *Social Science Computer Review, 31,* 371–385. http://dx.doi.org/10.1177/0894439312458760

Göritz, A. S., & Luthe, S. C. (2013b). Lotteries and study results in market research online panels. *International Journal of Market Research, 55,* 611–626. http://dx.doi.org/10.2501/IJMR-2013-016

Göritz, A. S., & Neumann, B. P. (2016). *The longitudinal effects of incentives on response quality in online panels*. Manuscript in preparation.

Göritz, A. S., & Wolff, H. G. (2007). Lotteries as incentives in longitudinal web studies. *Social Science Computer Review, 25,* 99–110. http://dx.doi.org/10.1177/0894439306292268

Göritz, A. S., Wolff, H. G., & Goldstein, D. G. (2008). Individual payments as a longer-term incentive in online panels. *Behavior Research Methods, 40,* 1144–1149. http://dx.doi.org/10.3758/BRM.40.4.1144

Groves, R. M. (1987). Research on survey data quality. *Public Opinion Quarterly, 51,* S156–S172. http://dx.doi.org/10.1093/poq/51.4_PART_2.S156

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly, 72,* 167–189. http://dx.doi.org/10.1093/poq/nfn011

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Hubbard, R., & Little, E. L. (1988). Promised contributions to charity and mail survey responses. *Public Opinion Quarterly, 52,* 223–230. http://dx.doi.org/10.1086/269096

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5,* 213–236. http://dx.doi.org/10.1002/acp.2350050305

Little, R., & Yau, L. (1996). Intent-to-treat analysis for longitudinal studies with drop-outs. *Biometrics, 52,* 1324–1333. http://dx.doi.org/10.2307/2532847

Manfreda, K. L., Bošnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research, 50,* 79–104.

Marcus, B., Bošnjak, M., Lindner, S., Pilischenko, S., & Schütz, A. (2007). Compensating for low topic interest and long surveys: A field experiment on nonresponse in web surveys. *Social Science Computer Review, 25,* 372–383. http://dx.doi.org/10.1177/0894439307297606

McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics, 16,* 163–200. http://dx.doi.org/10.1007/BF01065485

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55,* 107–122. http://dx.doi.org/10.1007/BF02294746

Muthén, L. K., & Muthén, B. O. (2007). *MPlus user's guide*. Los Angeles, CA: Author.

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46,* 1023–1031. http://dx.doi.org/10.3758/s13428-013-0434-y

Porter, S. R. (2004). Raising response rates: What works? *New Directions for Institutional Research, 2004,* 5–21. http://dx.doi.org/10.1002/ir.97

Scherpenzeel, A., & Toepoel, V. (2014). Informing panel members about study results. In M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, & P. J. Lavrakas (Eds.), *Online panel research: A data quality perspective* (pp. 192–213). Chichester, United Kingdom: Wiley. http://dx.doi.org/10.1002/9781118763520.ch9

Stoel, R. D., van den Wittenboer, G., & Hox, J. (2003). Analyzing longitudinal data using multilevel regression and latent growth curve analysis. *Metodologia de las Ciencias del Comportamiento, 5,* 21–42.

Tuten, T. L., Galešić, M., & Bošnjak, 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,* 377–384. http://dx.doi.org/10.1177/0894439304265640

Warriner, K., Goyder, J., Gjertsen, H., Hohner, P., & McSpurren, K. (1996). Charities, no; lotteries, no; cash, yes: Main effects and interactions in a Canadian incentives experiment. *Public Opinion Quarterly, 60,* 542–561. http://dx.doi.org/10.1086/297772

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. Los Angeles, CA: University of California. Retrieved from https://www.statmodel.com/download/Yudissertation.pdf

Yu, J., & Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research, 20,* 36–44. http://dx.doi.org/10.2307/3151410

Zeigarnik, B. W. (1927). Das Behalten erledigter und unerledigter Handlungen [Recall of finished and interrupted actions]. *Psychologische Forschung, 9,* 1–85.

Received June 30, 2015
Revision received March 22, 2016
Accepted March 22, 2016 ■