Journal of Official Statistics, Vol. 22, No. 2, 2006, pp. 191–210

An Experimental Study on the Effects of Personalization, Survey Length Statements, Progress Indicators, and Survey Sponsor Logos in Web Surveys

Dirk Heerwegh[1] *and Geert Loosveldt*[2]

This study tests the effect of several factors on web survey response rates and on some other aspects of web survey data quality. The factors reviewed are e-mail personalization, survey length statements in the e-mail contact, use of progress indicators, and display of survey sponsor logos. Hypotheses are tested in a web survey experiment conducted with a university student sample ($n = 2{,}520$). The results indicate that, of all considered factors, only personalization significantly increases the web survey response rate.

Key words: Online survey; response rate; personalization; progress indicator; survey methods; social exchange theory; survey data quality; experiment.

1. Introduction

Together with coverage-associated problems and sampling difficulties, high unit nonresponse may be a serious reason for many survey researchers not to employ web surveys. Cook et al. (2000) for instance, report an average web survey response rate of 39.6%. Fortunately, methodological research aimed at improving web survey response rates has been conducted on a relatively large scale.

Some researchers have concentrated on technical factors influencing the response rate, such as login procedures (Crawford et al. 2001; Heerwegh and Loosveldt 2002), visual design of questionnaires (Dillman et al. 1998b; Lozar et al. 2002), and progress indicators (Couper et al. 2001; Crawford et al. 2001; Conrad et al. 2003). Other researchers have focused on such factors as incentives (Frick et al. 2001; Downes-Le Guin et al. 2002; Bosnjak and Tuten 2003; Porter and Withcomb 2003b) or the timing of sending out reminder e-mails (Crawford et al. 2001). Researchers have also paid attention to the effects of the content of the e-mail invitation itself on web survey response rates. The effects of survey length statements have been researched (Crawford et al. 2001), as have the influences of the social status of the solicitor (Guéguen and Jacob 2002b; Porter and Withcomb 2003a) and the effects of sending the solicitor's picture along with the e-mail request (Guéguen and Jacob 2002a). Recently, some research has also been conducted

[1] Katholieke Universiteit Leuven, Center for Survey Methodology, E. Van Evenstraat 2B, 3000 Leuven, Belgium Email: dirk.heerwegh@soc.kuleuven.be
[2] Katholieke Universiteit Leuven, Center for Survey Methodology, E. Van Evenstraat 2B, 3000 Leuven, Belgium Email: geert.loosveldt@soc.kuleuven.be.

regarding personalization of the e-mail invitations (Pearson and Levine 2003; Heerwegh et al. 2005; Porter and Withcomb 2003a).

In this study, we discuss a number of relevant web survey design and administration issues that potentially influence the response rate or other measures of data quality. More specifically, we will experimentally investigate the effects of personalization of the e-mail invitation, survey length statements, the use of progress indicators, and survey organization logos.

2. Theoretical Background and Hypotheses

The main theoretical framework that will be used throughout this study is social exchange theory. This theory states that "actions of individuals are motivated by the return these actions are expected to bring" (Dillman 2000, p. 14). Three elements are crucial for predicting an individual's actions: rewards, costs, and trust. Rewards refer to expected gains, costs refer to what one has to give up or spend to obtain the rewards, and trust refers to the belief that the rewards will outweigh the costs in the long run (Dillman 2000, p. 14).

2.1. Personalization of the Salutation in e-mail Invitations

Personalizing correspondence increases the degree to which recipients perceive the importance and value attached to them. In the context of social exchange theory, the recipient is shown positive regard, which has some reward value (Dillman 2000, p. 15). Therefore, the general expectation is that personalizing invitations will increase the response rate. This effect is extensively demonstrated in the realm of mail surveys. Dillman (2000, p. 152), for instance, argues that personalization in general public mail surveys increases response rates by 5 to 11%. Other researchers also have studied the effect of personalization (in a variety of definitions) on mail survey response rates and most of them have found the expected increase in response rate (e.g., Boser 1988; Dodd and Markwiese 1987; but see Eisinger et al. 1974; Andreasen 1970). Recently, researchers have begun investigating whether the same is true for web surveys that use e-mail invitations.

In a sample of university alumni, Pearson and Levine (2003) varied the degree of personalization in e-mail invitations to participate in a web survey. Sample units in the control condition received a nonpersonalized salutation ("Dear Stanford Alum"), (it should be noted that the term "alum" is the informal form of "alumnus," and consequently, even the impersonal salutation is still relatively informal in this study) whereas those in the three treatment groups received some form of personalized invitation. One group received a formal salutation (e.g., "Dear Mr. Smith"), another group a familiar one (e.g., "Dear John"), and yet another group received a familiar salutation without "dear" (e.g., "John"). It was found that personalization, regardless of type, increased the response rate by a few percentage points, although the increase was not statistically significant (Pearson and Levine 2003, p. 4).

In a less complex experiment conducted among university students, Heerwegh et al. (2005) defined two experimental groups. One group received an impersonal salutation ("Dear Student"), while the other received a personalized one (e.g., "Dear John Smith"). They observed a significant increase in the response rate as it rose from 49.1% in the

impersonal salutation group to 57.7% in the personalization group. In contrast, Porter and Whitcomb (2003a) did not observe a significant effect of personalization on the response rate. The high school students receiving the personalized salutation (e.g., "Dear John") did not respond to the survey in larger numbers than those receiving the impersonal invitation ("Dear Student"). In line with the general theory, the current study expects personalization to increase the web survey response rates.

Although personalization might increase response rates, it could also compromise privacy (Joinson, Woodley, and Reips, forthcoming). Joinson et al. (forthcoming) found some indications that personalization reduced self-disclosure on a sensitive question (income) in that the "I prefer not to say" option was more often chosen. The current study takes the same position and assumes that personalization might compromise privacy. Consequently, it is expected that personalization will reduce self-disclosure on sensitive questions in the current survey.

2.2. Survey Length Statements

One obvious component of respondent costs is the amount of time a survey takes to complete. A longer survey entails a higher cost, which could translate into lower response rates. In a study of Crawford et al. (2001) a long and a short survey length statement in the e-mail invitation were compared. The long survey length statement (mentioning that the survey would take 20 minutes) yielded a significantly lower login rate than the short survey length statement (avowed survey length of 8–10 minutes), but it generated fewer break-offs, resulting in comparable response rates (Crawford et al. 2001, p. 153).

The discussion of the Crawford et al. (2001) study shows there are different web survey participation rates. For reasons of clarity, these rates are briefly defined here. The login rate is defined as the proportion of sample cases that starts the web survey by logging on to it. The completion rate is defined as the proportion of *respondents* (i.e., sample units who started the survey) who reach and submit the final page of the web survey. The break-off or dropout rate is defined as the proportion of *respondents* that do not reach the end of the survey. The web survey response rate is defined as the proportion of *sample cases* that reach and submit the last page of the web survey. The completion or retention rate is hence not equal to the (complete) response rate. The completion or retention rate refers to the percentage of sample units completing the survey after initiating it. By contrast, the (complete) response rate refers to the percentage of sample units completing the survey upon being invited to answer the survey (perhaps excluding noncontacted units, ineligible units, etc.).

The mentioned effects of the survey length statements can be understood in terms of social exchange theory. By lowering the perceived costs in terms of completion time, more initial response is generated (the login rate increases). However, as the respondent realizes that the survey is in fact longer, the trust placed in the researcher by the respondent is revoked, leading to halting survey participation (increasing the break-off rate and depressing the response rate).

It seems clear that the perceived respondent cost should be as small as possible to elicit high initial response (high login rates). At the same time, the chance of the respondent feeling deceived and dropping out of the survey must be minimized. One could remain

vague about the survey length and state that it has been kept "as short as possible." Such a statement could make the survey appear short, which would decrease the perceived costs of participation and lead to an increased login probability. In addition, the vague length statement "the survey has been kept as short as possible" could invoke the principle of reciprocation (Cialdini 1984). This principle is closely related to social exchange theory (Groves and Couper 1998, p. 33). The vague statement carries with it the connotation that the survey researcher has made an effort to keep the questionnaire short – a behavior that could be responded to by cooperation in greater numbers. It is therefore expected that the vague survey length statement "the survey has been kept as short as possible" will elicit higher login rates than a more specific survey length statement simply mentioning the estimated time needed to complete the survey if this amount of time is considerable (e.g., 20 minutes or more). (In our study, the respondents in the specific survey length condition will be told that the survey will take 20 to 25 minutes; see below.)

While the vague survey length statement should, similarly to a short length statement, yield a higher login rate than a longer length statement, we expect that with the vague length statement we can avoid the increased break-off rates associated with a short length statement. With a short specific survey length statement, it can in principle be objectively determined when the survey is starting to take longer than indicated in the solicitation e-mail. A vague survey length statement, in contrast, does not define such a clear line to cross. Respondents might hence remain in the survey longer, inhibiting the break-off rate from increasing. If this is the case, the net result of the vague survey length statement should be a higher response rate than one obtained by the specific length statement informing sample cases that the survey will take 20 to 25 minutes.

2.3. Progress Indicator

While both previous factors are thought to influence the decision to start a web survey, this factor only affects whether or not respondents remain cooperative and complete the survey. The underlying mechanism concerning progress bars refers to respondent motivation. As respondents progress through a survey questionnaire, fatigue and loss of motivation may lead them to quit the survey perhaps just a few questions before the end (Dillman et al. 1998a, pp. 12–13; Couper et al. 2001, p. 232). Displaying a progress bar may keep respondents motivated to complete the survey as the end of it draws visibly nearer with every question answered. Because of its presumed positive effect on the completion likelihood, use of a progress indicator is often recommended for page-by-page web surveys. In single-page web surveys, where all survey questions are listed on a single web page, the scrollbar may function as a progress indicator, thereby making an additional, more explicit progress indicator redundant (Dillman et al. 1998a, p. 13).

Respondents can use progress bars to periodically reassess how much more time and effort will be required to finish the survey. In terms of social exchange theory, this means that the costs-rewards analysis can be updated periodically. If the progress indicator suggests higher than anticipated costs of continued participation, the probability of breaking off survey participation should increase; if lower, the probability should decrease. An illustration of this effect is to be found in the study by Conrad et al. (2003). They compared three types of progress indicators, each using a different transformation so

they would exhibit fast or slow progress early or late in the questionnaire. It was found that displaying fast progress early in the questionnaire (the fast-to-slow progress bar), which leads to a favorable impression regarding the required time and effort to complete the remainder of the survey, had the most favorable outcome in terms of break-off rates. Still, even this best-performing type of progress indicator (fast-to-slow) did not consistently generate significantly better data quality than when no indicator was displayed at all.

The problematic connection between progress indicators and data quality is also retrieved in research conducted prior to the Conrad et al. (2003) study. In one study, Couper et al. (2001, p. 243) found that offering a progress indicator did increase the completion rate by 3.5 percentage points, but this increase was not statistically significant. The authors argued that the positive effect of the progress indicator might have been cancelled out to some degree by the longer download times associated with the (graphical) progress indicator (Couper et al. 2001, p. 243). In an attempt to keep download times constant, Crawford et al. (2001, p. 149) implemented a text-based progress indicator and offered it in the case of both experimental conditions, although in one of the conditions the text was rendered in the background color so the information would not be visible. Despite the careful setup of this experiment, the results indicated that displaying the progress indicator significantly increased the break-off rate instead of decreasing it (Crawford et al. 2001, p. 156). This could have been caused by an improper calibration of the progress indicator inducing a false sense of higher burden to complete the survey (Crawford et al. 2001, p. 157). In the current study, we will use data from survey pretests to calibrate the progress indicator on the basis of actual working times (rather than on the basis of the number of questions), which should ensure a proper calibration of the progress bar. Then, the fast-to-slow transformation will be applied since this transformation was shown to be the most effective in the Conrad et al. (2003) study. Given the proper calibration and transformation of the progress indicator, and in line with the theory, we expect a decreased break-off rate to result from the use of the progress indicator.

It was mentioned that the progress indicator could keep respondents motivated to answer the survey. Hence, it is sometimes assumed that progress bars decrease item nonresponse rates, although empirical tests failed to show significant effects (Couper et al. 2001, p. 232; Conrad et al. 2003). In the current study, we adhere to the theoretical assumption and expect the progress indicator to decrease the item nonresponse rate.

Progress bars have been reported to influence the subjective experience of the survey. Conrad et al. (2003) found that although respondents given the (fast-to-slow) progress indicator needed more time to complete the survey, they, after completing the survey, estimated the survey duration significantly lower than those not given this progress bar. This finding is surprising because psychological studies have found that greater time awareness (due to the presence of clocks) can lead to a subjective feeling that time is moving slower (and hence to an *over*estimation of duration), rather than to an increase in the accuracy of time interval estimates (Conti 2001). Perhaps the incongruence is due to the fact that a progress indicator is not a clock; although both instruments can give an impression of how fast (or how slow) time is moving. The current study hypothesizes that progress indicators will stimulate a positive evaluation of the survey duration by the respondents. This prediction runs counter to the contention that administering time-keeping stimuli increases time awareness, leading to an overestimation of the duration.

Our preference is based on the empirical evidence found by Conrad et al. (2003), and on the consideration that progress bars are not identical to clocks, although they can both serve to display how fast (or slow) time is moving.

2.4. *Using the Survey Sponsor's Logo in the Questionnaire*

Besides using a progress indicator, another potential strategy for keeping respondents in the survey is to use the authority principle (Cialdini 1984) and to display the logo of the survey sponsor on each survey questionnaire screen. In terms of social exchange theory, this is a means of establishing (and reinforcing) trust (Dillman 2000, p. 20). The effects of a logo are not straightforward, however. If the respondents hold the organization in high esteem, then repeating the logo on each survey screen could decrease break-off rates. Conversely, if the respondents do not hold the organization in high esteem, or even have serious doubts about its legitimacy, then a logo might produce the opposite effect. In this study, it is hypothesized that displaying a logo during the survey will increase the completion rate of the current survey. The reason for this directionality has to do with the specific sample. As will be detailed below, the survey sponsor was a university, and the sample consisted of its own first-year students. Thus, it is reasonable to assume that most respondents hold a positive attitude toward the survey sponsor.

3. Sample and Study Design

The sample was randomly drawn from the official database of all first-year students of the Katholieke Universiteit Leuven, Belgium. The Katholieke Universiteit Leuven is, with its 30,000 students, the largest university in Flanders (the Northern and Dutch-speaking part of Belgium). A "first-year student" is defined as a student enrolling for a first-year curriculum at the University – not necessarily for the first time. From the total database ($n = 6{,}304$), students unlikely to speak Dutch, students over 20 years old and students not enrolling for a full-time first-year academic curriculum were excluded. After this operation, the total number of eligible first-year students was 5,040 (i.e., the sample frame). From this frame, a simple random sample of $n = 2{,}520$ was drawn. In addition to several background variables (such as major, date of birth, etc.), the database contained the e-mail address of each student, as well as his or her first and last name. The university provides free high-speed Internet access for all enrolled students in a large number of computer labs and libraries, and also in the housing units of all students residing in the vicinity of the city of Leuven. Therefore, it was assumed that most of the sample units would have the opportunity to access the Internet at a low or at no financial cost at all.

To test the effects of the various factors on survey data quality, a $2 \times 2 \times 2 \times 3$ experimental design was used. In this design, personalization was varied. This represents the first experimental factor comprising two levels. Prior to the fieldwork, 50 percent of the sample units were randomly selected and assigned to the no-personalization condition. The remaining 50% were allocated to the personalization condition. In the no-personalization group, the salutation used in all e-mail contacts read "Dear student," whereas in the other group it read "Dear [*First name*] [*Last name*]" (e.g., "Dear John Smith").

The second factor experimentally manipulated was the type of survey length statement given in the e-mail invitations. Prior to the fieldwork, but independently of the assignment to the levels of the first factor, half of the sample units were randomly assigned to the specific survey length condition, the other half to the vague survey length condition. In the specific length statement condition, and the e-mail message – both the initial invitation and the reminder – mentioned that completing the survey would take "approximately 20 to 25 minutes." In the vague statement condition, the e-mail message stated that the survey questionnaire had been kept "as short as possible." In both conditions, it was said that the precise survey duration would depend on the answers given to some questions.

The third factor reflects whether or not sample units were given the choice regarding the display of the progress indicator ("Progress indicator on demand"). Prior to the fieldwork and independently of the assignment to the levels of both previous factors, everyone in three sample units was randomly assigned to the choice condition, while the remaining units were fitted into the no-choice condition (see Figure 1). In the choice condition, a sentence absent in the other condition appeared on the survey login page. This sentence appeared in blue type as opposed to black type for all other text information on this page, in an attempt to draw attention to it. The sentence briefly explained to the respondents that they could opt for a progress indicator to be displayed. In order to have it displayed, they were instructed to select a checkbox. (In order to prevent returning respondents in the choice condition from changing their choice over multiple visits, which would confound

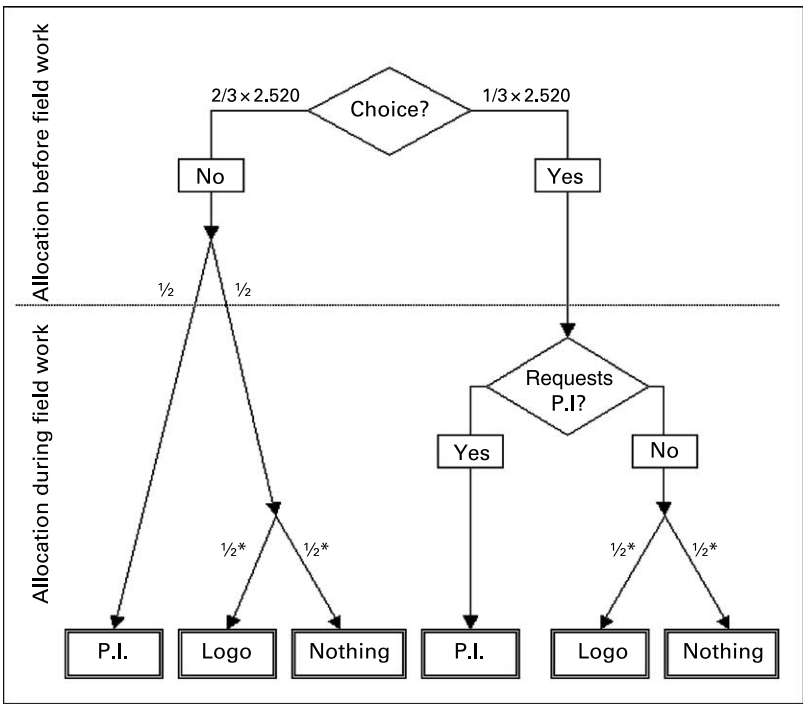

Fig. 1. Procedure to determine whether a progress indicator (P.I.), a survey sponsor logo (Logo), or nothing would be displayed during the survey. NOTE: Randomization procedures carried out during the fieldwork are marked with ∗

our analyses, the respondents' first-visit choice was recorded. Upon subsequent visits, their preference would be fetched from the database rather than asking again whether they would like to have the progress indicator displayed.) Inclusion of this factor allows a test of how many respondents actually expect or would like to have a progress indicator displayed.

The fourth factor refers to which motivator (a progress indicator, a logo or nothing) was displayed, either by design or by choice. At this point, it becomes clear that the value of the fourth experimental factor is not always determined at random. In the choice condition, display of the progress indicator depends on a volitional act of the sample unit. If the respondent selects the checkbox, the progress indicator is displayed (see Figure 1). Conversely, if the checkbox is not selected, the progress indicator is not displayed. Instead, the university logo or nothing would be presented. It was decided to display either the logo or the progress indicator, but never both simultaneously, to keep the survey screens from becoming visually too complex. This means that, at the time respondents logged on to the survey, the web survey software had to randomly decide whether to show nothing at all or the logo, to those who did not select the checkbox. A randomization procedure was devised so that half of the respondents not selecting the checkbox would receive the logo while the other half would receive nothing (see Figure 1).

In the group of respondents not given the choice regarding progress indicator display, half of the sample units were randomly set to have the progress indicator displayed prior to the fieldwork (see Figure 1). For the remaining sample units, the allocation to either a logo or nothing was delayed until the respondent logged in. At that time, the web survey software allotted the respondent to one of the two groups (logo or nothing) in exactly the same manner as used to allocate the respondents in the choice condition who chose not to have the progress indicator displayed (see Figure 1). Thus, regardless of the choice condition, the fourth factor can assume three values: progress indicator, logo, or nothing (see Figure 1). However, for those in the choice condition, the value is determined by choice, whereas in the other condition it is determined by design. Also regardless of the choice condition, if the logo was displayed, it appeared in exactly the same location as where the progress indicator would otherwise have been. (That position was the upper right corner of each screen. Each survey screen displayed one survey question or up to 10 survey items if an item battery was used. Item batteries appeared as "grid" or "matrix" questions: The survey items were listed in the rows and the common response options were arranged in the columns.)

The survey questionnaire was equipped with the client side paradata software as described by Heerwegh (2003). This software allows recording of response times to individual survey questions, the sequence in which multiple items on a single screen are answered, etc. The survey was briefly fielded for pretesting purposes. It was put on a local intranet server and a small number of university staff was asked to fill in the survey. Using the response times collected with the client side paradata, the progress indicator was calibrated. This was done by dividing the average amount of time spent on each individual survey question by the total average working time. After calibration, the progress indicator was set to perform a transformation so it would behave similarly to the fast-to-slow progress indicator of Conrad et al. (2003). (This was accomplished by taking the squared root of the working time proportion as determined in the pretest. Thus, when

the respondent had completed 50% of the questionnaire in terms of working time, the progress bar indicated a progress status of 71% ($\sqrt{.50 \times 100}$).) The progress indicator consisted of a graphical representation of the survey progress (a horizontal colored bar), supported by a line of text indicating how much of the survey questionnaire was completed. As people progressed through the survey, the originally totally gray bar would gradually, from left to right, color blue (see Figure 2, left panel).

The university logo (see Figure 2, right panel) was based on the official logo, with its size adjusted to approximate the space taken up by the progress indicator. It also used the same colors as the progress indicator. In order to keep average download times constant across experimental groups, the progress indicator as well as the logo were downloaded by all respondents, regardless of their condition or choice. Depending on the specific condition, the progress indicator, the logo, or both would be rendered invisible, either by resizing the progress indicator and/or the logo to a very small size (1 pixel by 1) or by displaying the text (displayed below the progress bar) in the background color. (To keep the download time of the graphical part of the progress indicator as short as possible, it actually consisted of only two very small graphic files of 1 pixel by 1 (one file being a blue pixel, the other a gray pixel – these are the colors used in the progress indicator). These graphic files were adjusted to the appropriate size by the web survey software. The university logo would also not require much download time because it would normally be cached by the respondent's computer so that it would be downloaded only once and thereafter retrieved from the respondent's own computer hard drive.)

4. Survey Questionnaire and Fielding Procedure

The questionnaire was designed at the Center for Population and Family Research (Katholieke Universiteit Leuven). The survey topic was young adults' attitudes toward marriage and divorce. Questions probing for respondents' sexual attitudes and behaviors were included in the questionnaire. The survey questions were distributed across 13 web pages, making it a page-by-page web survey. (Each web page contained a series of questions (or item batteries). The questions were interspersed with enough white space to ensure only one question could be seen on the screen at a time. This mimics a true page-by-page design in which each question is placed separately on a single page, while avoiding the large number of server connections associated with such a true page-by-page design.) Depending on the skip pattern, 231 to 284 survey items were to be filled in by the respondent. Skips were taken care of by the web server, and response to branching questions was made mandatory to insure a proper skip pattern. When such a mandatory question was left blank, a prompt would be displayed and the respondent would be urged to answer the question before he or she could move on to a subsequent question (these were "hard prompts"; see DeRouvray and Couper 2002). Response to other survey

Fig. 2. Progress indicator (translated from Dutch) and University logo

questions was not compulsory. Leaving these questions unanswered did not trigger a prompt. The item nonresponse was accepted and the respondent was allowed to continue to a subsequent question.

Using conditional questions poses some problems when it comes to calibrating the progress indicator, because skips cannot be anticipated. Hence, the progress indicator calibration is the same for every respondent, regardless of his or her pathway through the survey questionnaire. In this survey, questions could simply be skipped; there were for instance no simultaneous alternative pathways. It was decided to assume that every respondent would have to fill in the entire survey. This would somewhat slow down the overall progress indication, but the indicator would show big "leaps" each time a set of questions was skipped. This was assumed to be a better situation than a "hanging" or "stalling" progress indicator when a respondent was guided to a set of questions that were not skipped for him or her, but could be skipped by the other respondents.

The invitations to respond to the web survey were sent out by e-mail on Wednesday, December 3, 2003. One week later an e-mail reminder was sent to the sample units who had not yet completed the survey. The survey was closed on Wednesday, January 14, 2004.

5. Results

Of all the 2,520 e-mail invitations sent, 41 (1.63%) proved undeliverable. Of the remaining 2,479 sample units, who are assumed to have received but therefore not necessarily read the e-mail message, 1,602 sample units (64.62%) logged on to the web survey and 1,421 (57.32%) completed it. Thus, the login rate equals 64.62%, the overall break-off rate 11.30% (100 – [1,421/1,602] × 100), and the response rate 57.32% (1,421/2,479 × 100).

Of the 1,602 sample units logging on to the survey, 1,151 (71.85%) did so prior to the reminder e-mail (7 days of fieldwork, response Wave 1). The remaining sample units who logged on to the survey (451 or 28.15%) did this after the reminder e-mail and before the end of the fieldwork (36 days of fieldwork, response Wave 2).

In order to prevent a possible confounding of the experimental factors and the potential differences between respondents who did and did not choose to display the progress indicator in the choice condition, the analyses to be presented will exclude the cases from the choice condition. This means the sample size drops from 2,520 to 1,680. In this random subgroup of the sample, 29 e-mails bounced (1.73%). Of the remaining 1,651 cases, 1,085 (65.72%) logged on to the survey, and 961 (58.21%) completed it (11.43% break-off rate). In this subsample, 772 (71.15%) cases logged on before the first reminder, while the remaining cases that logged on did so after the reminder e-mail ($n = 313$, 28.85%). These rates and figures are very similar to those observed in the entire sample.

5.1. *Effects of Personalization*

As expected, the login rate is significantly higher in the personalization condition (69.09%) than in the impersonal one (62.35%; $\chi^2(1) = 8.3281$; $p = 0.0039$, $n = 1{,}651$). Also, personalization influences the break-off rate. Whereas the break-off rate equals 8.95% in the personalized group, it amounts to 14.17% in the other group ($\chi^2(1) = 7.3036$; $p = 0.0069$, $n = 1{,}085$). A higher login rate and a lower break-off

rate combine into a significantly higher response rate when personalization is applied than when it is not (62.91% vs 53.51%, respectively; $\chi^2(1) = 14.9864$; $p = 0.0001$, $n = 1{,}651$). In the personalization group, 71.40% of the responses came in before the reminder and 28.60% after the reminder. In the other group, these percentages are 70.87% and 29.13%, respectively ($\chi^2(1) = 0.0370$; $p = 0.8475$, $n = 1{,}085$). This indicates that late (post-reminder) respondents react in much the same way to personalization as do early (prereminder) respondents.

Since this questionnaire contained sensitive questions on sexual behavior, one of these was subjected to a closer inspection in order to evaluate the hypothesis that personalization would lead to increased nondisclosure. (The selection criterion was that the question should be sensitive, that it should include response alternatives allowing the respondent to avoid answering the question (*I don't know*, *I'd rather not answer*), and that it should be possible to continue the survey without selecting any response option. The currently selected question was the only one exhibiting these characteristics.) Question 37 probed for the sexual inclination of the respondents. The response alternatives were *heterosexual*, *homosexual*, *bisexual*, *I don't know*, and *I'd rather not say*. In addition, respondents could leave the question blank without selecting any of these response options and move on to a subsequent question. Table 1 details the responses to this question by experimental condition.

The differences across experimental conditions observed in Table 1 are not statistically significant (Fisher's exact test, $p = 0.6220$, $n = 961$). (The same conclusions are reached when analyzing the data by response wave.) Because respondents may perceive the heterosexual category as the more socially desirable one, it is conceivable that personalization may have led to more nonheterosexual respondents selecting this option in order to hide their true sexual inclination. An analysis did not substantiate this hypothesis, however. In the personalization condition 95.46% claimed to be heterosexual and in the impersonal condition 97.45% did so ($\chi^2(1) = 2.6474$; $p = 0.1037$, $n = 939$). This analysis hence indicates that the salutation does not affect the propensity to select the heterosexual category as opposed to other categories indicating different sexual inclinations. The same conclusions are reached when analyzing the data by response wave.

Although probably not as sensitive as the sexual inclination question, there was a question on political preference. The question asked, "If there should be federal (Belgian) elections tomorrow, which one of the following political parties would you vote for?" The "*I don't know*" response alternative was offered and respondents could skip to the subsequent question without answering the current question.

Table 1. Personalized salutation and selection of "I don't know," "I'd rather not say" and submission of no selection to a question on sexual inclination (n and %)

	Disclosed sexual inclination	Chose 'I don't know'	Chose 'I'd rather not say'	Submitted no selection	Total
Personalized	507 (97.69%)	7 (1.35%)	1 (0.19%)	4 (0.77%)	519
Impersonal	432 (97.74%)	7 (1.58%)	2 (0.45%)	1 (0.23%)	435
Total	939 (97.74%)	14 (1.46%)	3 (0.31%)	5 (0.52%)	961

Although the differences across conditions are statistically significant (See Table 2; $\chi^2(2) = 7.3566$; $p = 0.0253$, $n = 916$), when the same analysis is repeated by response wave, the p-values rise above the 0.05 alpha level (0.1094 for Wave 1 respondents, 0.1455 for Wave 2 respondents). These data do not corroborate the hypothesis that personalization would lead to increased use of a nondisclosure option. These data in fact suggest the opposite since personalization increases self-disclosure in this question.

Using the cases that completed the questionnaire, the overall item nonresponse and *don't know* rate were also calculated (excluding questions that were not applicable). Personalization does not influence the overall item nonresponse rate or the overall *don't know* rate. In the personalized group, the overall item nonresponse rate equals 2.18% while it is 1.71% in the other group (Kruskal-Wallis $\chi^2(1) = 0.1893$; $p = 0.6635$, $n = 961$). The *don't know* rate equals 0.56% in the personalized condition and 0.59% in the impersonal one (Kruskal-Wallis $\chi^2(1) = 0.2075$; $p = 0.6487$, $n = 961$). (The same conclusions are reached when analyzing the data by response wave.)

5.2. *Effects of the Survey Length Statement*

As expected, the vague length statement produces a somewhat higher login rate (67.23%) than the specific length statement (64.21%). The difference is not statistically significant, however ($\chi^2(1) = 1.6765$; $p = 0.1954$, $n = 1{,}651$). The vague length statement generates higher break-off rates than the specific length statement, although this difference is not statistically significant either (12.09% vs 10.73% respectively; $\chi^2(1) = 0.4950$; $p = 0.4817$, $n = 1{,}085$). Consequently, the net effect of administering a vague length statement is a slight, but not statistically significant increase in the response rate. The response rate rises from 57.32% to 59.10%, but the Chi-square test fails to reach statistical significance ($\chi^2(1) = 0.5414$; $p = 0.4618$, $n = 1{,}651$).

In the vague length statement group, 71.84% of the responses came in before the reminder and 28.16% after the reminder. In the other group, these percentages are 70.43% and 29.57%, respectively ($\chi^2(1) = 0.2619$; $p = 0.6088$, $n = 1{,}085$). This indicates that late (post-reminder) respondents react much the same way to the survey length statements as do early (prereminder) respondents.

5.3. *Combined Effects of Personalization and Survey Length Statements*

A logistic regression model was tested to determine the effect of personalization and survey length statements on the login and (complete) response probability. It should be

Table 2. Personalized salutation and selection of "I don't know" and submission of no selection to a question on political preference (n and %)

	Disclosed political preference	Chose 'I don't know'	Submitted no selection	Total
Personalized	440 (84.78%)	68 (13.10%)	11 (2.12%)	519 (54.01%)
Impersonal	353 (79.86%)	84 (19.00%)	5 (1.13%)	442 (45.99%)
Total	793 (82.52%)	152 (15.82%)	16 (1.66%)	916

Table 3. Effect of personalization and survey length statements on login and complete response probabilities

	Odds ratio login probability	Odds ratio complete response probability
Personalization (1 = yes)	1.350**	1.473***
Survey length statement (1 = "20 to 25 minutes")	0.874	0.930
n	1,651	1,651

*$p < .05$, **$p < .01$, ***$p < .001$.

noted that the login probability refers to the probability of starting the web survey (this refers to the login rate), and that the (complete) response probability refers to the probability of reaching and submitting the final web survey page (this refers to the response rate).

Table 3 shows that personalization has a significant effect on the login probability, while controlling for survey length statement (which does not have a significant effect on the login probability). The same goes for the probability of completing the web survey (i.e., sample cases reaching and submitting the final web survey page). This analysis hence confirms the results from the bivariate analyses.

5.4. Effects of the Progress Indicator[3]

5.4.1. How Many Respondents Request a Progress Indicator?

As a preliminary analysis regarding progress bars, this section looks at how many survey respondents actually want to receive feedback on their survey progress. Of all 840 subjects assigned to the choice condition, 517 logged on to the survey and made a choice regarding the progress indicator. (In this condition, 12 e-mail messages were undeliverable. Thus, in this condition the login rate equals 62.4%. This rate is very similar to the overall login rate.) Of these 517 respondents, the vast majority chose to have the progress indicator displayed (77.4%) by selecting the checkbox on the login page, while the rest did not indicate this wish (i.e., they did not select the check box). The proportion of sample units choosing the progress indicator is not affected by the specificity of the information concerning survey length in the e-mail invitation. If only vague information is offered, about 77.1% of all sample units choose the progress indicator. This rate is nearly identical to that in the specific information condition (77.7%; $\chi^2(1) = 0.0285$; $p = 0.8659$, $n = 517$). The percentage of respondents requesting a progress indicator does vary with response wave. In the first wave, 81.00% of the respondents request the progress bar, but in the second wave only 67.39% do that ($\chi^2(1) = 10.7045$; $p = 0.0011$, $n = 517$). (The units that did not select the checkbox were randomly distributed over the "logo" (11.4%) and the "nothing" (11.2%) condition by the web survey software.) This suggests that most respondents want to be informed about their survey progress. As mentioned, these cases from the choice condition are not included in any of the other analyses presented because

[3] Please note that all analyses in this subsection compare the progress indicator condition with the condition in which nothing was shown (no logo either). This avoids confounding between being shown nothing and being shown the logo.

of a possible confounding between the effects of those factors and the potential differences between sample cases choosing and sample cases not choosing to have a progress indicator displayed. These cases were used only in this specific analysis to determine the percentage of respondents that requested a progress bar.

5.4.2. Effect of the Progress Indicator On Survey Completion

It was hypothesized that the progress indicator would reduce survey break-off. In the progress indicator group, the break-off rate is 11.27% ($n = 559$) while it is 12.55% in the group that was shown nothing ($n = 263$). Although the direction of the effect is as expected, the difference is not statistically significant ($\chi^2(1) = 0.2829$; $p = 0.5948$, $n = 822$). (The same conclusions are reached when analyzing the data by response wave.)

Since this was a page-by-page web survey, survey break-off can be pinpointed to a specific exact web survey page. Figure 3 shows the break-off profile for the progress indicator condition versus the condition in which nothing was shown. The figure only includes sample cases that started the web survey and displays only the web survey pages that could not be skipped by any of the respondents.

Figure 3 shows that the break-off profile is very similar across both groups. Chi-square tests revealed that on none of the various web survey pages did the break-off rates differ significantly across both groups considered. (The same conclusions are reached when analyzing the data by response wave.) Hence, this analysis confirms that the progress indicator does not influence web survey break-off.

5.4.3. Combined Effect of the Progress Indicator and the Survey Length Statements On Survey Completion

In theory, both progress bars and survey length statements can influence the survey break-off rate. So far, no empirical support for this assumption has been found, but it is possible that the progress bars interacted with the survey length statements (see also Crawford et al. 2001). An analysis showed, however, that the break-off rates are similar across all groups. In the vague survey length statement, the break-off rate equals 11.59% in the progress bar group as opposed to a rate of 11.72% in the group where no progress indicator was present. In the specific survey length statement group, the break-off rate is 10.95% in the progress indicator group, while it is 13.33% in the group where no such indicator was available. In a

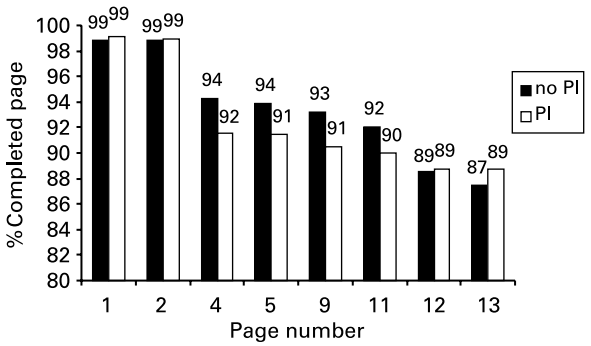

Fig. 3. Dropout profile across conditions (no progress indicator vs progress indicator present)

logistic regression model, there was no significant main effect of the progress indicator on the completion likelihood and there was no significant main effect of the survey length statement either. In addition, there was no significant interaction effect between the survey length statement and the progress bar on the completion probability. This analysis thus indicates that the break-off rates are not under the influence of either of these factors or their combination.

In an effort to further analyze any possible interaction effects, we looked at the break-off profile in each of the groups. Figure 4 shows that when respondents are told the survey will take 20 to 25 minutes, the break-off profile is similar regardless of whether or not a progress indicator is present. A series of Chi-square tests confirmed that the break-off rates are not significantly different on any of the web survey pages across both considered groups.

In Figure 5, however, there are indications that the progress indicator generates a different break-off profile when the respondents are told the survey was kept as short as possible. On web survey pages 4 and 5, the break-off rates in the indicator group are significantly higher than in the group where no progress bar was shown (progress indicator vs no progress indicator: page 4: 9.78% vs 3.91% resp., $\chi^2(1) = 4.1402$; $p = 0.0419$, $n = 404$; page 5: 10.14% vs 3.91% resp., $\chi^2(1) = 4.5373$; $p = 0.0332$, $n = 404$).

A supplementary analysis including only those who started but did not finish the survey (i.e., survey dropouts) yielded results pointing in the same direction. In the specific survey length statement condition, the average time at which break-off occurred equals $9'36''$ when no progress indicator was present ($n = 18$) and $8'19''$ when a progress indicator was shown ($n = 31$). This difference is not statistically significant (Kruskal-Wallis $\chi^2(1) = 0.1318$; $p = 0.7166$). In contrast, in the vague survey length statement condition there is a significant difference in the mean time at which survey break-off occurred across the two groups. In the group not shown a progress bar ($n = 14$), the average time until survey break-off equals $15'13''$, while it is only $3'03''$ in the group where a progress indicator was displayed ($n = 32$; Kruskal-Wallis $\chi^2(1) = 7.8751$; $p = 0.0050$).

These analyses provide some basis for the assumption that respondents periodically reassess the costs-rewards analysis during survey participation. Those that were told that the survey was kept as short as possible apparently assumed that it would be short.

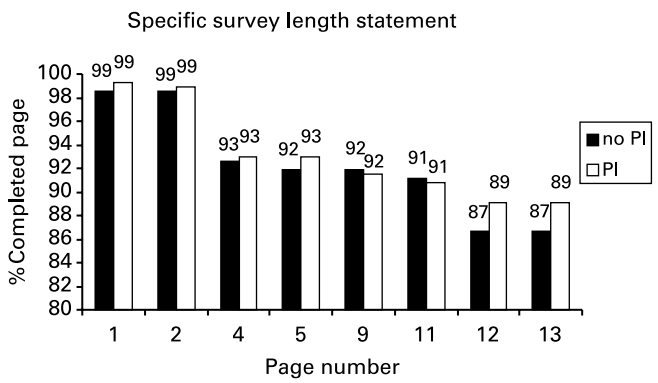

Fig. 4. Dropout profile across conditions (no progress indicator vs progress indicator present) for the group of respondents told that the survey would take 20 to 25 minutes

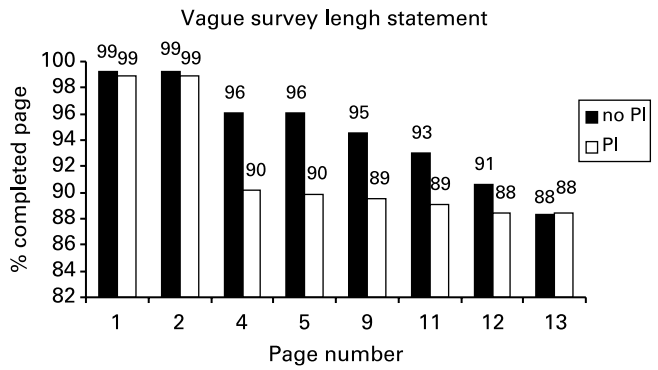

Fig. 5. Dropout profile across conditions (no progress indicator vs progress indicator present) for the group of respondents told that the survey was kept as short as possible

However, when they were given feedback on their progress, they were able to determine on that basis that the survey would in fact take longer than they had initially anticipated, which could explain the rapid survey dropout observed in that group. Those not given the progress indicator but also told that the survey was kept as short as possible did not have a means to determine that the survey would take longer than anticipated. This could explain the more evenly spread break-off profile in this group.

5.4.4. Effect of the Progress Indicator On the Item Nonresponse Rate

For each respondent reaching and submitting the final web survey page, an item nonresponse rate was calculated. This ratio reflects the number of unanswered questions that should have been answered taking into account skip patterns. It was found that the item nonresponse rate is lower when a progress indicator is present than when nothing is present (1.67% vs 2.63% respectively, Kruskal-Wallis $\chi^2(1) = 8.1181$; $p = 0.0044$; $n = 726$). Thus, similarly to Conrad et al. (2003), we find an effect of the progress indicator on item nonresponse when compared to the situation where nothing is displayed. (The same conclusions are reached when analyzing the data by response wave.) It should be noted however, that from a practical point of view the difference between the two conditions is small.

5.4.5. Effect of the Progress Indicator On Survey Duration Estimates

Conrad et al. (2003) found that the fast-to-slow progress indicator reduced the respondents' survey duration estimation after completing the survey, although this runs counter to the time awareness hypothesis. At the end of the current survey, respondents were also asked to estimate the total amount of time it had taken them to complete the survey. Similarly to Conrad et al. (2003), we find that the progress indicator yields the shortest survey time estimate (26.79 minutes vs 27.95 minutes in the nothing displayed group). The difference between these estimates is however not statistically significant (Kruskal-Wallis $\chi^2(1) = 2.2962$; $p = 0.1297$; $n = 702$). (When including only the respondents who completed the web survey in a single session, no significant difference was found either.) The reason for the direction of the difference could be that the progress indicator group did effectively complete the survey in a shorter amount of time than the

group that was shown nothing. In the progress indicator group, the average survey duration was 36 minutes and 53 seconds ($n = 487$). In the group where nothing was displayed ($n = 226$), the average survey duration was 39 minutes and 14 seconds. (These times exclude outliers. Outliers were defined as cases with a total survey duration smaller or larger than the average survey time (2,707.71 seconds) minus or plus twice the standard deviation (3,349.28 seconds), respectively.) Although this difference is not statistically significant (Kruskal-Wallis $\chi^2(1) = 2.3573$; $p = 0.1247$), the direction is consistent with the respondents' subjective evaluation of the survey duration.

In an effort to further investigate the time awareness hypothesis, the final open-ended question, positioned at the very end of the survey and probing for concluding remarks or comments on the survey, was used to dummy code a variable indicating whether the respondent had stated that the survey was too long. The progress indicator showed to significantly decrease the percentage of respondents saying that the survey took too long (6.25% vs 13.48% in the progress indicator vs the nothing displayed group, respectively; $\chi^2(1) = 10.5112$; $p = 0.0012$, $n = 726$). These findings are inconsistent with the time awareness hypothesis. If that hypothesis were correct, we should have observed more, not fewer, remarks that the survey took too long in the progress indicator condition. Thus, although progress bars may influence time perception, it does not seem that they would slow down the subjective experience of time. One needs to be aware, however, that we used a "fast-to-slow" progress bar. Using a different type of indicator might produce the effect of slowing down time perception.

5.5. Effects of the Survey Sponsor Logo

Display of the survey sponsor logo instead of displaying nothing did lead to the anticipated break-off rate reduction, although the effect does not reach statistical significance. In the group not shown a logo (and no progress indicator either), the break-off rate equals 12.55%. In the group where the logo was displayed, the break-off rate equals 10.65% ($\chi^2(1) = 0.4636$; $p = 0.4959$, $n = 526$). (The same conclusions are reached when analyzing the data by response wave. There was no evidence of an interaction effect between survey length statement and display of the survey sponsor logo.)

As mentioned before, the break-off rate when showing the progress indicator equals 11.27%. This is not significantly different from the 10.65% break-off rate observed in the survey sponsor logo condition ($\chi^2(1) = 0.0707$; $p = 0.7904$, $n = 822$).

6. Discussion and Conclusion

Drawing mainly on social exchange theory, this study investigated a number of factors potentially capable of influencing web survey response rates and a number of other measures of web survey data quality. This study was fielded within a student population. Obviously, it is necessary to replicate the study for other populations to assess the generalizability of the findings and to fine-tune the methodological strategies. Nevertheless, we believe this study has generated a number of findings that could be relevant and useful to other researchers.

Personalization of the salutation used in the e-mail invitation significantly increased the web survey response rate. The magnitude of the effect (9.4 percentage points) is such that

this effect can be considered practically relevant. We could not find evidence for the hypothesis that personalization would lead to decreased self-disclosure. If anything, the data seems to suggest the opposite. Nevertheless, the cautious approach would be to refrain from using personalization if the survey topic warrants anonymity (e.g., if a survey is conducted on a sensitive private matter), or if the target group is likely to be sensitive to their privacy and anonymity. It is possible that the students were convinced that their university would not harm their privacy, and were therefore not scared off by the personal approach even though the survey covered very intimate matters. In other target populations, such trust may not exist. Therefore, other researchers should thoroughly evaluate whether their survey would benefit from personalization.

In contrast with the expectations, survey length statements were unable to significantly influence the login or response rates. Perhaps a stronger manipulation of the length statements or a more elaborate study design (e.g., a very long statement vs a short statement vs a vague statement) could lead to more interesting results. Also contrary to what was expected, survey sponsor logos displayed on each screen did not decrease the break-off rate. It is possible that the logo had only a marginal impact because of the familiarity of the students with the logo and the university. Other, more general populations would perhaps be more impressed by the sponsorship, leading to a more marked effect of the logo on the break-off rate.

Although the progress indicator was calibrated using pretest data and even though it used the most effective transformation as shown in previous research, it could not significantly decrease the break-off rate. Some evidence was found, however, that the information conveyed by the progress indicator interacts with survey duration information given in the e-mail invitation. This provides some basis for the theoretical reasoning that respondents periodically reassess the cost of participation during the act of answering the survey and that they link this back to what they were told before they commenced the survey. The fact that a large majority of the respondents given the choice requested the progress bar also indicates that people want to be informed about their survey progress. The progress indicator reduced the item missing data rate and the probability that respondents would spontaneously mention that the survey was too long. The latter observation is inconsistent with the time awareness hypothesis, which predicted that progress bars would slow down respondents' time perception. These results warrant further research on progress indicators. Future research should perhaps also pay attention to the possible interactions with survey length statements and actual survey length. The present study indeed suggests that progress indicators do not operate in a vacuum, but they may interact with other design elements.

7. References

Andreasen, A.R. (1970). Personalizing Mail Questionnaire Correspondence. Public Opinion Quarterly, 34, 273–277.

Boser, J.A. (1988). Teacher-Education Graduate Surveys: Variables Related to Response Rate. Journal of Educational Research, 81, 369–373.

Bosnjak, M. and Tuten, T.L. (2003). Prepaid and Promised Incentives in Web Surveys. An Experiment. Social Science Computer Review, 21, 208–217.

Cialdini, R.B. (1984). Influence. How and Why People Agree to Things. New York: William Morrow and Company.

Conrad, F., Couper, M.P., Tourangeau, R. and Peytchev, A. (2003). Effectiveness of Progress Indicators in Web Surveys: It's What's Up Front That Counts. Paper presented at the ASC's 4th International Conference. Warwick University, UK, September 17–19.

Conti, R. (2001). Time Flies: Investigating the Connection Between Intrinsic Motivation and the Experience of Time. Journal of Personality, 69, 1–26.

Cook, C., Heath, F., and 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., Traugott, M.W., and Lamias, M.J. (2001). Web Survey Design and Administration. Public Opinion Quarterly, 65, 230–253.

Crawford, S.D., Couper, M.P., and Lamias, M.J. (2001). Web Surveys. Perceptions of Burden. Social Science Computer Review, 19, 146–162.

DeRouvray, C. and Couper, M.P. (2002). Designing a Strategy for Reducing "No Opinion" Responses in Web-Based Surveys. Social Science Computer Review, 20, 3–9.

Dillman, D.A. (2000). Mail and Internet Surveys: The Tailored Design Method. New York: John Wiley.

Dillman, D.A., Tortora, R.D., and Bowker, D. (1998). Principles for Constructing Web Surveys. Technical Report No. 98-50. Pullman WA: Washington State University Social and Economic Sciences Research Center.

Dillman, D.A., Tortora, R.D., Conradt, J., and Bowker, D. (1998). Influence of Plain Vs Fancy Design on Response Rates for Web Surveys. Paper presented at the Joint Statistical Meetings, Dallas, TX.

Dodd, D.K. and Markwiese, B.J. (1987). Survey Response Rate as a Function of Personalized Signature on Cover Letter. Journal of Social Psychology, 127, 97–98.

Downes-Le Guin, T., Janowitz, P., Stone, R., and Khorram, S. (2002). Use of Pre-Incentives in an Internet Survey. Journal of Online Research. Available at http://www.ijor.org/ijor_archives/articles/Use_of_pre-incentives_in_an_internet_survey.pdf.

Eisinger, R.A., Janicki, P.W., Stevenson, R.L., and Thompson, W.L. (1974). Increasing Returns in International Mail Surveys. Public Opinion Quarterly, 38, 124–130.

Frick, A., Bächtiger, M.-T., and Reips, U.-D. (2001). Financial Incentives, Personal Information, and Drop Out in Online Studies. In U.-D. Reips and M. Bosnjak (eds). Dimensions of Internet Science. Berlin: Pabst Science Publishers, 209–219.

Groves, R.M. and Couper, M.P. (1998). Nonresponse in Household Interview Surveys. New York: Wiley.

Guéguen, N. and Jacob, C. (2002a). Social Presence Reinforcement and Computer-Mediated Communication: The Effect of the Solicitor's Photography on Compliance to a Survey Request Made by E-Mail. CyberPsychology and Behavior, 5, 139–142.

Guéguen, N. and Jacob, C. (2002b). Solicitation by E-Mail and Solicitor's Status: A Field Study of Social Influence on the Web. CyberPsychology and Behavior, 5, 377–383.

Heerwegh, D. (2003b). Explaining Response Latencies and Changing Answers Using Client Side Paradata from a Web Survey. Social Science Computer Review, 21, 360–373.

Heerwegh, D. and Loosveldt, G. (2002). Web Surveys. The Effect of Controlling Survey Access Using PIN Numbers. Social Science Computer Review, 20, 10–21.

Heerwegh, D., Vanhove, T., Matthijs, K., and Loosveldt, G. (2005). The Effect of Personalization on Response Rates and Data Quality in Web Surveys. International Journal of Social Research Methodology: Theory and Practice, 8, 85–99.

Joinson, A.N., Woodley, A., Reips, U.-D. (In Press). Personalization, Authentication and Self-Disclosure in Self-Administered Internet Surveys Computers in Human Behavior (forthcoming).

Lozar Manfreda, K., Batagelj, Z., and Vehovar, V. (2002). Design of Web Survey Questionnaires: Three Basic Experiments. Journal of Computer-mediated Communication, 7.

Pearson, J. and Levine, R.A. (2003). Salutations and Response Rates to Online Surveys. Paper presented at the Fourth International Conference on the Impact of Technology on the Survey Process, September 19, University of Warwick (England).

Porter, S.R. and Withcomb, M.E. (2003a). The Impact of Content Type on Web Survey Response Rates. Public Opinion Quarterly, 67, 579–588.

Porter, S.R. and Withcomb, M.E. (2003b). The Impact of Lottery Incentives on Student Survey Response Rates. Research in Higher Education, 44, 389–407.

Received March 2004
Revised February 2005