Erschienen in: Journal of the American Society for Information Science and Technology ; 58 (2007), 11. - S. 1653-1660
https://dx.doi.org/10.1002/asi.20651

Forced-Response in Online Surveys: Bias from Reactance and an Increase in Sex-Specific Dropout

Stefan Stieger
Core Unit for Medical Education (BEMAW), Medical University of Vienna, Spitalgasse 23, BT87, A-1097 Vienna, Austria. E-mail: stefan.stieger@meduniwien.ac.at

Ulf-Dietrich Reips
Department of Psychology, University of Zurich, Binzmuhlestrasse 14, CH-8050 Zurich, Switzerland

Martin Voracek
Department of Basic Psychological Research, School of Psychology, University of Vienna, Liebiggasse 5, Rm 03-42, A-1010 Vienna, Austria

Due to computer technology, a forced-response can be easily achieved in online questionnaires and is frequently used to gather complete datasets. An Internet-based quasi-experiment was conducted on the student server at the University of Vienna to study the influence of forced-response on dropout, demographic reports, and the content of the results. Forced-response was shown to substantially increase dropout. In addition, forced-response interacted with reported sex in eliminating a naturally occurring sex difference in dropout that was observed for the questionnaire whenever responses did not need to be enforced. Also reported sex turned out to have a mediating effect on time of dropout: Men dropped out earlier than did women. Further analyses revealed a reactance effect, as predicted by reactance theory. It is concluded that data from online questionnaires with forced-response designs are in danger of being hampered by dropout and reactance.

Introduction

A particular feature of surveys conducted over the Internet is that programming technology makes it very easy to force respondents to answer the questions. There is, in fact, no counterpart to this forced-response phenomenon in offline research, except for questionnaires and tests that are presented on a computer. To our knowledge, only few research projects have examined Internet-based forced-response surveys in detail on the basis of empirical data (see Reips, 2002a, p. 248 for a theoretical discussion). This appears surprising, especially as this design is often used in online questionnaires to obtain complete datasets [About 50% of the almost 400 international online surveys and experiments currently archived at the Web Experiment List (http://genpsylabwexlist.unizh.ch/) use this design.]. In an online study aiming at the influence of methodological variables on dropout, O'Neil, Penrod, and Bornstein (2003) showed that forcing a response to demographic questions, which were asked on the last page of an online questionnaire, had a minimal effect on dropout. DeRouvray and Couper (2002) examined the effect of alerting respondents to unanswered questions in a comparison of various strategies aimed at reducing "no-opinion" responses. They found that alerting respondents to unanswered questions reduced the missing data rate. The missing data rate was lowest when respondents were alerted to unanswered questions and were not provided with a "decline to answer" option. DeRouvray and Couper concluded that the reminder prompt is useful in motivating respondents to give complete answers, but they could not demonstrate that data quality was unaffected. It seems likely that strongly encouraging respondents to provide an answer only satisfies the investigators' desire for complete datasets, but does not necessarily provide good data quality.

Certain similarities to the design discussed here are found in research on forced-*choice* designs, whose aim is to force a selection between two or more alternative responses, whereas the issue here is the forced-response itself. That is, forced-choice is a specific item-response format whereas forced-response solely refers to a design feature (i.e., obligatory response to a question to be able to move on to the next question) of survey-based online studies, which is independent from the item-response format. The forced-response mode therefore can include open-ended response formats as well as radio-button selections (i.e., forced-choice items) and still other response formats.

Along the lines presumed by Reips (1997, 2002c), we assumed that a forced-response design may create reactance in some participants. Reactance manifests itself as any behavior or action directly contradicting norms, and it occurs when pressure is exerted on individuals to adopt specific views or attitudes (reactance phenomenon; Brehm, 1966). As a consequence, reactance may either strengthen or weaken expected behavior or effects or may lead respondents to give answers at random. Relatedly, under reactance conditions, responses may show a boomerang effect; that is, the elicited behavior is in the opposite direction of theory-based expectations (Cohen, 1962; von Cranach, Irle, & Vetter, 1965). Forced-response may be particularly receptive to wrong answers in the case of demographic questions because these will reveal personal information of the respondent. Some of our own past work has shown that personalization reduces nonresponse and ameliorates data quality (Frick, Bächtiger, & Reips, 2001; Joinson & Reips, 2007; Joinson, Woodley, & Reips, 2007). Personalization is given through any request to respondents for their personal demographic information and also includes study invitations which contain prospective participants' names, operational designations, or the like—in short, any kind of information that allows personal identification. Consequently, in the present study we set out to ask participants for demographic data of various degrees of revealing personal information such as sex, age, and faculty. Additional information, such as connection speed to the Internet and cost of Internet connection, is derived from information that is automatically collected by the Web server. The influence of forced-response designs on dropout and replies to the questions thus can be researched in its interaction with demographic information.

The present study made use of data available from an online survey on research into jealous reactions to different types of partner infidelity, as investigated by previous evolutionary psychological accounts (Buss, Larsen, Westen, & Semmelroth, 1992; Buss et al., 1999; DeSteno & Salovey, 1996a, 1996b; Harris, 2003; Harris & Christenfeld, 1996b; Schmitt et al., 2003, 2004; Voracek, 2001; Voracek, Stieger, & Gindl, 2001). This line of inquiry concerns aspects of sexual behavior and intimate relationships. Importantly so, it hypothesizes that men and women react differently to different types of infidelity: Men, relative to women, tend to be more upset and distressed by instances of sexual than emotional infidelity of their partners whereas women present the reverse pattern (i.e., relative to men, they tend to be more upset and distressed by instances of emotional than sexual infidelity of their partners). This sex-specific pattern has been found by many studies, including those cited earlier in this paragraph. Given this research background, it appears likely that reactance to forced-response may interact with participants' sex when investigating sex differences in jealous reactions to different infidelity types. The present study therefore examined aspects of the forced-response in more detail and presents their effects on the results of the replies and the dropout rates.

These considerations lead to the following research questions. First, what is the dropout pattern with a forced-response design? Toward this end, forced-response dropout rates were compared with non-forced-response dropout rates. Second, which demographic variables have an effect on dropout and how does this depend on forced-response? And third, does the forced-response design lead to reactance effects on the replies to the questions?

Method

Design

The University of Vienna's student server (UNET server) was used to conduct this study. Using the "closed participant pool technique" (Voracek et al., 2001) that ensures participant membership in a specific population, students of the University of Vienna were randomly selected and contacted via e-mail in February 2000. Basically, in the case of the "closed participant pool technique," the entire population of prospective participants is known and can be enumerated, along with their basic demographic information such as sex, age, and so forth. This type of study setting typically applies to university-student populations. They were asked to participate by clicking a hyperlink that was embedded in the body of the e-mail message. To clearly authenticate participants, an identifier key technique was used. Each invitation e-mail contained a string of characters unique to the participant that was saved to the log file when the questionnaire was accessed. Therefore, we were able to make a connection between data entered in the online questionnaire and data from the online address book of the UNET server and thus were in a position to identify multiple submissions or online deception phenomena such as "gender switching."

To assess reactance, the online questionnaire was programmed in forced-response mode (i.e., an answer had to be given to each question, otherwise the respondent would not advance to the next question). An error page was generated in cases of attempted skipping and displayed a request to completely answer the question. Because these events were logged, it was possible later to identify cases when participants attempted to skip questions (for log file analysis, see Reips & Stieger, 2004). Making an attempt to skip questions was used as a criterion to compare participants who saw an error page and therefore knew it was a forced-response design (forced-response group) with those who never saw an error page because they answered all items without attempting to skip any item (non-forced-response group). Naturally, participants in the non-forced-response group did not know that the questionnaire was programmed in forced-response mode. Because of the nonrandom assignment of participants to groups, the design was quasi-experimental. Although this design is inferior to a classical experimental design because the forced-response and non-forced-response groups may systematically differ by self-selection, we could find no differences with respect to sex, $\chi^2 = 0.27$, $df = 1$, $p = .603$, time between study invitation and study participation, $t = 0.71$,

$df = 4{,}407$, $p = .476$, and whether participants stated a link to their own personal Web site in the online address book, $\chi^2 = 0.41$, $df = 1$, $p = .522$. There was a very small effect (effect size $\eta^2 = .008$) for minimum age, $F(1, 4407) = 36.60$, $p < .001$, as extracted from the inscription number every student receives upon first registration at the university.

Participants

About equal proportions of the participant sample were female (50.4%, $n = 1{,}795$) and male (49.6%, $n = 1{,}768$). An additional 846 participants dropped out before the information about participants' sex was collected. For comparison, the base rates among all UNET account holders at the University of Vienna ("closed pool") were 56.4% females versus 43.6% males at the time of data collection.

Mean reported age of participants was 25.26 years (range = 18–77 years; $Mdn = 24$), relative to a mean age of 26.67 years (range = 18–85 years; $Mdn = 25$) of the total population of UNET account owners at the time of data collection.

Table 1 shows the distribution of participants' affiliation with departments and the overall distribution of students' affiliation with departments at the University of Vienna. Students from the social sciences and basic and integrative sciences were overrepresented in our sample whereas students from the departments of medicine, the humanities, and the natural sciences were somewhat underrepresented. The uneven distribution may have been caused by a higher affinity with the Internet in some subdepartments such as computer science (which is part of the department of social sciences).

For 31.8% ($n = 1{,}275$) of the participants, access was free at the time of participation because they logged in from computer pools within the university. A modem connection via telephone line was used by 42.4% ($n = 1{,}696$) of the participants, given the telephony market in Austria at the time they had to pay a fee that depended on the duration of their call. In addition, 25.8% ($n = 1{,}035$) of the participants accessed the study via their Internet provider (e.g., Chello, UniADSL) and therefore paid for the connection as part of their monthly flat rate.

TABLE 1. Breakdown by faculty of the participants in the online study and the total number of students at the University of Vienna

	Participants	All students
Basic and integrative sciences	1,417 (40.6%)	26,646 (29.3%)
Social science	559 (16.0%)	7,745 (8.5%)
Humanities	393 (11.3%)	18,455 (20.3%)
Medicine	391 (11.2%)	12,569 (13.8%)
Law	390 (11.2%)	11,801 (12.9%)
Natural sciences	283 (8.2%)	12,177 (13.4%)
Catholic theology	41 (1.3%)	1,403 (1.5%)
Protestant theology	9 (0.3%)	238 (0.3%)
Total	3,483 (100.0%)	91,034 (100.0%)

Note. Basic and integrative sciences (e.g., Psychology, Education science, Journalism, Political sciences), Social science (e.g., Business studies, Computer science, Sociology), Humanities (e.g., Studies of languages, Studies of music, History), Medicine (Human medicine, Dentistry), Natural sciences (e.g., Biology, Pharmacy, Mathematics, Physics). The other departments imply only one subdepartment of the same name.

In terms of speed, 35.3% ($n = 1{,}414$) of the participants connected with high (10 MB/s), 22.4% ($n = 896$) with medium (128 KB/s), and 42.3% ($n = 1{,}696$) with low speed (max. 56 KB/s) to the Internet. For the remaining 403 participants, no information about access and speed could be retrieved because the IP address was not associated with a domain name.

Materials

Apart from the demographic items, the questionnaire comprised 10 items. A brief introduction presented the topic of the study without touching on the specific background (partnership, infidelity, sexual vs. emotional jealousy). Demographic data (age, sex, faculty, and whether the respondent already had taken part in an online study) were then requested, followed by four items (two dichotomous ones, two on a 6-point Likert scale) taken from Buss et al. (1992) as well as Harris and Christenfeld (1996a). A guideline for the second part of the questionnaire was then followed by another six items (on a 9-point Likert scale) taken from DeSteno and Salovey (1996b) as well as Buss et al. (1999). The introduction to some items was very similar in wording and consisted of small text pieces of several lines (called scenarios; see Figure 1). Because of the very similar scenarios and fine-grained recording of paradata (Reips, 2002a), we chose a one-screen-one-item design (see Figure 1). A progress display that showed at all times the number of questions remaining was used to give participants feedback as to where he or she stood in the questionnaire at any given moment. Respondents had the opportunity to leave their own e-mail address and add any comments at the end of the questionnaire. The questionnaire took about 6 ($Mdn = 6.08$) min to complete.

Procedure

Overall, e-mail messages were sent to 18,525 students (42% of those who owned a UNET account at the time of data collection). The questionnaire's first page was accessed 4,905 times. A total of 496 datasets had to be discarded for reasons of multiple submissions or gender switching. The remaining 4,409 datasets (23.8% corrected response rate) were used in our statistical analyses. One reason why the participation rate is somewhat low may be that some students did not have the chance to hear about this study because they did not check their e-mail messages on a regular basis. There also is evidence that many e-mail accounts are "sleeping" (i.e., the account holder used the student account for a short time period but then moved to another e-mail provider; Voracek et al., 2001).

At the time, the University of Vienna's online address book contained the participant's name, e-mail address, and inscription number, from which it was often possible to derive the participant's true sex. A possible "gender

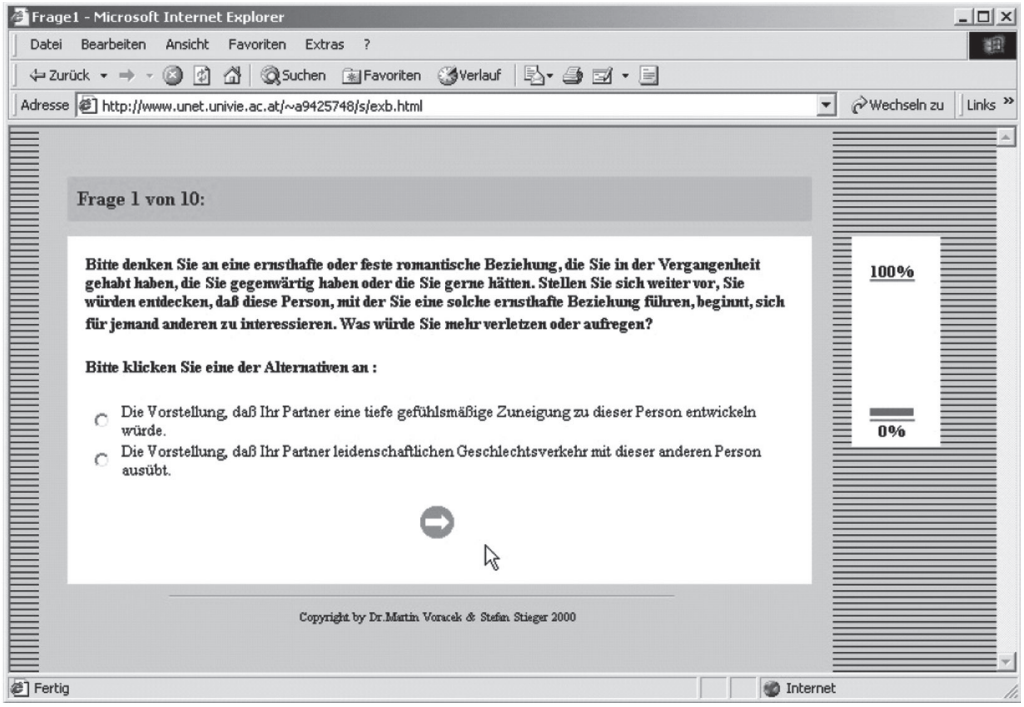

FIG. 1. Screenshot of Item 1.

switching phenomenon" was observed for 4.3% of the men and 4.9% of the women. Other explanations for the mismatch between stated sex and real sex as registered in the university database (e.g., forwarding, response by another member of the same household) cannot be excluded; however, cases with a mismatch were excluded from analyses. The identifier key made it also possible to identify multiple submissions (3.8% overall), which also were excluded from further analyses. Multiple submissions were defined as filling in the questionnaire with an identifier key that had been used before. We excluded all such cases from analyses, although there may have been cases of messages forwarded to someone else.

The forced-response design caused 18.2% ($n = 803$) of the participants to make at least one attempt to skip a question, and 13.7% ($n = 602$) made their first attempt on the second page—the one that asked for demographic data. This means that 75.0% of all attempted skipping took place on this page.

Results

Dropout Caused by the Forced-Response Design

All in all, 803 respondents (of 4,409) produced 863 errors. Of these, 394 (49.1%) respondents dropped out immediately after the first error they provoked. Therefore, this portion of respondents produced a total of 394 errors. Another 121 (15.1%) respondents dropped out later in the online questionnaire, producing a total of 143 errors. The remaining 288 (35.8%) respondents filled in the online questionnaire to the very end and produced 319 errors.

The "natural" dropout portion of respondents (i.e., those who never saw an error page) was 764 (of 4,409). Dropout rates were 18.6% in the group of respondents who did not attempt to skip an item (non-forced-response group) versus 64.1% in the group who did so at least once (forced-response group).

Which Variables Have an Effect on the Dropout Rate?

A Cox regression was calculated by the inclusion method to more precisely determine the factors that really do influence the dropout. The calculation included the following independent variables:

- *Flat rate:* Indicates whether the participant has a dedicated line for which she or he pays a fixed monthly charge, whether the use of the Internet is charged on a time basis (non-flat-rate), or whether no charges at all accrue (direct use of the university's computer pool).
- *Forced-response:* If the respondent attempted to skip an item, an error page that requested completion was generated. Forced-response also was recorded to identify those participants who had attempted to skip an item and to identify the item(s).
- *Speed:* What connection speed to the Internet did the participant's computer have at the time of the survey? The possible choices were "very fast" (i.e., direct at the university), "medium" (e.g., cable modem, ADSL), and "poor" (i.e., via a modem and telephone network).
- *Sex:* The sex of the participant.
- *Age:* The age of the participant.
- *Faculty:* The university faculty to which the participant belongs.

The Cox regression yielded a significant result, $\chi^2 = 262.21$, $df = 15$, $p < .001$. A significant effect on the dropout was shown for the variables of sex and forced-response (cf. Table 2). We will now look more closely at the interacting influences of forced-response and sex on dropout.

Reactance as a Reason for Dropout in a Forced-Response Design

To clarify the trends of the significant differences, Kaplan–Meier survival curves were calculated for the variables of forced-response and sex.

Figure 2 shows that participants in the forced-response group aborted the survey as early as the page displaying the demographic data more frequently than did participants in the non-forced-response group, log-rank test: $\chi^2 = 549.93$, $df = 1$, $p < .001$ (i.e., those who did not generate an error page and so were unaware that they were answering forced-response questions). If the entire questionnaire is considered, the survey was aborted directly after attempted skipping in 76.5% of cases ($n = 394$). This result suggests a reactance effect on dropout, as predicted, and thus might be taken as empirical evidence for the effectiveness of techniques that rely on producing some dropout to achieve better data quality, such as the high-hurdle technique (Reips, 1997, 2002b, 2002c).

TABLE 2. Cox regression with the covariables of sex, flat rate, forced-response, speed, age, and faculty.

	b	*p*	exp(*b*)
Forced-response	1.459	<.001	4.301
Sex	.383	<.001	1.467
Age	.005	.541	1.005
Speed	−.005	.939	0.995
Flat rate	−.028	.704	0.973
Faculty–Catholic theology		.462	
Faculty–Protestant theology	.366	.415	1.443
Faculty–Law	−9.238	.936	<0.001
Faculty–Social sciences	.325	.151	1.385
Faculty–Medicine	.237	.284	1.268
Faculty–Basic and integrative sciences	.347	.130	1.415
Faculty–Humanities	.053	.797	1.054
Faculty–Natural sciences	.274	.241	1.315

Note. b = logistic regression weights (nonstandardized); p = significance level; exp(b) = partial odds ratios.

Sex-Specific Dropout

Figure 3 shows the results of a Kaplan–Meier survival curve with sex as the group variable. It transpires that men aborted the survey more quickly than did women, log-rank test: $\chi^2 = 18.62$, $df = 1$, $p < .001$—a fact already shown by the Cox regression.

Furthermore, sex-specific dropout was examined separately in the forced-response and non-forced-response groups. In addition to those who aborted the survey (i.e., dropout), those who carried on answering it (i.e., carry-on) also were more closely scrutinized for their sex. The criterion for the carry-on condition was that at least the next question after the question where the attempted skipping took place had to be answered.

Table 3 shows that confrontation by an error page significantly increased the dropout. In the non-forced-response group, men showed a 1.78 times greater "risk" of aborting the survey than did women ($OR = 1.78$; 95%-$CI = 1.38, 2.30$). The contents of this effect parameter based on the odds ratio correspond to one based on a Cohen's d value of $d_1 = 0.32$ (Hasselblad & Hedges, 1995). Consequently, the sex difference is significant in the non-forced-response group, which indicates the natural baseline without forced-response intervention. The comparisons in the forced-response group

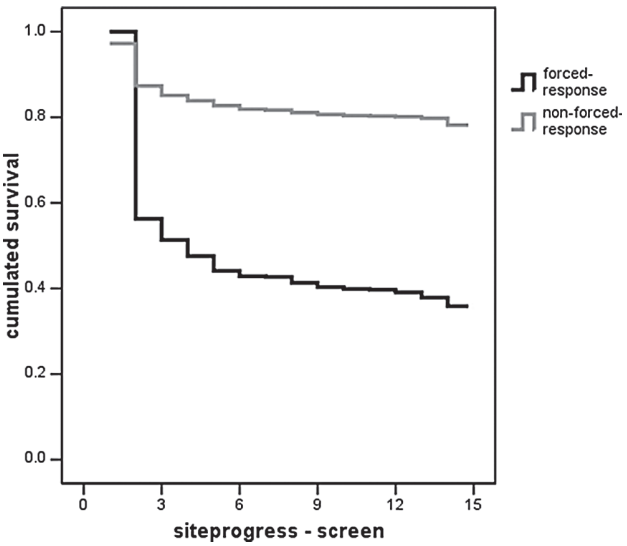

FIG. 2. Kaplan–Meier survival curve with forced-response as the grouping variable.

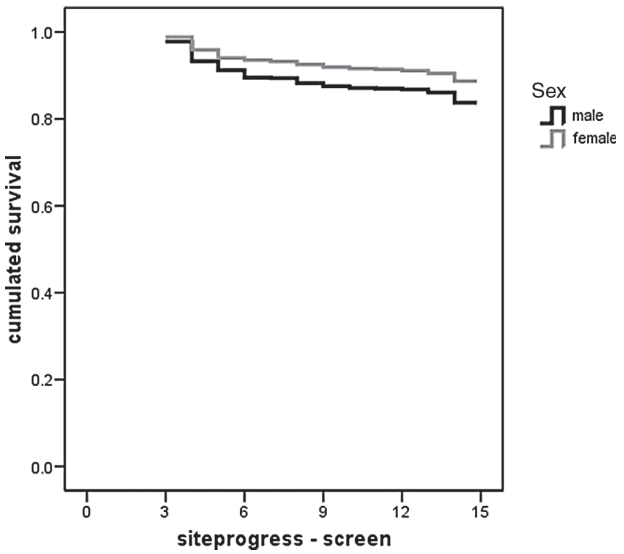

FIG. 3. Kaplan–Meier survival curve with sex as the grouping variable.

TABLE 3. Frequencies in the forced-response and non-forced-response group separated by whether participation was aborted (dropout) or not (carry-on).

	Sex (total)				Sex[a]		
	Female	Male			Female	Male	Total
Forced-response group	237 (52.4%)	215 (47.6%)	Dropout	394 (49.1%)	95 (45.5%)	114 (54.5%)	209 (100%)
			Carry-on	409 (50.9%)	215 (52.6%)	194 (47.4%)	409 (100%)
			Total	803 (100%)			
Non-forced-response group	1,558 (50.1%)	1,553 (49.9%)	Dropout	764 (21.2%)	100 (37.2%)	169 (62.8%)	269 (100%)
			Carry-on	2,842 (78.8%)	1,458 (51.3%)	1,384 (48.7%)	2,842 (100%)
			Total	3,606 (100%)			

[a] Sex in the dropout conditions was determined from the forenames in the online address book of the University of Vienna.

($OR = 1.33$; 95%-$CI = 0.95, 1.86$; $d_1 = 0.16$) and those for the dropouts (forced-response vs. non-forced-response group: $OR = 1.41$; 95%-$CI = 0.97, 2.03$; $d_1 = 0.19$) and the carry-on condition (forced-response vs. non-forced-response group: $OR = 1.05$; 95%-$CI = 0.85, 1.30$; $d_1 = 0.03$) yielded no significant sex differences, indicating that a natural sex difference in dropout was leveled by the forced-response design.

We will now examine whether the forced-response design also affected the quality of responses to the survey.

Does the Forced-Response Design Bias Results?

Next, we examined whether the forced-response design led to divergent results in the subsequent answers (i.e., signs of reactance). According to reactance theory, participants who show reactance no longer answer in a manner consistent with the expected response behavior for the content domain at hand. They either strengthen or weaken expected response behavior or answer at random. Sometimes they even answer in opposition to the theory-based expectations (i.e., boomerang effect). Furthermore, we assume that reactance will most likely affect the item where the error occurred; that is, where attempted skipping took place (i.e., strong reactance). Items following this item will show less and less influence by reactance, with increasing distance.

To examine this point more precisely, the two dichotomous Items 1 and 2 and the Likert-scaled Items 3, 4, and 10 were analyzed in more detail. Items 5 to 9 were dropped because of too few cases of attempted skipping. Figure 4 shows the *OR*s for Items 1 and 2, compared for both groups. If reactance took place, then participants would select answers to Items 1 and 2 randomly, with a resulting *OR* of about 1. Exactly this pattern was found for Items 1 and 2 in the forced-response group, but not in the non-forced response group (see Figure 4). For the item following the item where attempted skipping took place, weaker reactance should occur. Figure 4 shows that the *OR*s for the follow-up item were around those for the non-forced-response group.

For Items 3 to 10, it was possible to analyze a possible reactance effect over time. Therefore, we computed means for Items 3 to 10 separately for the forced-response and non-forced-response groups and for any distance between the analyzed item and the item where attempted skipping took place. For example, if analyzing Item 3, we compared four different distances: (a) the means on Item 3 when attempted skipping took place on exactly this item (distance = 0); (b) the means on Item 3 when attempted skipping took place on the item before, which is Item 2 (distance = 1); (c) the means on Item 3 when attempted skipping took place on Item 1 (distance = 2); and (d) when attempted skipping took place on the demographic questions (distance = 3). This was done for all items. For all distances, a mean difference score

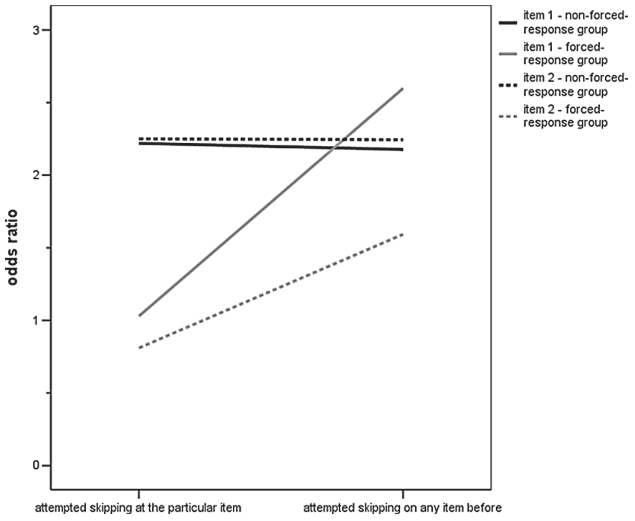

FIG. 4. Reactance effect for Items 1 and 2.

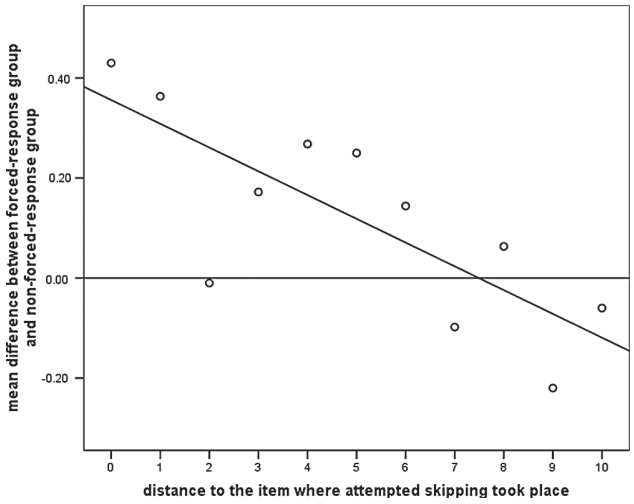

FIG. 5. Reactance effect over time for Items 3 to 10.

between forced-response and non-forced-response groups was computed. A first comparison shows the highest difference between both groups when attempted skipping directly took place at the item of interest, mean difference = 0.43; range 1–5. This difference reached statistical significance, one-sample t test: $t = 5.41$, $df = 2$, $p = .032$. As predicted, the difference declines with temporal distance to the attempted skipping, standardized Jonckheere-Terpstra statistic $= -3.10$, $p = .002$ (see Figure 5).

The analyses conducted here also showed that forced-response has a substantial effect on dropout behavior that is further moderated by the participant's sex. Sex-specific dropout was observed among those participants who did not attempt to skip an item: Males dropped out about twice as frequently as did females. This sex difference was no longer seen among participants in the forced-response group. In summary, the topic of jealousy/sexuality may have had a differential effect on the participants, but the forced-response design leveled out this difference. Further analyses of the answers to Items 1 to 10 revealed a tendency in response behavior that is consistent with reactance (but not the boomerang effect), according to reactance theory. The strongest reactance effect was found on the items where attempted skipping took place.

Discussion

We found that a forced-response design has a major impact on the replies to the items and the dropout. The latter also shows a sex-specific difference, but only for those participants who did not generate an error page (i.e., those who did not attempt to skip an item and were thus unaware of the forced-response character of the survey). This sex difference was not found for the participants in the forced-response group (i.e., the forced-response design distorted the sex-specific dropout pattern observed for this questionnaire).

A comparison of the replies to Items 1 to 10 showed that the use of the forced-response design did invoke a reactance effect that was strongest on the items where the attempted skipping took place and gradually disappeared with distance (i.e., the larger the distance between analyzed item and item where attempted skipping took place, the less the mean difference on replies between the forced-response and non-forced-response groups). This pattern of results was predicted from reactance theory.

Forced-response can be used as a method of implementing the high-hurdle technique (Reips, 2002a, 2002b, 2002c). This technique says that a hurdle placed at the beginning of a Web survey filters out those participants with low motivation (for a checklist of possibilities, see Reips, 2002a, p. 249), leading to lower dropout during the study and possibly improved data quality. A comparison with offline surveying may illustrate the process: In face-to-face communications, anonymity is lost due to the presence of the experimenter. The loss of anonymity produces a greater commitment to participating in the survey. The anonymity of online surveys means that this motivator is lost, so that it makes sense to either increase the motivation or to create a new "hurdle" to encourage persons with low motivation to abort the survey. Similar results were reported by Heerwegh and Loosveldt (2002, 2003), who showed that the need to log in manually to participate in an online study rather than having free access to the study may reduce the response rate but improves the data quality. Here, too, the technique used appears to act as a hurdle.

We suggest that dropout in online studies should be considered in a differentiated way. The "good dropout" of participants who would have a negative impact on data quality due to their poor motivation (e.g., people who are interested only in the design or programming of the online questionnaire) must be distinguished from the "bad dropout" due to the poor design of the questionnaire, programming errors, lack of progress displays, and so on. Thus, the aim for designers of online questionnaires would be to maximize the good dropout and minimize the bad dropout; that is, to hold motivated participants in the online questionnaire and to discard participants with low motivation. Seen from this perspective, a very low dropout—such as that observed in the average laboratory study—even would appear to be suspicious and possibly indicate motivational confounding (Reips, 1997, 2002c).

It would be an obvious idea to compare the present results with the motivation and satisfaction concept by Herzberg, Mausner, and Synderman (1959), which makes a distinction between hygiene factors (i.e., dissatisfiers) and motivators (i.e., satisfiers). Transposed to the study presented here, factors that avoid bad dropout, such as an appealing design or no programming errors, can be considered to be hygiene factors. They need not necessarily improve the quality of the data, but they prevent participants from dropping out prematurely. On the other hand, forced-response and the high-hurdle technique maximize the good dropout. That is, transposed to the concept of Herzberg et al. (1959), it may not increase the motivation directly but cause participants with low motivation to abort the survey prematurely and thus have a positive impact on the data quality. As O'Neil et al. (2003) noted,

"Thus, to further explain dropout, future research should add additional measures […] of motivation" (p. 225). The general tenor in online research—to develop strategies to reduce dropout—should consequently be reconsidered from this new perspective.

Furthermore, an important side result sends a positive message regarding reliability of responses to those considering Internet-based research methods. With the help of the identifier key method, it was possible to detect upper boundaries for gender switching and multiple submissions. These types of unwanted behaviors were rare (<5%).

Future research on reactance following forced-response should use experimental designs to confirm the effects found in this study. Fruitful approaches could include the search for moderators of the reactance effect on dropout as well as on the given answers to the items. To broaden the scope of the present study, another field of research could be effects of reactance when responding to registration forms on Web sites. If people are forced to provide certain demographics to register for a Web-based service, are reactance effects stronger? In addition, do they show signs of reactance such as deception?

Currently, the forced-response design of online studies can be recommended only if (a) it is important to have a complete set of replies from the participants (e.g., for semantic differentials, in the construction of Guttman scales, for multivariant analyses, pairwise comparisons, and for designs in which the questions have to be presented in a specific sequence); (b) a high return is expected, and an increased dropout does not present a problem; and (c) the distribution of participants' sex is not the principal concern of the study.

References

Brehm, J.W. (1966). A theory of psychological reactance. New York: Academic Press.

Buss, D.M., Larsen, R.J., Westen, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. Psychological Science, 3, 251–255.

Buss, D.M., Shackelford, T.K., Kirkpatrick, L.A., Choe, J.C., Lim, H.K., Hasegawa, M., Hasegawa, T., & Bennett, K. (1999). Jealousy and the nature of beliefs about infidelity: Tests of competing hypotheses about sex differences in the United States, Korea, and Japan. Personal Relationships, 6, 125–150.

Cohen, A. (1962). A dissonance analysis of the boomerang effect. Journal of Personality, 3, 75–88.

Cranach, M. von, Irle, M., & Vetter, H. (1965). Zur Analyse des Bumerang-Effektes [Analysis of boomerang effect]. Psychologische Forschung, 28, 535–561.

DeRouvray, C., & Couper, M.P. (2002). Designing a strategy for reducing "no opinion" responses in Web-based surveys. Social Science Computer Review, 20, 3–9.

DeSteno, D.A., & Salovey, P. (1996a). Evolutionary origins of sex differences in jealousy? Questioning the "fitness" of the model. Psychological Science, 7, 367–372.

DeSteno, D.A., & Salovey, P. (1996b). Genes, jealousy, and the replication of misspecified models. Psychological Science, 7, 376–377.

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

Harris, C.R. (2003). A review of sex differences in sexual jealousy, including self-report data, psychophysiological responses, interpersonal violence, and morbid jealousy. Personality and Social Psychology Review, 7, 102–128.

Harris, C.R., & Christenfeld, N. (1996a). Sex, jealousy, and reason. Psychological Science, 7, 364–367.

Harris, C.R., & Christenfeld, N. (1996b). Jealousy and rational responses to infidelity across sex and culture. Psychological Science, 7, 378–379.

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

Heerwegh, D., & Loosveldt, G. (2002). Web surveys: The effect of controlling survey access using PIN numbers. Social Science Computer Review, 20, 10–21.

Heerwegh, D., & Loosveldt, G. (2003). An evaluation of the semiautomatic login procedure to control web survey access. Social Science Computer Review, 21, 223–234.

Herzberg, F., Mausner, B., & Synderman, B. (1959). The motivation to work. New York: Wiley.

Joinson, A.N., & Reips, U.-D. (2007). Personalized salutation, power of sender and response rates to Web-based surveys. Computers in Human Behavior, 23, 1372–1383.

Joinson, A.N., Woodley, A., & Reips, U.-D. (2007). Personalization, authentication and self-disclosure in self-administered Internet surveys. Computers in Human Behavior, 23, 275–285.

O'Neil, K.M., Penrod, S.D., & Bornstein, B.H. (2003). Web-based research: Methodological variables' effects on dropout and sample characteristics. Behavior Research Methods, Instruments, & Computers, 35, 217–226.

Reips, U.-D. (1997). Das psychologische Experimentieren im Internet [Psychological experimenting on the Internet]. In B. Batinic (Ed.), Internet für Psychologen (pp. 245–265). Göttingen, Germany: Hogrefe.

Reips, U.-D. (2002a). Internet-based psychological experimenting: Five dos and five don'ts. Social Science Computer Review, 20, 241–249.

Reips, U.-D. (2002b). Standards for Internet-based experimenting. Experimental Psychology, 49, 243–256.

Reips, U.-D. (2002c). Theory and techniques of conducting Web experiments. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), Online social sciences (pp. 229–250). Seattle, WA: Hogrefe & Huber.

Reips, U.-D., & Stieger, S. (2004). Scientific LogAnalyzer: A Web-based tool for analyses of server log files in psychological research. Behavior Research Methods, Instruments, & Computers, 36, 304–311.

Schmitt, D.P., Alcalay, L., Allensworth, M., Allik, J., Ault, L., Austers, I., et al. (2003). Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands. Journal of Personality and Social Psychology, 85, 85–104.

Schmitt, D.P., Alcalay, L., Allensworth, M., Allik, J., Ault, L., Austers, I., et al. (2004). Patterns and universals of mate poaching across 53 nations: The effects of sex, culture, and personality on romantically attracting another person's partner. Journal of Personality and Social Psychology, 86, 560–584.

Voracek, M. (2001). Marital status as a candidate moderator variable of male–female differences in sexual jealousy: The need for representative population samples. Psychological Reports, 88, 553–566.

Voracek, M., Stieger, S., & Gindl, A. (2001). Online replication of evolutionary psychological evidence: Sex differences in sexual jealousy in imagined scenarios of mate's sexual versus emotional infidelity. In U.-D. Reips & M. Bosnjak (Eds.), Dimensions of Internet science (pp. 91–112). Lengerich, Germany: Pabst.