Implementation of the forced answering option within online surveys: Do higher item response rates come at the expense of participation and answer quality?

Jean Philippe Décieux[1], Alexandra Mergener[2, 3],
Kristina Marliese Neufang[3], and Philipp Sischka[1]

[1]*Research Unit INSIDE, Université du Luxembourg, Luxembourg*
[2]*Section 2.2 "Qualifications, Occupational Integration and Employment", Federal Institute for Vocational Education and Training (BIBB), Bonn, Germany*
[3]*Chair of Methodology and Empirical Social Research, University of Trier, Germany*

Online surveys have become a popular method for data gathering for many reasons, including low costs and the ability to collect data rapidly. However, online data collection is often conducted without adequate attention to implementation details. One example is the frequent use of the forced answering option, which forces the respondent to answer each question in order to proceed through the questionnaire. The avoidance of missing data is often the idea behind the use of the forced answering option. However, we suggest that the costs of a reactance effect in terms of quality reduction and unit nonresponse may be high because respondents typically have plausible reasons for not answering questions. The objective of the study reported in this paper was to test the influence of forced answering on dropout rates and data quality. The results show that requiring participants answer every question increases dropout rates and decreases quality of answers. Our findings suggest that the desire for a complete data set has to be balanced against the consequences of reduced data quality.

Keywords: Forced Answering, Online Survey Research, Dropout, Nonresponse, Response Quality, Faking, Random Answering

The forced answering option

Within the last decade, online surveys have become a well-established method for data gathering and are especially used for asking questions on personal or sensitive topics (Mohorko, de Leeuw, & Hox, 2013). With this increasing use, the number of poorly operationalized online questionnaires also rises. Researchers are paying increasing attention to implementation of these questionnaires and point out that "it is important to pay careful attention to

question format and implementation details" (de Leeuw, Hox, & Boevé, 2015, p. 12). One example of a critical implementation detail is the forced answering (FA) option. As the name implies, under the FA option, respondents are forced to answer each question in order to proceed through the questionnaire. Most online survey software packages allow researchers to enable this option by simply ticking a box. Typically, the idea behind using FA is to avoid missing data (e.g. Valentijn et al., 2015, p. 3). However, avoiding missing data may come at the cost of data quality and unit nonresponse, or participants ceasing to respond to the questionnaire. The aim of the present paper was to test the hypothesized impact of FA on the quantity and, crucially, the quality of data.

An increase in the use of the FA option has been observed (Stieger, Reips, & Voracek, 2007) and some software manuals (Questback, 2013) as well as studies (e.g. Albaum, Roster, & Smith, 2014; Albaum, Roster, Wiley, Rossiter, & Smith, 2010; Albaum, Wiley, Roster, & Smith, 2011; de Leeuw et al., 2015; Roster, Albaum, & Smith, 2014; Valentijn et al., 2015) recommend the use of the FA option because it significantly reduces item nonresponse rates which is considered an indicator of data quality (Groves, 2004). However, researchers have expressed concern that forcing answers might result in increased reactance (Brehm, 1966) because respondents typically have plausible reasons for not answering questions, thus requiring responses leads to dropout and/or poor data quality (Dillman, Smyth, & Christian, 2014; Dillman, Tortora, & Bowker, 1998; Jacob, Heinz, & Décieux, 2013; Stieger et al., 2007).

A vast amount of methodological research has focused on the impact of design aspects of internet-based surveys on data quality and dropout rates (e.g. de Leeuw et al., 2015; Dillman et al., 2014; Galesic & Bosnjak, 2009; Marcus, Bosnjak, Lindner, Pilischenko, & Schütz, 2007; Steinbrecher, Roßmann, & Blumenstiel, 2015). Less attention has been paid to the impact of requiring online survey respondents to answer or enter an appropriate response to every question in order to see the next question and proceed through the questionnaire. This feature – called forced answering (or FA), forced response, or hard reminder – is specific to online surveys and frequently used. One alternative to the FA option is the soft reminder option, a setting in which respondents have the option of leaving a question blank after a warning message has been displayed, thus reducing the possibility of a respondent accidentally skipping an item (DeRouvray & Couper, 2002). There are some similarities between FA and forced-choice designs. However, in the case of forced-choice designs the aim is to enforce respondents to select one response out of two or more proposed alternative responses and thereby create an exclusive answer (single choice without a neutral category). Forced-choice is a specific item-response format, but FA is solely a design feature. That is, FA is completely independent from the item-response format and can include different question formats including forced choice-, multiple response, -and open-ended questions (Stieger et al., 2007).

The use of FA has been criticized from an ethical perspective. Researchers have expressed concern that enforcing answers goes against ethical norms of voluntariness and might result in increased reactance either because respondents

typically have plausible reasons for not answering questions (e.g., not understanding the question, a lack of appropriate categories, or for personally sensitive issues; Dillman et al., 2014; Kaczmirek, 2005; Peytchev & Crawford, 2005; Schnell, Hill, & Esser, 2011; Skarbek-Kozietulska, Preisendörfer, & Wolter, 2012; Stieger et al., 2007).

Concerns have also been expressed in relation to data quality. A "prefer not to answer" (PNA) option may reduce negative effects of FA on data quality (Dillman et al., 2014). Currently, only a few studies have researched the impact of FA on different quality parameters (dropouts, item nonresponse, and quality of the answers). DeRouvray and Couper (2002) and Ganassali (2008), for example, investigated the effect of displaying a warning message when a respondent has skipped a question. They found a decrease of item nonresponse when a warning message was displayed. However, in both scenarios – with and without the alert message – respondents were allowed to leave the question blank. These findings were endorsed by Al Baghal and Lynn (2014), who compared two different types of messages when questions were left unanswered: a reactive probe placed directly after the missed question and a probe at the end of the questionnaire reminding respondents to answer all important questions. Authors concluded that the direct probe was much more effective, reducing item nonresponse rates to a very low level, close to the levels obtained in face-to-face interviews.

O'Neil, Penrod, and Bornstein (2003) showed a minimal effect on dropouts if respondents were forced to answer demographic questions on the last page. Stieger et al. (2007) reported higher dropout rates for an FA group (64.1 percent) compared to a non-forced answering (NFA) group (18.6 percent). In contrast, Albaum et al. (2010) did not find effects of the FA or the neutral PNA options on dropouts, but did find effects on item nonresponse. They showed that FA significantly reduced item nonresponse. Furthermore, no differences in response content for the two groups were found. Roster et al. (2014) used a 2x2x2 factorial design with FA/NFA, PNA/no PNA-option, and low/high sensitive topic in two different cultural groups (U.S. and Hong Kong). They found no effect of FA on dropouts and a significantly lower item nonresponse under FA conditions. The PNA option had no effect on item nonresponse or dropouts. Albaum et al. (2014) examined the influence of FA, topic sensitivity, and the PNA option on retrospectively reported motives (intrinsic or extrinsic) for participating their survey and found no significant effect of FA on participation motives. De Leeuw et al. (2015) showed that explicitly offering a PNA option in an online survey increases the amount of missing data and often decreases reliability of the answers. Allowing respondents to skip a question but following up with a polite probe produced the highest response rates.

The studies of Albaum et al. (2010, 2011, 2014, Roster et al. 2014) as well as of de Leeuw et al. (2015) used only panel populations. Panel populations usually differ from other populations in their likeliness to complete a survey. Motivated to earn panel points or other incentives, they are much more likely to complete the survey (Göritz, 2007). The use of panel populations might explain

the different results in dropout rates compared to other studies (Stieger et al., 2007) and might bias the conclusion by Albaum and colleagues (2010, 2011, 2014, Roster et al. 2014) that FA has no effect on dropout rates.

In our own previous experiments, we did not use panel populations and still did not find a significant effect of FA on the quantity of the "final" dropouts (Décieux, Mergener, Sischka, & Neufang, 2015; Mergener, Sischka, & Décieux, 2015). However, regarding the allocation or timing of the dropouts, we found a significant effect on unit nonresponse: Respondents drop out earlier when they are forced to answer. Our findings may thus confirm the proposition by Stieger et al. (2007), who claimed that FA can be used as a high-hurdle technique that filters out participants with low motivation.

In addition to dropouts and item nonresponse, FA could have an effect on quality of answers. Apart from a comparison of the frequency and distribution of some answer categories under the NFA and FA conditions (Albaum et al., 2010; Décieux et al., 2015; Mergener et al., 2015), no study has examined this crucial issue to date.

This short overview shows that there is no common agreement about the consequences of the implementation of an FA or PNA option on data quality. Inconsistent results have been found for dropout rates. In addition, no study has systematically examined effects of FA formats on answer quality. Given the rise in the popularity of online surveys in general and the use of the FA option in particular, the impact of FA on data quality needs to be addressed. The main objective of this paper is to fill this gap in knowledge for the scientific community as well as practitioners in general. This paper presents analyses of the consequences of the implementation of the FA option on dropouts and answer quality. Our major hypothesis is that forcing respondents to answer will cause reactance, which in turn will decrease data quality and may also result in higher dropouts.

In line with Reips (1997, 2002), Steinbrecher et al. (2015), and Stieger et al. (2007), we assume that FA creates reactance in some respondents with the result that these respondents cease participation in the study earlier and more frequently. Early dropouts can be interpreted in two different ways. In some cases, the early dropout may be a "good dropout" as it filters out unmotivated participants who could have a negative impact on data quality. In other cases, early dropouts may be considered "bad dropouts" as the respondents would have been crucial to obtain valid data and their answers are lost. These "bad dropouts" are the consequence of a poor questionnaire design, programming errors, lack of progress displays, or the lack of freedom to not answer a question. Apart from losing data due to dropout rates, we also assume that forcing the respondents to answer may result in more random answering or faking, impacting data quality.

Forced answering and Reactance

Reactance takes place when an individual (in our case the respondent) is pressured to adopt specific views or attitudes (reactance phenomenon; Brehm, 1966). It appears when an individual's freedom is threatened and cannot be

directly restored. The individual may feel pressured and may subjectively value imposed alternatives as less attractive than denied options (Jonason & Knowles, 2006; Miron & Brehm, 2006, Sischka, Décieux, Neufang, & Mergener, 2015).

The FA option could be experienced by the respondent as a loss of freedom, as (s)he is denied the choice to leave a question unanswered. This results in an internal pressure to disclose information that (s)he actually does not want to offer. This may be felt strongly, especially when sensitive or personal topics are concerned. For these topics, online surveys have become the preferred mode of data collection as the communication with a machine provides greater anonymity and results in an increased willingness to answer (Décieux, 2012; Décieux & Hoffmann, 2014; Sakshaug, Yan, & Tourangeau, 2010; Taddicken, 2008). If respondents experience FA as a loss of freedom, there are some possible actions to restore this freedom. They can restore it by performing a behaviour related to the threat (Brehm & Brehm 1981; Dillard & Shen, 2005). In the context of a survey, this could be quitting the survey.

A phenomenon interrelating with reactance is survey satisficing. This means that with a decrease in motivation (for example, driven by a rising reactance), respondents adopt an answering strategy that is less demanding of mental effort. Participants may begin giving more unreflected answers of lower quality and validity or, if the option is offered, may show an increased tendency to select "don't know" as a response option (Krosnick, Narayan, & Smith 1996; Vanette & Krosnick 2014).

Present study

In the present study, we explore the effect of different response scenarios on completion rates and data quality, a process mediated through reactance, which reduces willingness to proceed through the survey. The following hypotheses were derived for the study:

Hypothesis 1: Forcing respondents to answer all questions results in increased unit nonresponse rates.

Furthermore, we assume that a reactance effect will reduce the willingness to offer valid information. Participants may answer questions randomly or not honestly. This is especially important because in online surveys the barriers for faking responses are lower compared to other data collection methods (Jensen & Thomsen, 2014). Therefore, we presume that:

Hypothesis 2: Forcing respondents to answer all questions results in decreased answer quality (through random answers or faking of answers).

Dillman et al. (2014; 1998) suggested offering a neutral category such as a "prefer not to answer" option or something similar as a way to reduce the risk of increased dropouts when using FA. Offering a neutral "do not know" or

"do not want to answer" option under the mandatory response condition might result in an increase in item nonresponses as respondents have been provided with a no-answer option (Albaum et al., 2014; Albaum et al., 2010; Albaum et al., 2011; de Leeuw et al., 2015; Roster et al., 2014). However, this "no" or "neutral" answer option may neutralise the possible reactance effect by freeing respondents from being forced to answer. Following from this, we suggest that PNA option relieves respondents from the pressure to have to answer all questions. Furthermore, pressure is defused when sensitive or personal questions are concerned. Therefore, we propose:

> *Hypothesis 3: Offering a "prefer not to answer" option weakens the effects of Hypothesis 1 and 2.*

Method

Participants

About 20,000 students at two German universities were contacted via e-mail and invited to participate in our survey. They could start the questionnaire by clicking on a hyperlink embedded in the body of the e-mail message. The sample consisted of 625 students. The mean age was 25 years ($SD = 4.5$) and 331 (53 percent) of the participants were women.

Survey design

We used a topic appealing to young students (i.e. partnership and sexuality) in order to achieve a high participation rate. This topic also offered the opportunity to ask personal and sensitive questions. It was presented in a brief introduction without revealing the actual research interest. Anonymity was assured and no incentives were offered. Subsequently, age and gender were requested. The questionnaire contained approximately 100 items with different types of response formats (e.g. multiple response format, open-ended questions, and five-point Likert scales) with an average response time of 12 minutes (median 11 minutes). In the first half of the questionnaire, students were asked about issues related to partnership (e.g. satisfaction with partnership, conflicts with partner, and characteristics of their partner). The second half of the questionnaire contained questions about their sexual behaviour (e.g. use of contraceptive, masturbation frequency, number of sexual relationships). The end of the questionnaire contained items testing quality of the answers. To analyse the impact of the FA option and test the hypotheses listed above, we randomized participants across three different experimental conditions.

Under condition I, respondents were completely free to choose to answer or to skip any question (non-forced answering, NFA).

Condition II was programmed with the forced answering format (forced answering, FA). If respondents wanted to skip a question (or item), an error page was generated and displayed reminding the participant to provide answers to all questions.

Condition III was also programmed with the FA format, but participants also had the choice of selecting a response from the "prefer not to answer" (PNA) category, such as "not specified" or "I don't want to answer", which was presented in the response option set. This allowed participants to reach the next page and the next question without being forced to give a substantive reply. Condition III forced participants to answer to all questions, but in fact allowed item nonresponse answers (forced answering with PNA-option, FA/PNA).

The participants in all conditions received the same instructions. The different conditions were programmed starting from the third page of the questionnaire. As a result, participants in condition II and III were not made aware that they were not permitted to skip any items or leave any items blank until after they tried to do so. First, a large item battery including 11 different items pertaining to personal beliefs was presented. Then, the respondents were randomized across the three conditions. All attempts of skipping a question were logged. In a second step, this allowed the separation of respondents into those who tried to leave a question blank and those who did not. The latter group thus completed the entire questionnaire "voluntarily". This leads to the research design displayed in Figure 1. The two groups marked with the asterisk and that derived from conditions I and II are explained in more detail below.
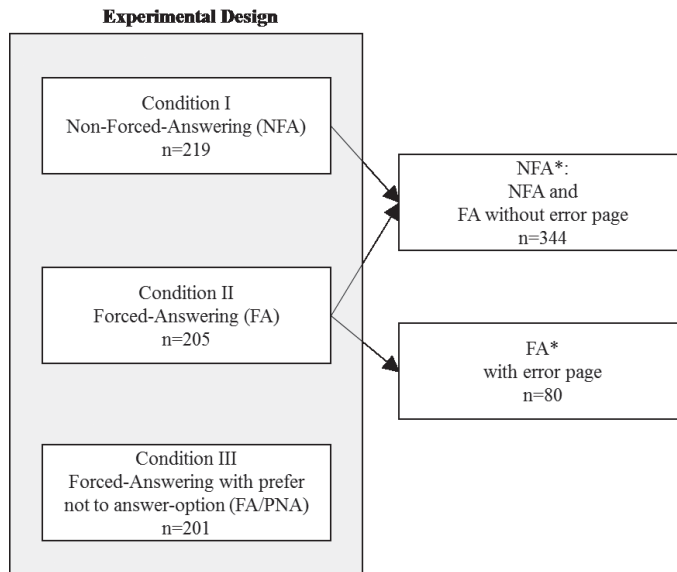

Figure 1. Research design

Measures

To test hypotheses 2 and 3 concerning the quality of the responses, we included special items and scales in the questionnaire (see Appendix for each original item).

Instructed response items. Instructed response items are a tool to identify inattentive or random answers by allowing for only one correct answer (e.g. "We want to test your attention, please click the answer category 'partially agree'"). If respondents choose an incorrect answer, it is assumed that they were not attending to the content of the item. A positive feature of instructed response items is that if a respondent chooses an incorrect answer, there is little doubt that (s)he is responding carelessly or dishonestly (Meade & Craig, 2012). We used items similar to those items used by Marjanovic, Struthers, Cribbie, and Greenglass (2014) and that have just one correct answer. As recommended by Meade and Craig (2012), we interspersed three instructed response items.

Self-reported faking. We also implemented a self-report for faking at the end of the survey. To improve the performance of this indicator and to reduce the possibility of a stigma

of having to confess that questions were not answered honestly, we opted for a neutrally formulated stimulus (Jensen & Thomsen, 2014; Meade & Craig, 2012): "Within this survey we asked some questions that many people would consider very private and highly personal. A common reaction to this is that people do not answer honestly. Therefore, we would like to know: How many questions did you not answer honestly?"

Self-reported sensitivity. Furthermore, we asked participants: "How personal did you find the questions in this survey?" with the response options 1 = "not personal at all" to 5 = "very personal".

Self-reported attention. We also included a self-report measure of attention: "How much attention did you pay to this study?" with response options of 1 = "almost no attention" to 5 = "a lot of attention".

Self-reported interest. Finally, we asked participants: "How interesting did you find the survey?" with the response options 1 = "not interesting at all" to 5 = "very interesting".

Results

Part 1 refers to hypothesis 1, which stipulates that implementing an FA option has a negative effect on dropout rates. Part 2 refers to hypothesis 2, which addresses how the use of FA affects answer quality. Hypothesis 3, concerning the moderating effect of the "prefer not to answer" option on hypotheses 1 and 2, is reviewed in both parts.

Dropout

In terms of dropout rates, there were no significant effects across the three conditions. About a quarter of participants aborted the whole questionnaire. The dropout rates under condition II (FA) were at 14.1 percent higher than for the other two conditions (condition I: 13.7% and condition III: 9%), however, the difference is not significant (Table 1).

Table 1
Dropouts, attempts to skip, and faking across condition

	condition I	condition II	condition III	Total	
	Percentages				Cramer's V
Dropouts	13.7	14.1	9.0	12.3	.071
Attempts to skip	36.1	39.0	16.9	30.9	.210***
Instructed responding (wrong)	12.5	9.5	11.6	11.2	.039
Faking	20.7	24.3	13.0	19.3	.118**
	Means (SD)				Eta²
Item nonresponse	1.28 (5.40)	0 by def.	1.71 (4.96)	1.00 (4.31)	.028***
Sensitivity	3.00 (1.08)	3.02 (1.11)	3.01 (1.12)	3.01 (1.10)	.000
Attention	4.40 (.59)	4.51 (.58)	4.40 (.60)	4.44 (.59)	.007
Interest	3.97 (.77)	4.09 (.65)	4.03 (.68)	4.03 (.70)	.005

Note. *p< 0.1, **p<0.05, ***p<0.01.

To precisely interpret their dropout-behaviour, we needed to distinguish between respondents who tried to skip at least one question under condition II and those who did not. Those who attempted to skip a question got a warning message reminding them to answer each item. Those who never tried to skip a question were never presented with this error page. The majority of respondents never tried to leave a question blank and were thus never presented with the warning message, so they did not notice that they were in condition II, the FA condition, and therefore were not truly forced to answer any questions. However, 80 (39 percent) of the total of 205 respondents in the FA condition II attempted to skip a question and thereby realized that they were required to answer all questions.

To gain a better understanding of the dropouts, it is important to divide the group of respondents under condition II into two subgroups: Those who tried to not answer at least one question and received a warning message, and those who answered all the questions without trying to leave any blank. This procedure results in three groups for analysis:

1. "Not forced to answer" (NFA*) includes all respondents that were not explicitly forced within their participation. This respondents from condition I ($n=219$) are included in this group, as are those respondents from condition II who did not see a warning message ($n=125$; total NFA group $n=344$).
2. "Forced to answer" (FA*) includes the respondents of condition II that were provided a warning message and thus were made aware and reminded that they were required to answer questions to continue through the questionnaire ($n=80$).
3. "Forced to answer with 'prefer not to answer'" (FA/PNA) includes the respondents from condition III ($n=201$).

Table 2 shows a significant increase of dropout rates when the respondents were really forced, or were told they are required to respond to items to continue. The percentage of those who quit the survey increased to 35 percent.

Table 2
Dropouts, attempts to skip, and faking by group (NFA, FA*, and FA/PNA)*

	NFA*	FA*	FA/PNA	Total	
	Percentages				Cramer's V
Dropouts	9.0	35.0	9.0	12.3	.264***
Instructed responding (wrong)	11.6	8.0	11.6	11.2	.033
Faking	22.0	25.0	13.0	19.3	.115**
	Means (SD)				Eta²
Item nonresponse	.81 (4.35)	0 by def.	1.71 (4.96)	1.00 (4.31)	.017***
Sensitivity	3.03 (1.09)	2.88 (1.08)	3.01 (1.10)	3.01 (1.10)	.001
Attention	4.46 (.58)	4.38 (.63)	4.40 (.60)	4.44 (.59)	.003
Interest	4.04 (.72)	3.94 (.70)	4.03 (.68)	4.03 (.70)	.002

Note. *p< 0.1, **p<0.05, ***p<0.01.

Here we detected a negative effect of the FA option on the quantity of our data. Our analysis showed that forcing respondents to answer each question resulted in their cancelling the survey more often (higher dropout rates, as predicted by H1). In the two non-forced conditions, NFA* and FA/PNA, dropouts did not differ significantly (as predicted by H3).

We also analysed the development of the dropout rates during the survey process when forced condition applied for the first time, and focussing on the exact point when questionnaire procedure was aborted. Figure 2 shows the relative cumulative percentage of dropouts of each question (Figure 2).
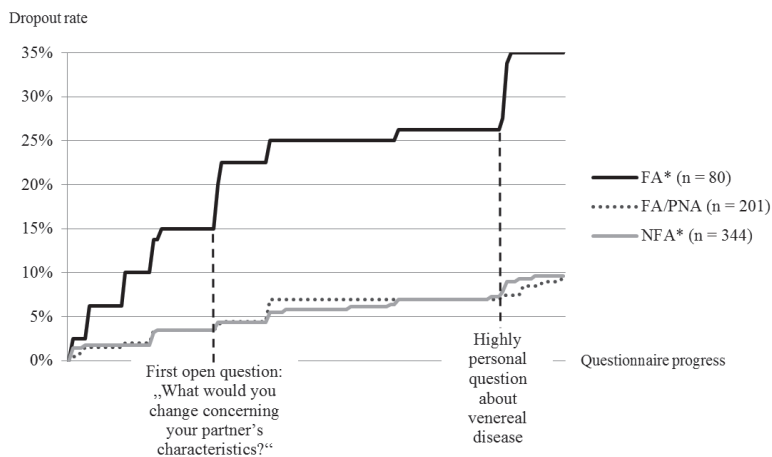

Figure 2. Dropout process by group

Before analysing dropout rates under different conditions, it is important to know that the most dropouts occur at the first hurdle at the beginning of the survey. At this point, participants had not been forced to answer any items. The respondents were then asked to rate 11 different items concerning their personal beliefs. About 16 percent of all respondents aborted the questionnaire at this point. Nearly all participants of the NFA* group who continued the survey after this complex item battery completed the survey (in both groups NFA* and FA/PNA only about 9% dropped out additionally).

Figure 2 illustrates that dropout patterns are very similar for NFA* group and FA/PNA condition. However, these patterns are quite different from the dropout pattern for the FA* group, in which forcing participants to answer appears to encourage "bad drop-outs". Dropout rates increased to a larger degree when respondents were forced to answer the questions. A relatively conspicuous first jump in dropouts (+ 7.5 percentage points) in this group was detected when respondents were given an open-ended, personal question. Respondents were asked to state which characteristics of their partner they would like to change.

The second biggest peak in dropout rate occurs when a highly personal question concerning own experiences with venereal diseases was asked (+ 8.7 percentage points). Thus, respondents in the FA condition aborted the survey even though they had completed it nearly to the end.

Answer quality

Following our hypotheses 2 and 3, we analysed whether forcing the respondents to answer results in a reactance effect of giving random answers or faking more often. However, there was no significant difference in responses to the instructed response items across the groups. Corresponding to that, 94.9 percent of participants indicated that they gave "quite a bit of attention" or "a lot of attention" to the questions asked. A possible explanation is that participants were paying a lot of attention due to the survey's topic. This is substantiated by the finding that 80.9 percent of participants stated that they found the survey "interesting" or "very interesting". The second indicator for answer quality is the likeliness to fake within the three conditions. Our hypothesis is that forcing the respondent to an answer results in faking more often than when answers are not forced, especially when questions addressing highly personal and sensitive topics are asked. Indeed, 68.9 percent of respondents stated that they found the questions "very personal", "personal", or "a bit personal". In the question that asked how often respondents did not answer honestly, the percentages of fakers was at highest in the FA* (25 percent) group and at lowest in the FA/PNA (13 percent) condition. For self-reported sensitivity, attention or interest (see Tables 1 and 2) we found no significant differences among the three conditions or the three groups, but we detected a significant effect on the mean number of item nonresponse. Although this number is per definition 0 when respondents are forced to answer each question, item nonresponse significantly increases to a mean of 1.67 when a PNA option is offered (and an answer of "prefer not to answer" is considered as missing data).

Discussion

Only one-third of our FA subsample was actually forced to answer one or more questions. Most of the respondents answered even the highly personal questions without trying to skip the question. One reason for that might be the fact that the sample consists of students, who are used to answering to questions in surveys. Another explanation might be that students are so used to FA format that they expect it and automatically answer all questions. This interpretation is supported by the fact that only about 34 percent of the respondents in condition I made use of their option to not answer every question. In a next study, we plan to make the NFA group aware of the "not forced answering" response possibility. It would also be interesting to test the use of a soft reminder against the use of a hard reminder.

Regarding dropout rates, our findings show that after the first hurdle in form of a large item battery (before the start of the conditions), low-motivated participants are filtered out ("good dropouts"). Those who are generally willing to finish the questionnaire will stop proceeding through the questionnaire at some point if they are forced to answer ("bad dropouts"). Highly personal and open-ended questions are especially likely to provoke such a premature abort (H1).

Concerning the quality of the answers, our data did not show any significant effect for the instructed response items. These results might be different in surveys with low-involvement topics. In addition, our results show that the percentages of respondents who reported faking responses differed significantly between the FA* (highest faking) and FA/PNA (lowest faking) conditions. Therefore, our data show the proposed significant reactance effect on faking behaviour when respondents are forced to answer to questions (H2). Furthermore, in the FA* group there are far more dropouts. These respondents already showed reactance by quitting their participation and by quitting they did not get to the faking self-report question. Therefore, our results indicate that voluntariness is associated with honesty.

De Leeuw et al. (2015) show that offering a PNA option decreases the consistency of answers and scales, hence data reliability is decreased by giving respondents an easy way out of the question-answer process and this leads to satisficing (de Leeuw et al., 2015). We also detected the tendency to make use of this category resulting in an higher item nonresponse, but in comparison to Albaum et al. (2010) who suggested that the use of a PNA category "can erode the advantages of forced answering" (p. 291), we conclude that the effects of this neutral response category are more positive. Depending on question type, item topic, and question difficulty as well as on respondent characteristics (e.g. involvement, experience or knowledge), respondents may genuinely not know the answer. Not giving the participants the opportunity to express their nescience will also threaten answer quality as participants are forced to fabricate answers (Converse, 1976; Converse & Presser, 1986; Jacob, Heinz, & Décieux, 2013). In addition, results in the present study showed that the PNA response category can neutralize the shown reactance effect by freeing respondents from the mandate to answer each question to continue the questionnaire (H3).

In conclusion, FA has a negative influence on answer quality and dropout rates by systematically producing "bad dropouts" in form of a reactance effect, which results in a decreasing power of estimates. This reactance effect has to be balanced against higher item-response rates.

Limitations

Some limitations in the research presented here have to be addressed. As Stieger et al. (2007), we used a survey topic that is more personal than most. Our results could be different with topics that are less personal. Additionally, the PNA-option could have different moderation effects of FA on dropouts, faking, and item nonresponse if it available only after a warning message is displayed. Furthermore, we did not address the issue of reliability of the answers.

Another limitation of our study is the specificity of our research population (students), which might be more open and experienced in surveys. An interesting approach for a further study would be to test the hypotheses of this study within a more heterogeneous sample. We suggest that this would enforce the detected reactance effect. Furthermore, previous studies have shown that the barriers for faking in web surveys are very low and that it is difficult to estimate the extent of faking. Because of those concerns, we cannot guarantee that the faking self-report is an accurate estimate. However, following Jensen and Thomsen (2014) we propose this self-report as an indicator for detecting lies. A final limitation of our study is that the instructed response items might also to some extent lead to reactance, because the items might give respondents the feeling of being tested (Meade & Craig, 2012).

References

Al Baghal, T., & Lynn, P. (2014). *Using motivational statements in web instrument design to reduce item missing rates in a mixed-mode context*. (Understanding Society Working Paper Series, 2014-02). Colchester, United Kingdom: Institute for Social and Economic Research, University of Essex.

Albaum, G., Roster, C. A., & Smith, S. M. (2014). Topic sensitivity and research design: effects on internet survey respondents' motives. *Asia Pacific Journal of Marketing and Logistics, 26*, 147-161.

Albaum, G., Wiley, J., Roster, C. A., & Smith, S. M. (2011). Visiting item non-responses in internet survey data collection. *International Journal of Market Research, 53*, 687-703.

Albaum, G., Roster, C. A, Wiley, J., Rossiter, J., & Smith, S. M. (2010). Designing web surveys in marketing research: does use of forced answering affect completion rates? *The Journal of Marketing Theory and Practice, 18*, 285-294.

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

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

Converse, J. M. (1976). Predicting no opinion in the polls. *Public Opinion Quarterly, 40*, 515-530.

Converse, J. M., & Presser, S. (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Newbury Park: SAGE Publications.

De Leeuw, E. D., Hox, J. J., & Boevé, A. (2015). Handling Do-Not-Know Answers: Exploring New Approaches in Online and Mixed-Mode Surveys. *Social Science Computer Review,* 1-15.

Décieux, J. P., Mergener, A., Sischka, P., & Neufang, K. (2015). Higher response rates at the expense of validity? The consequences of the implementation of forced answering options within online surveys. *Paper presented at the General Online Research (GOR) Conference in Collogne 2015*.

Décieux, J. P. (2012). *Modeeffekte bei Onlineumfragen Ein multivariater Methodenvergleich unter Zuhilfenahme eines Propensity Score Matchings*. München: Akademische Verlagsgemeinschaft.

Décieux, J. P., & Hoffmann, M. (2014). Antwortdifferenzen im Junk & Crime Survey: Ein Methodenvergleich mit goffmanscher Interpretation. In M. Löw (Ed.), *Vielfalt und Zusammenhalt. Verhandlungen des 36. Kongresses der Deutschen Gesellschaft für Soziologie in Bochum und Dortmund 2012*. Frankfurt am Main: Campus.

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.

Dillard, J. P., & Shen, L. (2005). On the Nature of Reactance and its Role in Persuasive Health Communication. *Communication Monographs, 72*, 144-168.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method, 4. eds*. New Jersey: John Wiley & Sons.

Dillman, D. A., Tortora, R. D., & Bowker, D. (1998). *Principles for constructing web surveys.* Paper presented at the Joint Meetings of the American Statistical Association.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly, 73*, 349-360.

Ganassali, S. (2008). The Influence of the Design of Web Survey Questionnaires on the Quality of Responses. *Survey Research Methods, 2*, 21-32.

Göritz, A. S. (2007). Using online panels in psychological research. In A. Joinson, K. McKenna, T. Postmes & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology* (pp. 473-485). Northamptonshire: Oxford University Press.

Groves, R. M. (2004). *Survey errors and survey costs*. Hoboken, New Jersey: John Wiley & Sons.

Jacob, R., Heinz, A., & Décieux, J. P. (2013). *Umfrage: Einführung in die Methoden der Umfrageforschung*. München: Oldenbourg Verlag.

Jensen, C., & Thomsen, J. P. F. (2014). Self-reported cheating in web surveys on political knowledge. *Quality & Quantity, 48*, 3343-3354.

Jonason, P. K., & Knowles, H. M. (2006). A unidimensional measure of hong's psychological reactance scale. *Psychological reports, 98*, 569-579.

Kaczmirek, L. (2005). Web Surveys. A Brief Guide on Usability and Implementation Issues. *Usability professionals*, 102-105.

Krosnick, J. A., Narayan, S.. S., & Smith, W. R. (1996). Satisficing in surveys: Initial evidence. *New Directions for Evaluation, 70*, 29–44.

Marcus, B., Bosnjak, M., Lindner, S., Pilischenko, S., & Schütz, A. (2007). Compensating for Low Topic Interest and Long Surveys A Field Experiment on Nonresponse in Web Surveys. *Social Science Computer Review, 25*, 372-383.

Marjanovic, Z., Struthers, C. W., Cribbie, R., & Greenglass, E. R. (2014). The Conscientious Responders Scale: A New Tool for Discriminating Between Conscientious and Random Responders. *SAGE Open, 4*, 1-10.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological methods, 17*, 437-455.

Mergener, A., Sischka, P., & Décieux, J. P. (2015). To force or not to force. That is the question!": Die Auswirkungen des Einsatzes von Forced-Response-Fragen auf die Qualität der Befragungsergebnisse. *Verhandlungen der Kongresse der Deutschen Gesellschaft für Soziologie: "Routinen der Krise – Krise der Routinen" 37*.

Miron, A. M., & Brehm, J. W. (2006). Reactance theory: 40 years later. *Zeitschrift für Sozialpsychologie, 37*, 9-18.

Mohorko, A., de Leew, E., & Hox, J. (2013). Internet Coverage and Coverage Bias in Europe: Developments Across Countries and Over Time. *Journal of official statistics, 29,* 609-622.

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.

Peytchev, A., & Crawford, S. (2005). A typology of real-time validations in web-based surveys. *Social science computer review, 23*, 235-249.

Questback. (2013). *Enterprise Feedback Suite EFS Survey 9.1*. Köln.

Reips, U.-D. (1997). Das psychologische Experimentieren im Internet. In B. Batinic (Ed.), *Internet für Psychologen* (pp. 245–265). Göttingen: Hogrefe.

Reips, U.-D. (2002). Theory and Techniques of Conducting Web Experiments. In B. Batinic, U.-D. Reips & M. Bosnjak (Eds.), *Online social sciences* (pp. 229–250). Seattle: Hogrefe & Huber.

Roster, C. A., Albaum, G., & Smith, S. M. (2014). Topic sensitivity and Internet survey design: A cross-cultural/national study. *Journal of Marketing Theory and Practice, 22*, 91-102.

Sakshaug, J. W., Yan, T., & Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly, 74*, 907-933.

Schnell, R., Hill, P. B., & Esser, E. (2011). *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

Sischka, P., Décieux, J. P., Neufang, K., & Mergener, A. (in press). Der Reaktanzeffekt oder: Warum Gesundheitskampagnen scheitern können.

Skarbek-Kozietulska, A., Preisendörfer, P., & Wolter, F. (2012). Leugnen oder gestehen? Bestimmungsfaktoren wahrer Antworten in Befragungen. *Zeitschrift für Soziologie, 41*, 5-23.

Steinbrecher, M., Roßmann, J., & Blumenstiel, J. E. (2015). Why Do Respondents Break Off Web Surveys and Does It Matter? Results From Four Follow-up Surveys. *International Journal of Public Opinion Research, 27*, 289-302.

Stieger, S., Reips, U.-D., & Voracek, M. (2007). Forced-response in online surveys: Bias from reactance and an increase in sex-specific dropout. *Journal of the American Society for Information Science and Technology, 58*, 1653-1660.

Taddicken, M. (2008). *Methodeneffekte bei Web-Befragungen*. Köln: Herbert von Halem Verlag.

Valentijn, P. P., Vrijhoef, H. J. M., Ruwaard, D., de Bont, A., Arends, R. Y., & Bruijnzeels, M. A. (2015). Exploring the success of an integrated primary care partnership: a longitudinal study of collaboration processes. *BMC Health Services Research, 15*, 1-18.

Vannette, D. L. & Krosnick J. A. (2014). Answering Questions: A Comparison of Survey Satisficing and Mindlessness. In A. le, C. T. Ngnoumen, E. J. Langer (Eds.), *The Wiley Blackwell Handbook of Mindfulness* (pp. 312-327). Chichester, West Sussex: Wiley Blackwell.

Appendix

Instructed response item

"We want to test your attention. Please click the answer category 'partially agree'".

"Please check 'rather correct'".

"Please click 'I rather agree'".

Note. Each instructed response item was embedded in a larger item battery with five response categories.

Original items' text in German:

Self-reported faking: „Wir haben Ihnen im Laufe der Befragung einige Fragen gestellt, die viele Personen als hochpersönlich einstufen würden. Eine gängige Reaktion auf solche persönlichen Fragen ist nicht ehrlich zu antworten. Daher würden wir gerne wissen: Auf wie viele Fragen haben Sie bei dieser Umfrage nicht ehrlich geantwortet?"

Self-reported sensitivity: „Wir haben Ihnen im Laufe der Befragung einige Fragen gestellt, die viele Personen als hochpersönlich einstufen würden: Wie sensibel fanden Sie selbst die Fragen in dieser Umfrage?

Self-reported attention: „Wie viel Aufmerksamkeit haben Sie der Beantwortung der Fragen geschenkt? "

Self-reported interest: „Wie interessant fanden Sie diese Umfrage? "