Using Adaptive Questionnaire Design in Open-Ended Questions: A Field Experiment

Matthias Emde[1], Marek Fuchs[1]
[1] Institute of Sociology, Darmstadt University of Technology
Residenzschloss, 64283 Darmstadt (Germany)

Abstract
Previous research on open-ended questions revealed several factors influencing the answers provided by respondents in web surveys. Among others, motivating instructions have proven to elicit long and rich responses. Furthermore the visual design of the answer box size influences the answers provided by respondents. Small boxes seem to pose a lower response burden and therefore reduce item-nonresponse. Larger answer boxes increase burden but at the same time motivate respondents to provide more extended responses. Based on this evidence determining the ideal box size seems to require a trade-off balancing item-nonresponse and the length of the answers. Web surveys offer opportunities to deal with this dilemma and optimize answers to open-ended questions. In a survey among university freshman students we conducted an experiment using auto-adjusting answer boxes. In a second experiment, we used the amount of information respondents typed into the response box to an initial open-ended question to assign them later in the survey a custom-size answer box to a second open-ended narrative question. Results suggest that this adaptive enhancement in the visual design of text boxes improves data quality. Especially the adaptive assignment of answer boxes reduced item-nonresponse for those respondents who did not answer the first question while prolific writers typed more detailed answers into the second box. Results are discussed in light of the general question whether questionnaire design in general and visual design in particular should be used to influence and instruct respondents on how to answer an question or whether an adaptive design should be used that adjusts elements of the questionnaire to the behaviour of respondents.

Key Words: web surveys, narrative open-ended question, adaptive design, answer box

1. Introduction

Growing access and popularity make web surveys a more and more common survey mode. Web surveys are self-administered, costs are relatively low and responses are not affected by the presence of an interviewer. Also, due to the computer-assisted character, web surveys offer advantages in term of automated routing and branching and allow design opportunities in order to accomplish a more engaging survey experience for respondents than classic paper pencil surveys (Shropshire et al., 2009). With respect to open-ended narrative questions the literature is not fully conclusive; however, open-ended questions in web surveys seem to yield longer responses than in a paper-based survey mode (Kwak & Radler, 2002; Denscombe, 2008).

Answers to web survey questions are affected by the visual design of the questionnaire (Tourangeau, Couper, & Conrad, 2007). Visual design aspects like labels (Schwarz, Grayson & Knäuper, 1998), pictures (Toepoel & Couper, 2011), arrangement of answer categories (Smyth, Dillman, Christian, & Mcbride, 2009) the use of colors (Tourangeau et al., 2007) and the order of answer categories (Tourangeau et al., 2007) affect responses. Various studies have focussed on the design of associated answer boxes with respect to open-ended questions. Among others, the size of the input box (Christian & Dillman, 2004; Smyth et al., 2009), any label or mask attached to the box (Fuchs, 2007; Fuchs, 2009b; Couper et al., 2011) or the number of boxes offered in association with a question (Fuchs, 2009a; Keusch, 2012) have been proven to affect the quality and the extent of responses to open-ended questions.

In this paper, we will focus on ways to further improve the visual design of answer boxes for narrative open-ended questions. The results of two studies will help explore the influence of different answer box designs on data quality and the willingness of respondents to answer open-ended narrative questions. Moreover we will also test an adaptive answer box designs where groups of respondents were assigned custom-sized answer boxes based on answers provided to a previous open-ended in the same survey.

2. Research Questions

As web surveys are self-administered it is the respondent who has to transform the subjective response into its final format. For open-ended questions guidance on how the answer has to be formatted can be provided via verbal instructions and visual cues (Couper, Kennedy, Conrad, & Tourangeau, 2011). Formatting problems predominantly occur when open-ended questions are used to access information that require a specific format like the date of birth for instance (Christian, Dillman, & Smyth, 2007). To avoid formatting problems such answer boxes can be restricted to a specific format. By contrast, answer boxes to narrative open-ended questions are usually not restricted and respondents are not limited in neither a specific format nor the response length (Couper et al., 2011).

This study focuses on narrative open-ended questions which are especially used to gain detailed spontaneous responses and to explore new topics where exhaustive answer categories cannot be provided. Furthermore open-ended narrative questions allow respondents to freely answer in their own words and likewise produce long and rich responses. On the other hand respondents might have difficulties to articulate their own views properly which is why responses to these questions are affected by personal characteristics like age, sex and education (Desncombe, 2008; Stern, Dillman & Smyth, 2007). Also, responses to open-ended narrative questions are often prone to higher rates of item-nonresponse (Desncombe, 2008; Reja, Manfreda, Hlebec, & Vehovar, 2003). Despite these disadvantages and regardless the fact that open-ended narrative questions require extensive coding (Reja et al., 2003) they are a powerful tool to get valuable information from respondents.

Unfortunately item-nonresponse to open-ended narrative questions is considerable which tempted some web survey designers to use error messages or warnings in case a respondent chose not to answer a question. However, these messages increase respondent frustration and break-offs (Dillman & Smyth, 2007). More effective were follow-up probes to get respondents to answer the question or to increase the length or their

responses (Holland & Christian, 2009; Oudejans & Christian, 2010). Even though, these probes have proven effective, the increase in the extent of the responses is rather small. By contrast, motivational statements associated with the questions wording were reported to be more effective (Smyth et al., 2009).

Besides verbal instructions and follow-up probes, the visual design of answer boxes also affects responses to narrative open-ended questions. Christian and Dillman (2004) and Israel (2010) altered the answer box size and found longer answers and more reported topics to large answer boxes in comparison to smaller answer box sizes in paper-based self-administered surveys. Smyth, Dillman and Christian (2007) as well as Smyth, Dillman, Christian and Mcbride (2009) demonstrated the same pattern in a web survey setting, where at least late respondents, who were assumed to be less motivated, provided more extensive responses to larger answer boxes.

Generally it is assumed that respondents construe the answer box size as auxiliary information when interpreting the question and generating their answer. If a large box is displayed, respondents are assumed to perceive the task at hand as broader as compared to a small box. The scope of the concept addressed in the question is assumed to be more extensive and they get the impression that they have to provide a longer answer. By contrast, when providing a small box the scope of the question is limited and respondents are encouraged to shorten their answer. Consequently, small boxes seem to pose a lower response burden while large answer boxes increase this burden and therefore might provoke increasing item-nonresponse. Choosing the ideal box size for an open-ended narrative question therefore seems to require a trade-off balancing item-nonresponse and the extent of the answers provided.

According to this reasoning we expected an influence of the answer box size on the information provided by respondents. While a small answer box was supposed to pose a lower response burden and therefore produces less item-nonresponse lager boxes pose a higher burden to answer. Therefore we expected lower item-nonresponse rates to smaller box sizes in comparison to larger boxes. At the same time larger answer boxes indicate that a more detailed answer is expected encouraging respondents to type longer responses and report more topics. Thus, response length and reported topics were expected to increase with larger box sizes.

We further tested the use of counters indicating the number of characters left for the respondent while typing the answer. Just like the box size, the counter is important for the respondent as auxiliary information when interpreting the question and generating the answer. According to this, we expected the counter to work as a permanent stimulus that motivates respondents to match their answers with the expected response length and provide more topics, characters and characters per topic when exposed to an open-ended narrative question with a counter.

Even though the rapid development of the internet and internet data collection methods offers various graphic design features, previous studies mostly rely on a static web survey design using pre-defined input fields. In making use of web technologies it is possible to support and motivate respondents by visual design features like dynamic answer boxes. By using a dynamic box it is possible to start with a small box implicating a low response burden while motivating respondents to provide a longer answer by continuously increasing the box size with the respondent typing the answer. Each time the respondent completed a line of text in the answer box, an additional empty line was added at the end

of the answer box. Overall, when using a dynamic box we expected less item-nonresponse in comparison to a large answer box and more information in comparison to a small answer box.

In addition to these dynamic features that are assumed to enhance data quality in open-ended narrative questions web surveys offer the opportunity to adapt answer boxes to groups of respondents, according to their response behaviour in a previous open-ended question. An adaptive design changed the survey or at least the visual design of some questions, in order to offer a survey design that fits best the motivation and behaviour of groups of respondents. By using an adaptive assignment of answer box sizes for groups of respondents we expected to reduce item-nonresponse for those respondents who usually would not answer and encourage prolific writers to write even more detailed answers. Overall the adaptive box design with customized answer box sizes was presumed to reduce item-nonresponse and increase the quantity and quality of responses to narrative open-ended questions.

3. Methods

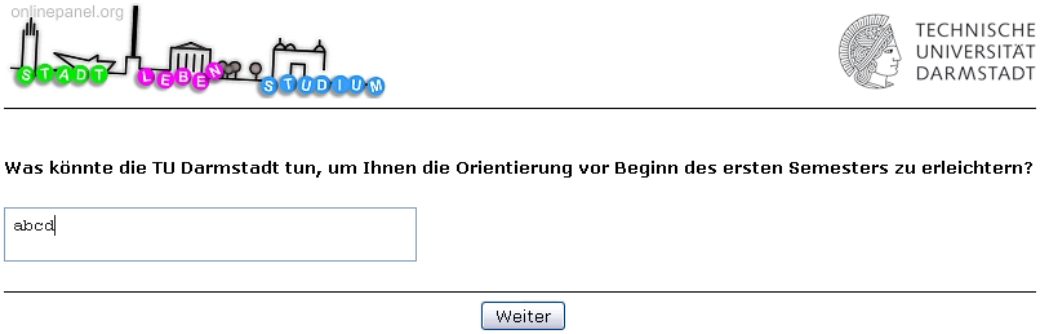
Study 1
In October 2011 4.342 university freshman students were invited by e-mail to join the online access panel at Darmstadt University of Technology. As an incentive a lottery drawing of five book vouchers was offered and a reminder was send after seven days of field work. Overall, 673 Students registered at the panel website and took the web survey on the topic of being a freshman student. No hitherto existing panel members were invited to the survey. The response rate amounts to 16 percent (AAPOR RR2).

Respondents were randomly assigned to one of four different versions of the same open-ended question in a between-subjects design: *"In what ways could the university assist new students?"* The design of the first experiment is displayed in Figure 1. Respondents were randomly assigned to either a small (a) or a large (b) answer box. In the third experimental condition we added a counter (c) to an answer box of the same size like the small box in experimental group (a). The counter was displayed after the respondent clicked into the field in order to isolate visual cues of the answer box and the displayed counter. The counter was displayed under the box using an initial start value of 250 characters. The initial value of the counter was about three times higher than the number of characters that fit into the visible answer box (2 lines of 42 characters each). The experimental condition (d) included a dynamic growing box. Starting with the same size as the small box, an additional empty line was added to the box for every row of text the respondent completed. Overall, answers were not limited by the box sizes and respondents were able to write up to 5,000 characters in every condition.

The validity of narrative open-ended question is difficult to measure. We gauge response quality of narrative open-ended questions by item-nonresponse, the number of characters and the number of topics reported. A topic was defined (same as Smyth et al., 2009) as a subject that answered the question and was independent of all other subjects mentioned within the same response. To measure elaboration, we use the characters per topic ratio indicating how many characters were used by the respondents to account every single topic.

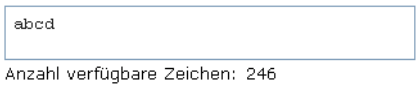
Figure 1: Four different answer box designs in Study 1

(a) Small box



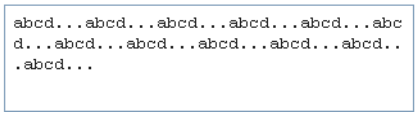
(b) Large box

(c) Small box with counter



(d) Dynamic box after 4 typed characters

(d) Dynamic box after 84 typed characters



Notes. small box = 2 rows with 42 characters each; large box = 8 rows with 42 characters each, small box with counter = 2 rows with 42 characters each; dynamicl box = starting with 2 rows with 42 characters each

Study 2

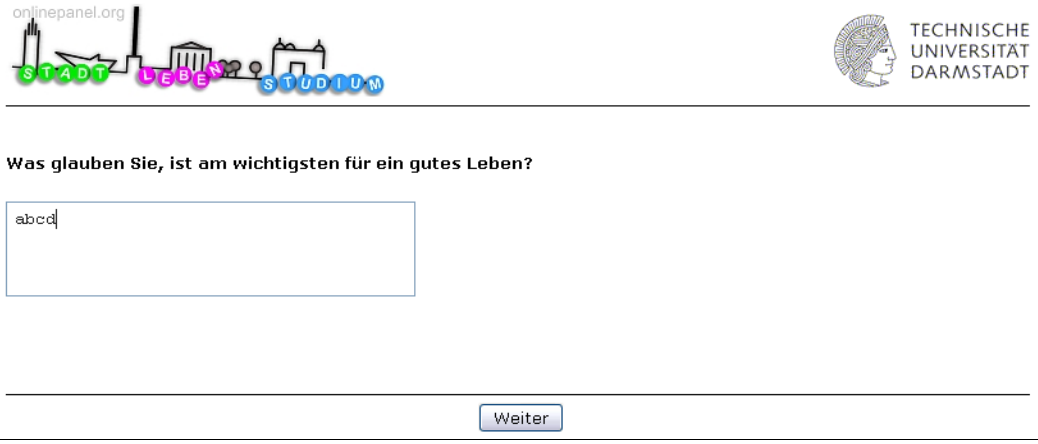
The second study was also carried out in the Darmstadt University of Technology online access panel in November 2011. For this study, 2446 Panel members were invited to a web survey on student life satisfaction. Actually 907 panelists took the survey resulting in a Response rate of 41% (AAPOR RR2).

In Study 2 respondents were exposed to an open-ended narrative question *"In your opinion, what is most important to lead a good life?"* Later in the questionnaire respondents were again asked an open-ended narrative question: *"In your opinion, what is essential to succeed in academic studies?"* After the first question 25 percent of the respondents were randomly assigned to a control group in a between-subjects design.

Figure 2: Adaptive answer box design in Study 2

First question

(a) Initial box (mid-size)



Second question

(b) Control group box (mid-size)

(c) Small box: no characters were typed into the initial box

(d) Mid-size box: up to 84 characters were typed into the initial field

(e) Large box: 85 and more characters were typed into the initial field

Notes. Small box = 2 rows with 42 characters each; Mid-size box = 4 rows with 42 characters each; Large box = 8 rows with 42 characters each

Respondents in the control group got the same answer box size for the second questions (b) like in the first question (4 rows with 42 characters each). The remaining 75 percent of respondents built the experimental group for the adaptive design. In the experimental group respondents who did not respond to the first open-ended question were assigned to a small answer box (c) in the second open-ended question (2 rows with 42 characters each). Respondents who filled up to half of the first answer box (up to 84 characters) were assigned to the same box size (d) for the second open-ended question (4 rows with 42 characters each) while respondents who filled more than half of the answer box to the first question (85 characters or more) were assigned to a answer box twice the size (e) of the first question (8 rows with 42 characters each). Response quality was measured like in study 1 considering item-nonresponse, the number of characters, the number of topics and the number of characters per topic.

4. Results

Study 1

In Study 1 respondents were randomly assigned to one of four designs: a small answer box, a large answer box, a small answer box combined with a counter and a small dynamic growing answer box (see Figure 1). We assumed that the reduced response burden associated with a small input box would apply to all three designs starting with a small box while the large box would provoke a higher item-nonresponse. In comparison to the small answer box, we expect the large box, the small box with a counter and the small dynamic box to motivate respondents to provide longer answers. Results on the proportion of item-nonresponses and the number of characters respondents typed into the open-ended questions are summarized in Table 1.

Table 1: Item-nonresponse, characters, topics and characters per topic by answer box type

	(a) Small box	**(b) Large box**	**(c) Small box & counter**	**(d) Dynamic box**
No response	33%	33%	36%	33%
Response	66%	66%	63%	66%
Characters	92 [d]	107 [(c),d]	101 [(b)]	80 [a,b]
Topics	1,3 [(c)]	1,3 [(d)]	1,2 [(a),d]	1,1 [(b),c]
Characters per topic	74	84	89	73
N	162	174	161	174

Notes. Pairwise Chi²-tests: a, b, c, d = p < .05; (a), (b), (c), (d) = p < .1. Letters denote significant differences to the indicated columns. Outliers were excluded at two standard deviations above the group mean for characters, topics and characters per topic

Contrary to expectations, the box size did not significantly influence item-nonresponse. For the small answer box as well as for the large box two out of three respondents (66%) answered the open-ended question. Thus the larger response box did not seem to pose a higher response burden in our study. Starting with the same small answer box size similar results were found for the small box with a counter (63%) and the dynamic growing box (66%).

The result concerning the length of the responses to the narrative open-ended question was inconclusive. While the small box yielded 92 characters on average (outliers were excluded at two standard deviations above the group mean), the large answer box version yielded 107 characters on average. However, this difference was not statistically significant. The average length of the answers for the version with the counter was higher as well (101 characters) compared to the small box (again not statistically significant). By contrast the dynamic growing box produced significantly shorter responses (80 characters) in comparison to the three other conditions. This difference was statistically significant in comparison to the large answer box and to the box with a counter ($p < .05$).

Slightly varying results were to be noticed for the number of topics reported by the respondents. The large box did not increase the number of reported topics compared to the small answer box while the number of reported topics for the counter version (1.2 topics, $p < .1$) differed on a marginally significant level while the dynamic answer box (1.1 topics, $p < .05$) differed significantly.

The number of characters per topic exhibited a more conclusive result. Respondents to the large answer box (84 characters per topic) used more characters per topic compared to the dynamic box (73 characters per topic, $p < .1$). The box with the counter produced 89 characters per topic and therefore significantly more characters per topic compared to the small box (74 characters per topic, $p < .1$) and the dynamic box (73 characters per topic, $p < .05$).

Study 2

In Study 2 members of the Darmstadt University of Technology online access panel were invited to a survey on student life satisfaction. Respondents first had to answer the initial question "*In your opinion, what is most important to lead a good life?*" using a mid-sized answer box (4 rows, 42 columns). The number of characters typed into this initial answer box was then used to assign respondents to one out of three versions of a second open-ended question *"In your opinion, what is essential to succeed in academic studies?"* with an adaptive answer box size based on the length of the response to the first question (see Figure 2).

Table 2: Item-nonresponse, characters, topics and characters per topic by first initial and second/adaptive question

	(a) First/initial question	**Second/adaptive question**	
		(b) Control group	**(c) Experimental group**
No response	13%	13%	15%
Response	87%	87%	85%
Characters	48	46[c]	53[b]
Topics	2.4[b,c]	2.1[a]	2.0[a]
Characters per topic	21[c]	22[c]	29[a,b]
N	800	209	591

Notes. Comparing initial question, control group and experimental group: one-way ANOVA, F-test; a,b,c = $p < .05$. Letters denote significant differences to the indicated columns. Outliers were excluded at two standard deviations above the group mean for characters, topics and characters per topic

Overall we found no differences in item-nonresponse in the second study (Table 2). Neither in comparing responses to the initial box size and responses in the control group nor in comparing the control group to the experimental group. As in Study 1, the box size is not affecting item-nonresponse in the adaptive design setting.

The overall length of the answer did not vary between the initial (48 characters) and the control group (46 characters) indicating that both questions are comparable in terms of the length of the answers they elicit. Answers to the adaptive box design in the experimental group were significantly longer (53 characters) in comparison to the control group ($p < .05$). Thus, in total we accomplished on average more characters using the adaptive box design. When comparing the number of topics for both questions, a significant decrease ($p < .05$) between the first initial question (2.4 topics) and the control group (2.1 topics) as appeared. Also, there was a significant difference to the experimental group (2.0 topics, $p < .05$). Because of the differences between the answers to the initial question and the control group in the reported topics these differences might be caused by the question itself and not necessarily by the differences in the answer box design. Therefore the characters per topic are a better measure for elaboration and data

quality as it resembles the answer differences affected by the visual design and not by the question wording. Even though the adaptive design has no assets in item-nonresponse, the overall number of characters and the proportion of characters per topic were significantly higher ($p < .05$) for the adaptive answer box design in the experimental group as compared to the initial question.

Table 3: Number of characters, topics and characters per topic by answer box type in the second adaptive question

	Control group (mid-size box)	Small box	Mid-size box	Large box
1st initial question				
(I) Characters initial question	50	-	33	120
(II) Topics initial question	2.6	-	2.1	3.6
(III) Characters per topic initial question	20	-	20	40
2nd adaptive question				
Characters adaptive answer box	46	28	47 (I)	114
Topics adaptive answer box	2.1 II	1.8	1.9 II	2.7 II
Characters per topic adaptive answer box	22	18	24 III	50 (III)
N	165	17	374	74

Notes. Pairwise Chi²-tests comparing the adaptive answer box with the answers given to the first initial question: I, II, III= p < .001; (I), (II), (III) = p < .05. Roman numbers denote significant differences to the indicated rows in the upper part of the table. Outliers were excluded at two standard deviations above the group mean for characters, topics and characters per topic and the number of cases (N) including all respondents who answered the second open-ended question

Table 3 summarizes results of a simulation. Results for the first/initial question are reported separately for the groups of respondents who were assigned to the three answer box designs for the second/adaptive question. Thus, we are able to compare response behavior for the first/initial and the second/adaptive question for nonrespondents to the first/initial question (small box), for respondents who provided a condensed answer (mid-size box) and for prolific writers (large box). This enables an assessment of the impact of the assigned box size on response behavior of respondents in the respective groups (control, small box, mid-size box and large box). In addition, responses by members of the control group are displayed. In the control group a slight decrease in the number of characters occurred when comparing the responses to the first (50 characters) and second open-ended question (46 characters, not significant). While the decrease in the number of topics (2.6 vs. 2.1 to the second question) is significant ($p < .001$) the number characters per topic did not differ significantly (20 vs. 22 characters per topic).

The small answer box was applied to all nonrespondents to the first question. Compared to respondents who answered the initial question (mid-size box and large box), respondents in this group typed in fewer characters (28) and topics (1.8) and also the characters per topic proportion (18) is the lowest for this group. However, we could motivate 17 former nonrespondents to answer the second open-ended question while item-nonresponse to the second question rarely occurred.

The mid-size box was applied to respondents who filled up to one half of the answer box to the first/initial question. In comparison to their answers to the first question, respondents provided significantly more characters to the second adaptive questions. The number of characters increased significantly ($p < .05$) from 33 to 47 and the number of characters per topic elevated from 20 to 24 ($p < .001$). Only the number of topics decreased from 2.1 to 1.9 ($p < .001$) for the mid-size answer box group.

The large answer box was presented with the second question if respondents provided a long response to the first open-ended question (85 or more characters). In this group the number of topics reported also decreased compared to the first/initial question (from 3.6 to 2.7 ($p <.001$), likely due to the subject of the second open-ended question). The number of characters provided in reaction to the large box (114 characters) in the adaptive box design setting did not change significantly in comparison to the initial question (120 characters). Therefore prolific writers still provided extensive information in the adaptive answer box design. However, the number of characters per topic for this group (50 characters per topic) was significantly higher ($p <.05$) than in the initial question (40 characters per topic) indicating that the large box increased elaboration for this group (plus 25% characters per topic). This increase was more pronounced in comparison to the control group (plus 10%) and the mid-size box group (plus 20%).

5. Summary and discussion

The purpose of both of our studies was to find ways to lower item-nonresponse and to increase the extent of information obtained in narrative open-ended questions in web surveys. In the first study we tested different techniques to lower the response burden and motivate respondents to provide long and rich responses. Previous studies have shown an influence of the answer box size on data quality in narrative open-ended questions. Indeed we could demonstrate that different box sizes affect the number of characters, the number of topics and number of characters per topic provided by respondents. Overall responses to the large box yielded more characters and characters per topic in comparison to a small answer box. However, contrary to expectations no differences in item-nonresponse between small and large boxes occurred suggesting that a large box-size did not pose a higher response burden in our study. The relatively small differences in the box size in our experiments might be an explanation for that. The fact that we experimented with a highly educated and motivated student sample might be another explanation.

In order to motivate respondents to provide high quality answers we further added a counter to a small answer box. Based on the small answer box the response burden was assumed to be rather low while respondents should be encouraged to provide extensive responses by the counter continuously indicating the number of characters left. The answer box with a counter produced slightly more characters and significantly more characters per topic in comparison to the answer box of the same size without a counter while no differences in item-nonresponse appeared. Therefore we assume the counter to be a valuable enhancement to every open ended question.

The results for the dynamic growing answer box revealed a shorter response length and fewer reported topics. So far, we have no compelling explanation for this finding. The dynamic answer box should at least accomplish results comparable to the small box. We hypothesize that the resizing of the dynamic answer box while typing might confuse

respondents and even lower their motivation to answer as they might feel misled. Overall, the dynamic box in our study seems to have a negative influence on data quality. Therefore further testing on dynamic growing answer boxes is necessary. Regarding the question-answer process the dynamic box has the disadvantage that the formatting of a response has already been finished before the respondent notices the appearance of additional rows in the dynamic design. Altering the box size during the typing process might therefore not be ideal.

In the second study we tested a design that avoided this problem of altering box sizes during the question-answer process by using custom-sized answer boxes for groups of respondents. Based on the number of characters respondents typed into the answer box of a first/initial open-ended question we assigned them later in the survey to a custom-size answer box to a second open-ended question in order to reduce item-nonresponse and increase the amount of information. Even though we gain a few respondents by the adaptive design, there is no substantial influence on item-nonresponse when using adaptive box sizes since we at the same time observed some additional nonrespondents to the second question that have in fact answered the first/initial question. However, the use of adaptive box sizes produces more information with respect to the number of characters and the characters per topic in comparison to the control group with a mid-size answer box. Compared to the first initial question, the adaptive box design yielded more characters per topic. This result is even more remarkable as respondents tend to write shorter answers to open-ended questions, when they appear later in the questionnaire (Galesic & Bosnjak, 2009).

The results from both our studies provide preliminary evidence that the visual design of answer boxes influences data quality in open-ended narrative questions. Overall, using larger answer boxes, adding a counter or using adaptive box sizes increases the number of characters per topic to open-ended narrative questions. Also, results provide preliminary support for the effectiveness of a web survey design that adapts the visual design of survey questions to the motivation and capabilities of the respondent. While previous studies in the design of open-ended narrative questions aimed to enhance the effectiveness of design features that were meant to influence the response behaviour (in particular of not so motivated respondents) the adaptive design changes the visual design of the questionnaire in order to get the most out of the respondent accepting his or her motivation and capabilities.

A limitation of both of our experiments is the use of a student sample. Students are a highly educated population and may be more able and willing to answer narrative open-ended questions. Whether our results will hold for web surveys in the general population remains to be tested.

References

AAPOR (2011). *Standard definitions. Final dispositions of case codes and outcome rates for surveys. Revised 2011*: AAPOR.

Christian, L. M., & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly, 68*(1), 57-80.

Christian, L. M., Dillman, D. A., & Smyth, J. D. (2007). Helping the respondents get it right the first time: the influence of words, symbols, and graphics in Web surveys. *Public Opinion Quarterly, 71*(1), 113-125.

Couper, M. P., Kennedy, C., Conrad, F. G., & Tourangeau, R. (2011). Designing input fields for non-narrative open-ended responses in web surveys. *Journal of Official Statistics, 27*(1), 65-85.

Denscombe, M. (2008). The length of responses to open-ended questions. A comparison of online and paper questionnaires in terms of a mode effect. *Social Science Computer Review OnlineFirst, 26*(3), 359-368.

Dillman, D. A., & Smyth, J. D. (2007). Design Effects in the Transition to Web-Based Surveys. *American Journal of Preventive Medicine, 32*(5), 90-96.

Fuchs, M. (2007). *Asking for Numbers and Quantities: Visual Design Effects in Web Surveys*. Paper presented at the Paper presented at the AAPOR Annual Conference.

Fuchs, M. (2009a). Differences in the visual design language of paper-and-pencil surveys vs. Web surveys. A field experimental study on the length of response fields in open-ended frequency questions. *Social Science Computer Review, 27*(2), 213-227.

Fuchs, M. (2009b). *Dynamic screen design in open-ended questions. A field-experiment on the visual design language of Web surveys*. Paper presented at the 1st International Workshop of Internet Science.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a Web survey. *Public Opinion Quarterly, 73*(2), 349-360.

Holland, J. L., & Christian, L. M. (2009). The influence of topic interest and interactive probing on responses to open-ended questions in web surveys *Social Science Computer Review, 27*(2), 196-212.

Israel, G. D. (2010). Effects of answer space size on responses to open-ended questions in mail surveys. *Journal of Official Statistics, 26*(2), 271-285.

Keusch, F. (2012). *Open-ended questions in web surveys: one large vs. ten small boxes*. Paper presented at the AAPOR.

Kwak, N., & Radler, B. (2002). A comparison between mail and Web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics, 18*(2), 257-273.

Oudejans, M., & Christian, L. M. (2010). Using interactive features to motivate and probe responses to open-ended questions. In M. Das, P. Ester & L. Kaczmirek (Eds.), *Social and rehavioral research and the internet* (pp. 215-244). New York: Routledge.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in Web questionnaires. *Metodoloski zvezki, 19*, 159-177.

Schwarz, N., Grayson, E. G., & Knäuper, B. (1998). Formal measures of rating scales and the interpretation of question meaning. I*nternational Journal of Public Opinion Research, 10*, 177-183.

Shropshire, K. O., Hawdon, J. E., & Witte, J. C. (2009). Web survey design. Balancing measurement, response, and topical interest. *Sociological Methods Research, 37*(3), 344-370.

Smyth, J. D., Dillman, D. A., & Christian, L. M. (2007). *Improving response quality in list-style open-ended questions in Web and telephone surveys*. Paper presented at the Annual Conference of the American Association for Public Opinion Research.

Smyth, J. D., Dillman, D. A., Christian, L. M., & Mcbride, M. (2009). Open-ended questions in web surveys. Can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly, 73*(2), 325-337.

Smyth, J. D., Dillman, D. A., Christian, L. M., & Stern, M. J. (2006). Effects of using visual design principles to group response options in Web surveys. *International Journal of Internet Science, 1*(1), 6-16.

Stern, M. J., Dillman, D. A., & Smyth, J. D. (2007). Visual design, order effects, and respondent characteristics in a self-administered survey. *Survey Research Methods, 1*(3), 121-138.

Toepoel, V., & Dillman, D. A. (2011). Words, Numbers, and Visual Heuristics in Web Surveys: Is There a Hierarchy of Importance? *Social Science Computer Review, 29*(2), 193-207.

Toepoel, V., & Couper, M. P. (2011). Can verbal instructions counteract visual context effects in web surveys? *Public Opinion Quarterly, 75*(1), 1-18.

Tourangeau, R., Couper, M. P., & Conrad, F. G. (2007). Color, labels, and interpretive Heuristics for response scales. *Public Opinion Quarterly, 71*(1), 91-112.