# Autosuggest in Dynamic Forms:

# The Effect of Response Formats on Response Time and Data Quality

Master's Thesis of:

**Sébastien Orsini, B.Sc.**
Institute of Psychology
Department of Cognitive Psychology and Methodology
University of Basel

November 2008

Thesis Supervisors:

**Peter Schmutz, M. Sc.**
Department of Cognitive Psychology and Methodology, University of Basel

**Prof. Dr. Klaus Opwis**
Department of Cognitive Psychology and Methodology, University of Basel

## Abstract

An experiment was conducted to test the performance of autosuggest, a response format based on AJAX (Asynchronous JavaScript and XML) technology, versus traditional response formats in the field of web forms or web surveys. The aim of the study was to examine, if autosuggest could excel radio buttons and drop-down boxes. In a web survey 412 participants had to search for target values with one of these three input types in long as well as in short word lists. Response times, number of answered items as well as cognitive load were assessed. Data show that radio buttons are useful for short word lists, whereas autosuggest excels drop-down boxes in long word lists. Regarding the dropout rate and coding effort none of the response formats differed. The amount of answered items is lower with autosuggest, foreshadowing that people are less familiar with it. Differences in response times influenced also participant's subjective feeling of time pressure. These results demonstrate that the use of autosuggest can improve the completion of web forms and web surveys, but depends on the list length and its familiarity.

# Autosuggest in Dynamic Forms:

# The Effect of Response Formats on Response Time and Data Quality

## *Introduction*

Nowadays user cannot surf in the Internet without abandoning personal information in online registrations, surveys and forms. A web designer's task is to make sure, that these interactions are conceived as easy as possible avoiding user's annoyance. Therefore, it is very important to provide the right input type to register users' answers.

Usual elements for a fill out of web forms are check boxes, text boxes, radio buttons, and drop-down boxes. Checkboxes in comparison with the other three response formats are utilized to select more than one answer from a list of response options, but quite often just one option is required. Therefore, it is a format that is relatively rarely seen in general web forms. In contrast to checkboxes, text boxes differ from the other response formats in its unlimited possibilities of answer options. The user is free to response whatever he likes to. Tough, it is desired to give users a selection of possible answers from which they can choose, so that webmasters are able to reduce the variability of answers (e.g., see Reja, Manfreda, Hlebec, & Vehovar, 2003).

An additional input format implements this requirement. The Web 2.0 movement reaches the whole World Wide Web (WWW) with the intention to enhance the interactivity. This implies amongst other a new form of input format, the so-called *autosuggest*. Its appearance is the same as a plain text box, but, since it is programmed with Asynchronous JavaScript and XML (AJAX), there is a constant exchange between browser and server in the background. For a technical overview about AJAX see Garrett (2005). As soon as the first letter is entered in the box, a drop-down list appears and shows a list of possible answers. The more letters are entered, the more of the drop-down list will be reduced. The user is still free to answer whatever he likes to, but the automatic limitation with possible propositions gets him to choose a given option. This specific kind of input type is mostly applied on search

boxes. But the autosuggest text input becomes more and more popular in forms and arises the question how successful is autosuggest as a new form of response format.

The goal of this study is to examine situations in which users fill out a form. Particularly, when users are only allowed to select one item per response format. Therefore, radio buttons, drop-down boxes, and autosuggest were compared in order to find out, if autosuggest can outclass traditional input formats.

*Visual design of surveys*

The knowledge that visual features of a questionnaire affect users' answers was particularly explored in paper and pencil and web surveys. Traditionally, paper and pencil questionnaires have been viewed as depending only upon words and sentences being the source of influencing users response. However, it has become clear that users' response also depends upon information communicated through visual effects. Smith (1995) for instance, lists several examples of unintentional layout changes producing differences in surveys. Empirical research on visual design suggests that numbers, symbols and graphics influence how users answer both paper and web surveys (Redline & Dillman 2002; Christian & Dillman 2004; Tourangeau, Couper, & Conrad 2004). In medical research, there are also several examples comparing visual features on response formats. Especially visual analog scales (VAS) were analyzed and compared with classical rating scales (e.g., see Brunier & Graydon, 1996; Grant et al., 1999). In recent years, these visual analog scales have become subject of online researchers (Couper, Tourangeau, & Conrad 2006), pointing out that different input formats influence participants' behavior. According to Christian, Dillman, and Smith (2007) for example a manipulation of the date format shows that writing effective forms may depend on the presentation of the answer format much more than the question wording itself.

*Radio buttons vs. drop-down boxes*

In the Internet, there is a wide common usage of radio buttons and drop-down boxes with little rules for their application. For example to request the gender, radio buttons as well as drop-down boxes are used. In longer word lists there is also no consistency. Technically seen both provide the selection of one option from a list, whereas drop-down boxes require less space on the screen. However, radio buttons require fewer mouse clicks to complete a request in a form than drop-down boxes. Online researchers have faced these response format decisions whether to use radio buttons or drop-down boxes. Beaumont, James, Stephens, and Ullman (2002) see the gain of drop-down boxes in needing less space than a list of radio buttons. According to them the advantage of radio buttons is the visibility of all options at once. Radio buttons are preferred to drop-down boxes as long as the lists are limited to four or five items. On the other hand, Miller and Jarett (2001) propose to use radio buttons only for lists with up to four items. Magee, Straight, and Schwartz (2001) as well as Dillman (2007) recommend avoiding the use of dropdown boxes, because of nonresponse and misuse.

Couper, Traugott, and Lamias (2001) argue in their work that radio buttons lead to less missing data and invalid answers than typing a number in text boxes. Heervegh and Loosveldt (2002) analyzed the differences between radio buttons and drop-down boxes. They discovered that no differences in number of nonsubstantial answers and response rate exist between the two input types, but the completion time differed. Radio buttons had a lower completion time than drop-down boxes. But the findings within the two assessed experiments were not consistent. Healy (2007) came to the same result according the number of nonsubstantial answers. In contrast to Heervegh and Loosveldt (2002), the completion time did not differ from each response format and drop-down boxes led to higher item nonresponse and partially to longer response times. This is also consistent with the Keystroke Level Model (KLM). The KLM (Card, Moran, & Newell, 1980) offers the possibility to calculate the expected time, which users need for filling out questions with drop-down boxes and radio

buttons. This model proposes standard times for error-free execution of operations. The only difference between the time needed for searching a word in lists with radio buttons compared to drop-down boxes is, following the KLM, the click to open the drop-down list. After that, both lists need the same completion time. Heervegh and Loosveldt (2002) argue that this time difference does not touch the overall completion time and neutralizes itself, because radio buttons need more time to be downloaded than drop-down boxes. This thesis supports also the findings of Couper, Tourangeau, and Conrad (2004), who found no significant differences in completion time between radio buttons and drop down boxes. The question which one of the two input formats is better has to stay open. But there seems to be a slight empirical reason to prefer radio buttons.

*Dynamic forms*

A coming approach in the WWW is the use of dynamic forms. They are programmed with AJAX, a Web 2.0 technique. The advantage of AJAX is that the single parts of web pages can be loaded directly depending on the users' current actions. As a result, it is unnecessary to reload the whole website, when the user only wants to refresh a part of the content. In combination with forms, Girgensohn, Zimmerman, Lee, Burns, and Atwood (1995) demonstrate that it frees the user from unnecessary distractions and the benefit is most apparent for complex data entry tasks.

Funke and Reips (2007) dealt with response formats in dynamic forms. They compared five input types as follows: plain HTML text boxes, radio buttons, and drop-down boxes as conventional input formats in comparison to two dynamic text boxes, namely autocomplete and autosuggest. Autocomplete offers with every new letter typed in automatically the ending of the current word. Autosuggest gives not only one, but also multiple word suggestions in the form of a drop-down list. They found an effect in average response time: radio buttons and drop-down boxes had lower response times than the other three text boxes. The dropout rate could not be examined and for item nonresponse no differences were found in coding effort,

particularly because of a too small-sized amount of item nonresponders. Additionally the list length was not varied, which we think might be a major factor to test the performance of different response formats.

According to the KLM, it seems reasonable that answering questions with few response possibilities take shorter with radio buttons than with drop-down boxes. In return, the use of radio buttons within long word lists is not meaningful for reasons of space. In previous studies (e.g., see Heervegh & Loosveldt, 2002), the time effect did not touch the overall completion time, because several items were placed on one site, leading to different download times. But this effect should be reduced, if one question per screen is displayed. This method increases additionally usability and records useful information about the dropout rate (Knapp & Heidingsfelder, 2001).

Even though drop-down boxes containing short word lists need longer to be answered than radio buttons, we suggest that they lead faster to the desired target values than autosuggest. Drop-down lists can be shown after one click, whereas autosuggest needs one to three typed letters unless the word list is enough reduced. Additionally, after the typing a change to the mouse must occur to select a possibility from the remaining list. The outcome of this will be an acceleration of the procedure with longer word lists compared to drop-down boxes. Finding a word in a long drop-down list takes much longer than reducing it before the target value is selected. Moreover, we suggest that neither the length of the word lists nor the position of the target value within the word lists should affect response times of autosuggest. Therefore it was hypothesized that:

*Hypothesis 1*: Questions with radio buttons will be answered faster than with drop-down boxes and questions with drop-down boxes will be answered faster than with autosuggest within short word lists.

*Hypothesis 2*: Questions with autosuggest will be answered faster than with drop-down boxes within long word lists.

*Hypothesis 3*: Response times of autosuggest will not differ from the word list length and the position of the target value.

Most studies have not found evidences for differing item nonresponse and dropout rate according to the response format. Also, Funke and Reips (2007) did not find any evidence for differing item nonresponse in dynamic forms, particularly because item nonresponse was quite low. However, two reasons speak for a higher dropout rate with autosuggest. First of all autosuggest needs more effort to be filled out. It is the only response format that needs a keyboard to type the letters. Radio buttons as well as drop-down boxes can be filled out easily only with a mouse. Second, the appearance of this dynamic text box is the same as a normal plain text box, which could be irritating or imply higher effort to answer it. Research about text boxes shows in web surveys (e.g., see Reja et al., 2003; Knapp & Heidingsfelder, 2001) that open-ended questions lead to a higher dropout rate. Therefore it was assumed that participants seeing a plain text box would drop out faster especially during the first pages of the survey:

*Hypothesis 4*: Autosuggest will lead to a higher dropout rate and a lower amount of answered items than radio buttons and drop-down boxes.

An essential aspect of web surveys as well as of forms is the coding effort of the obtained data. In this experiment the risk of unintentional clicking on wrong target values was apparent. Tough, it was not assumed that the amount of wrong coded data would differ. Funke and Reips (2007) did not have any significant differences in invalid answers in their experiment.

*Hypothesis 5*: The amount of wrong chosen target values will not differ regarding the response formats.

It was also important to show, which impact the different response formats would have on subjective feeling. Therefore the cognitive load of each participant was assessed. It was assumed that radio buttons in the short word list and autosuggest in the long word list would reduce the cognitive load, because the effort to find the target value is in each condition very low.

*Hypothesis 6*: The lowest cognitive load appears in the radio button / autosuggest condition.

## Method

### *Study Design*

To examine the advantages and disadvantages of autosuggest an online experiment in German language was conducted in September 2008. A 2 x 3 mixed design was chosen in order to assess effects of the factors *list length* (short vs. long word lists; within-subjects) and *response format* (radio buttons vs. drop-down boxes vs. autosuggest; between-subjects) (see Figure 1 and 2). Since radio buttons were not applied for the long word lists, the factor response format was divided into four conditions. For further explanations see section Apparatus. Dependent variables were response time, dropout rate, number of items answered, coding effort and subjective mental workload.

### *Participants*

The participants were recruited via email from a database with people interested in attending studies from the surroundings of Basel, Switzerland, and via links that were put on several online market places in Switzerland and Germany.

412 persons participated in the experiment. The sample consisted of 138 males ($M$ = 27.9 years, $SD$ = 11.3) and 226 females ($M$ = 27.1 years, $SD$ = 8.7). The mean of the age distribution was 27 years of age with a standard deviation of 10 years. Amongst all participants 20 memory sticks and one MP3 player were raffled off.

Of the 412 participants who started the survey, 39 participants were excluded. 25 of them had chosen exclusively wrong target values and most of them did not finish the survey. The other 14 selected more than half of the chosen values wrong. Looking at the 373 participants that were included in the data analysis, 142 did not finish the survey, representing a dropout rate of 38.1%. This is relatively high, but not further surprising, since the survey was extremely long in order to examine dropout effects, number of answered items and the cognitive load. The focus thereby was rather on the analysis of differences between response formats than representativeness.

*Apparatus*

Autosuggest was implemented using the scriptaculous AJAX framework (Fuchs, 2008). The drop-down list appeared after a time delay of 100ms as soon as at least one letter was typed in.

All participants had to find target values in short as well as in long word lists. Each of them was after the welcome page randomly selected to one of four different conditions. In the first condition all target values had to be found with autosuggest and in the second condition all had to be found with drop-down boxes. The third and fourth condition included radio buttons for the short word lists and autosuggest or drop-down boxes for the long word lists. Radio buttons were not applied for the long word lists, because no one would use such a long word list with radio buttons in a form. In total 96 target values had to be found to terminate the study. Each target value had to be selected alternately in a short or a long word list, whereas the survey began with a short word list.

The short word list consisted each time of 12 different answer options. These 12 values were randomly selected from a self-composed pool containing 100 meaningful words with seven to eight letters. This word list was chosen in such a length to justify the use of autosuggest. Using autosuggest only to request gender for example is not really meaningful. In order to avoid inconsistencies in browser behavior, not more than 12 values were chosen, as this leads to scroll bars in some browsers. The long word list was assessed for more practical reasons. It contained 246 country and territory names taken from ISO 3166-1. 10 countries were excluded, as their notation was too long and consisted of several combined words.

Both lists were alphabetically ordered and the target values were each time randomly selected for each list. To avoid position effects inside word lists, all target value positions were consistently arranged. After having 48 short and 48 long word lists, the order among themselves was randomly assigned. Each participant had the same order with a maximum of randomization, whereas the short and the long word list were presented alternately.


*Figure 1*. Example of searching a target value in the short word list condition using radio buttons, drop-down boxes or autosuggest (Zielwort = target value).

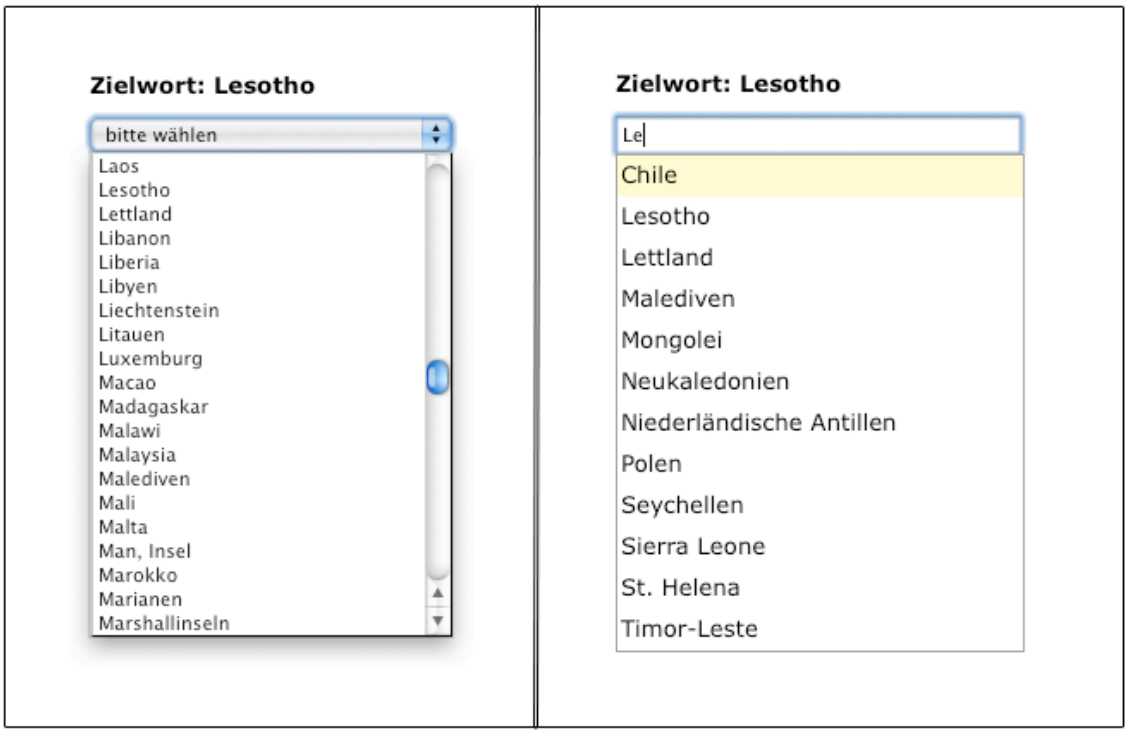


*Figure 2*. Example of searching a target value in the long word list condition using drop-down boxes or autosuggest (Zielwort = target value).

The subjective workload score was obtained by using the NASA Task Load Index (NASA TLX) (Hart & Staveland, 1988), which contains six subscales: mental demand, physical demand, temporal demand, own performance, effort, and frustration level. Normally, participants first rate all subscales and then decide which subscale of each paired combination of the 6 dimensions are more related to their personal definition of mental workload. For this survey only the rating was assessed, as it is easier to use for participants (Hart 2006). This combination of the NASA TLX is often called Raw TLX. For the survey an online version of NASA-TLX (Schmutz, Tuch, & Opwis, 2008) was used.

*Procedure*

The experiment started after a short instruction and some demographic questions with a site showing the first target value. Clicking on “next” conducted the participants to the next page showing again the target value with one of three possible input formats below. The page was active until the target value was found. It was not possible to skip the page without

clicking on a value, guiding the participants to the next target value. As soon as the site appeared, response time was measured. At the end of the survey the cognitive load was assessed.

## Results

### *Response formats and response times*

An alpha level of .05 was used for all statistical tests. All response times higher than 30 seconds were coded as a miss. It was assumed that values higher than 30 seconds would be an indication for a nonserious target value searching. Additionally, after transforming all single response times into logarithmic values a rather normal distribution could be achieved for the statistical analysis. In order to examine the effect of response formats on response times, the average response time of each response format was compared and not the completion time. The disadvantage of the completion time is the dependence on complete responders.

The average response times were first analyzed over both word lists using a one-way ANOVA for unrelated samples, with response format as the independent variable. This analysis was significant, $F(3, 376) = 20.67$, $p < .001$, indicating higher response times for conditions without radio buttons (drop-down, $M = 5066$ ms finding a target value, $SD = 192$; autosuggest, $M = 4705$ ms finding a target value, $SD = 240$) than with radio buttons (radio button / drop-down, $M = 4294$ ms finding a target value, $SD = 192$; radio button / autosuggest, $M = 3750$ ms finding a target value, $SD = 292$). To receive more precise results, the word lists were analyzed separately. Regarding the short word lists, data indicated a lower average response time for radio buttons than for drop-down boxes, $F(1, 368) = 67.43$, $p < .001$, and a lower average response time for drop-down boxes than for autosuggest, $F(1, 368) = 22.21$, $p < .001$. In the long word list condition the results were contrariwise. Using the independent $t$ test for equal variances, data showed that autosuggest led to a lower average response time than dropdown boxes, $t(358) = 7.84$, $p < .001$. All values are shown in Table 1.

Table 1.

*Average Response Times in ms to Find Target Values in each Condition*

| | Short word list | | | Long word list | |
|---|---|---|---|---|---|
| | *M* | *SD* | | *M* | *SD* |
| Drop-Down | 3327 | 187 | | 6864 | 239 |
| Autosuggest | 3950 | 161 | | 4863 | 200 |
| Radio Button | 2133 | 196 | Drop-Down | 6579 | 322 |
| Radio Button | 2030 | 117 | Autosuggest | 5055 | 251 |

*Note*. *M* = Mean; *SD* = Standard Deviation

In order to test, whether the average response time using autosuggest was independent of the list length in comparison to drop-down boxes and radio buttons, a 2 x 2 ANOVA for mixed designs was assessed, using list length as the related samples variable and response format as the unrelated samples variable. There was a significant List Length x Response Format interaction, $F(3,356) = 186.21$, $p < .001$. To compare average response times of both lists contrasts were calculated for each response format. All contrasts were significant, drop-down with $t(117) = 25.51$, $p < .001$, autosuggest with $t(119) = 9.34$, $p < .001$, radio button / drop-down with $t(63) = 24.01$, $p < .001$ and radio button / autosuggest with $t(57) = 29.34$, $p < .001$, indicating that longer word lists lead for all response formats to longer average response times. Tough, looking at Figure 3 it foreshadows not the same list length effect for all response formats. Autosuggest is less dependent on the influence of the list length comparing to the other three conditions.

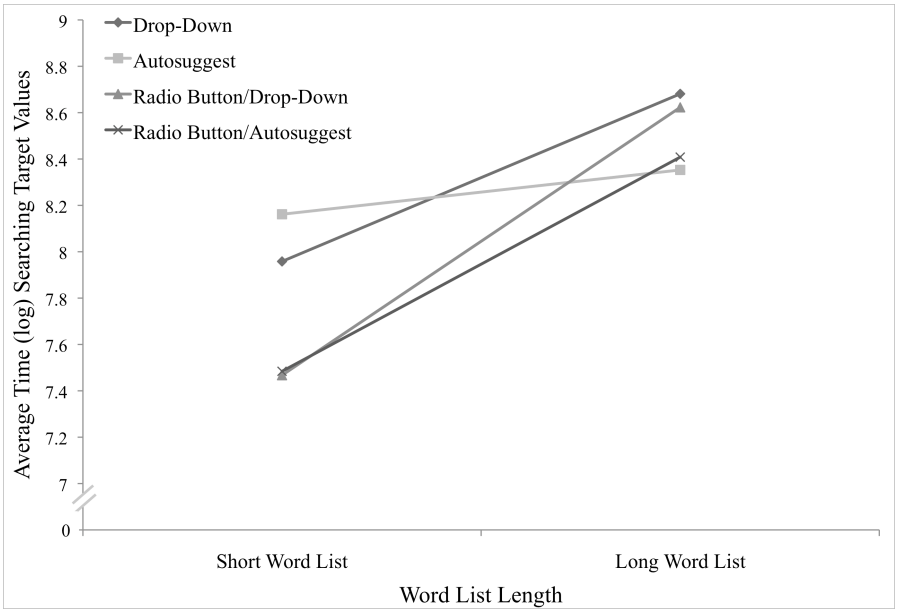


*Figure 3*. Relation between the word list length and the average logarithmic time searching target values.

It was hypothesized that the average response time using autosuggest would be independent of the target value position. Therefore all lists were divided in two groups, whereas one group contained the target values in the upper part of the word lists and the other in the lower part. The data were analyzed using again 2 x 2 ANOVA for mixed design, this time with the target value position as the related samples variable. There was a statistically significant main effect for the target value position, $F(1,346) = 108.08$, $p < .001$, as well as a significant interaction, $F(3,346) = 18.99$, $p < .001$. Comparing each response format with the two target value positions (see Table 2), following contrasts were significant indicating higher average response times for target values in the lower part of the word lists: drop-down with $t(113) = 6,67$, $p < .001$, radio button / drop-down with $t(63) = 9.47$, $p < .001$ and radio button / autosuggest with $t(57) = 4.31$, $p < .001$. Autosuggest did not differ according to the target value position, $t(113) = .22$, $p = .83$.

Table 2.

*Average Response Times in ms for Target Value Positions in each Condition*

| | Upper part | | Lower part | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| Drop-Down | 4763 | 207 | 5334 | 199 |
| Autosuggest | 4270 | 158 | 4234 | 160 |
| Radio Button / Drop-Down | 3992 | 180 | 4924 | 351 |
| Radio Button / Autosuggest | 3331 | 156 | 3519 | 187 |

*Note*. *M* = Mean; *SD* = Standard Deviation

*Response formats and dropout rate*

Turning to the amount of complete responders and participants who dropped out a contingency table with a $\chi^2$-test was calculated on the different response formats. Data show that there were no significant differences between the response formats, $\chi^2(3, N = 373) = 2.73$, $p = .44$. Analyzing just the data of participants, who dropped out, using one-way ANOVA for unrelated samples, with response format as independent variable, it revealed that in the autosuggest condition less items were answered than in the other conditions (see Figure 4). The main effect was not statistically significant, $F(3,138) = 1.69$, $p = .17$, but the relevant contrast showed that in the autosuggest condition less items were answered compared to the other conditions, $F(1,138) = 4.18$, $p = .02$. Especially within the first two presented target values more participants dropped out in autosuggest and radio button / autosuggest conditions (12 participants) than in drop-down and radio button / drop-down conditions (2 participants).

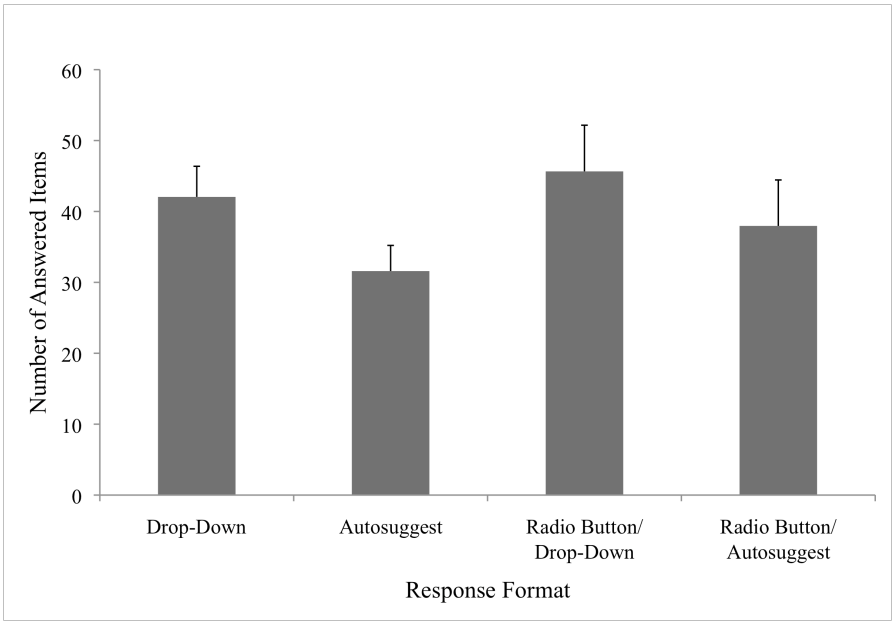


*Figure 4*. Mean and standard errors of items answered per condition for all participants, who dropped out.

*Response formats and coding effort*

Comparing the amount of all wrong selected target values for each response format condition data revealed no statistically significant differences in coding effort using again one-way ANOVA, $F(3,369) = .05$, $p = .99$ (drop-down, $M = 1.82$ wrong selected values, $SD = 2.42$; autosuggest, $M = 1.89$ wrong selected values, $SD = 2.33$; radio button / drop-down, $M = 1.95$ wrong selected values, $SD = 1.83$; radio button / autosuggest, $M = 1.91$ wrong selected values, $SD = 2.64$).

*Response formats and NASA TLX*

It was assumed that the radio button / autosuggest condition would have the lowest value on participants' cognitive load. Though, using one-way ANOVA there were no significant differences between response formats regarding mean score of all NASA TLX subscales, $F(3,294) = .79$, $p = .50$ (see Table 3).

Table 3.

*NASA TLX Subscores and Mean Score for each Response Format Condition*

| | Drop-Down | | Autosuggest | | Radio Button/ Drop Down | | Radio Button/ Autosuggest | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Mental Demand | 19.16 | 21.94 | 14.52 | 18.88 | 21.30 | 25.38 | 17.82 | 20.35 |
| Physical Demand | 8.97 | 15.71 | 12.80 | 19.88 | 9.16 | 18.55 | 8.26 | 12.69 |
| Temporal Demand | 43.05 | 28.69 | 40.24 | 29.23 | 53.14 | 29.86 | 39.68 | 27.60 |
| Own Performance | 53.61 | 34.35 | 47.81 | 31.77 | 48.09 | 35.00 | 50.52 | 33.30 |
| Effort | 24.64 | 24.37 | 25.51 | 26.87 | 26.88 | 25.45 | 23.94 | 21.07 |
| Frustration Level | 42.76 | 35.21 | 42.61 | 35.37 | 43.95 | 36.16 | 38.44 | 32.48 |
| Mean Score | 32.03 | 15.96 | 30.58 | 15.05 | 33.75 | 15.63 | 29.78 | 12.83 |

*Note*. *M* = Mean; *SD* = Standard Deviation

Nevertheless, temporal demand differed between the 4 conditions, $F(3,294) = 2.82$, $p = .04$. The values of the radio button / drop-down condition were higher compared to the other conditions. The felt time pressure of participants was especially diverse in the radio button conditions, whereas the radio button / autosuggest condition had lower values than the radio button / drop-down condition, $F(1,294) = 5,77$, $p = .02$.

## Discussion

In this experiment, exploring the difference between one dynamic and two classical response formats, it was shown that the optimal response format depends on the word list length.

Regarding the different response times, the advantage of each response format is clear. Radio buttons are a good option for short word lists, whereas drop-down boxes lead to longer response times. These findings have not been seen so clear in research yet. On the other hand, autosuggest is just an option for longer word lists. Although, the use of autosuggest is faster in short word lists than in longer ones, it is just effective if the amount of words in a list is so long that by typing in some letters in the response box a meaningful limitation of the drop-down list is reached. Compared to drop-down boxes autosuggest is useful as soon as the

procedure of limitation happens faster than the scrolling in the whole drop-down list to find the target value.

Another advantage of autosuggest is the independency of the target value position in short as well as in long word lists. While the order of words within a radio button or drop-down list must be planned, the word list order standing behind autosuggest is nonrelevant. This disadvantage of radio buttons and drop-down boxes can only be corrected in drop-down boxes with a pre-selection of the most important target values at the top of a drop-down list.

The hypothesis about the dropout rate stated that autosuggest would have more participants who drop out than the other response formats, because of its appearance. Tough, this study revealed no differences between the input types. Also the amount of unintentional clicking on wrong target values did not differ. Both result were already found in previous studies. But, looking more precisely at the moment when participants dropped out, it becomes clear that, as hypothesized, autosuggest leads people to drop out earlier. More people drop out during the first two presented questions in relation to the two classical response formats. It seems that autosuggest is not enforced in forms yet. The perceived appearance of autosuggest irritates the user and lets him believe that it could be an open-ended question with great effort to answer it.

The objective differences of effort regarding the response formats does not really concern participants subjective cognitive load. Their only subjective feeling that differs is temporal. Drop-down boxes with long drop-down lists put participants under time pressure. These findings underline once more the clear advantage of autosuggest for longer word lists.

Further studies should tie on these results with a more practical approach. Most forms are completed on one site, whereas here one response format after another was shown. Therefore the different download times of a normal form could not be examined. Additionally the impact of item nonresponse should be further analyzed, because in this experiment participants were forced to select the target values.

In this study important insights were achieved, showing that radio buttons are useful for short word lists, whereas autosuggest excels drop-down boxes in long word lists. The only disadvantage of autosuggest seems to be that users are not familiar with this response format in web forms yet. But, it is to assume that the more autosuggest will be used in the WWW in the next years, the more people will expect it also in web forms or web surveys. Perhaps it is even possible to reduce the dropout rate with autosuggest one day compared to radio buttons and drop down lists. It could also been considered that an introduction describing autosuggest before using it, could reduce the dropout rate in web forms or web surveys.

A lot of research was conducted in the last years finding differences on rating scales and response formats in web surveys. Regarding web forms just little research exists. Online research should tie on these results and especially towards dynamic forms further exploration is advised.

## References

Beaumont, A., James, J., Stephens, J., & Ullman, C. (2002). *Usable forms for the Web*. Birmingham: Glasshaus.

Brunier, G., & Graydon, J. (1996). A comparison of two methods of measuring fatigue in patients on chronic haemodialysis: visual analogue versus likert scale. *International Journal of Nursing Studies, 33*, 338-348.

Card, S.K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance with interactive systems, *Communications of the ACM*, *23*(7), 396-410.

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

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

Couper, M. P., Traugott, M. W., & Lamias, M. J. (2001). Web survey design and administration. *Public Opinion Quarterly*. *65*, 230-253.

Couper, M. P., Tourangeau, R., & Conrad F. G. (2004). What they see is what we get. *Social Science Computer Review*, *22*(1), 111-127.

Couper, M. P., Tourangeau, R., & Conrad F. G. (2006). Evaluating the effectiveness of visual analog scales. *Social Science Computer Review*, *24*(2), 227-245.

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

Fuchs, T. (2008). Script.aculo.us AJAX framework. Retrieved March 25, 2008 from http://www.scriptaculous.com/

Funke, F., & Reips, U.-D. (2007, March). *Dynamic Forms: online surveys 2.0*. Paper presented at the 9th Annual Conference of the German Society for Online Research (D.G.O.F.), Leipzig, Germany.

Garrett, J. J. (2005). *Ajax: A New Approach to Web Applications*. Retrieved November 5, 2008 http://www.adaptivepath.com/ideas/essays/archives/000385.php

Girgensohn, A., Zimmerman, B., Lee, A., Burns, B., & Atwood, M. (1995, June). *Dynamic forms: an enhanced interaction abstraction based on forms*. Paper presented at the 5th International Conference on Human-Computer Interaction (Interact '95), Lillehammer, Norway.

Grant, S., Aitchison, T., Henderson, E., Christie, J., Zare, S., Mcmurray J., & Dargie H. (1999). A comparison of the reproducibility and the sensitivity to change of visual analogue scales, borg scales, and likert scales in normal subjects during submaximal exercise. *CHEST, 116*, 1208-1217.

Hart, S.G., & Staveland, L. (1988). Development of NASA task load index TLX: results of empirical and theoretical research. In P. A. Hancock, and N. Meshkati (Eds.). *Human mental workload* (pp. 139-183). Amsterdam: Elsevier Science.

Hart, S. (2007). Nasa-Task Load Index (Nasa-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *USA*, *5*, 904-908.

Heerwegh, D., & Lossveldt, G. (2002). An evaluation of the effect of response formats on data quality in web surveys. *Social Science Computer Review*, *20*(4), 471-484.

Healy, B. (2007). Drop downs and scroll mice: the effect of response option format and input mechanism employed on data quality in web surveys. *Social Science Computer Review*, *25*(1), 111-128.

Knapp, F., & Heidingsfelder, M. (2001). Drop-out analysis: Effects of the survey design. In U.-D. Reips, and M. Bosnjak (Eds.). *Dimensions of Internet Science* (pp. 221-230). Lengerich: Pabst Science.

Magee, C.G., Straight, R.L., & Schwartz, L. (2001, May). *Conducting web-based surveys: lessons learned and keys to success with known populations*. Paper presented at the 56th Annual Conference of the American Association for Public Opinion Research (AAPOR), Montreal, Canada.

Miller, S., & Jarrett, C. (2001). *Should I use a drop-down? Four steps for choosing form Elements on the Web*. Retrieved July 23, 2007 from http://www.formsthatwork.com/ftp/dropdown.pdf

Redline, C., & Dillman, D. A. (2002). The influence of alternative visual designs on respondents' performances with branching instructions in self-administered questionnaires. In R. Groves, D. Dillman, J. Eltinge, and R. Little (Eds.). *Survey nonresponse.* (pp. 179-195). New York: John Wiley and Sons.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web questionnaires. *Advances in Methodology and Statistics (Metodološki zveski)*, *19*, 159-177.

Schmutz, P., Tuch, A. N., and Opwis, K. (2008). NASA-TLX online. [*webbased application*]. Unpublished internal application.

Smith, T. W. (1995). Little things matter; a sampler how differences in questionnaire format can affect survey responses. In *Proceedings of the American Statistical Association, Survey Research Methods Section* (pp. 1046-1051). Alexandria, VA: American Statistical Association.

Tourangeau, R., Couper, M. P., & Conrad F. (2004). Spacing, Position, and Order. Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly*, *68*, 368–393.