

Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator

ULF-DIETRICH REIPS

University of Zürich, Zürich, Switzerland

AND

FREDERIK FUNKE

University of Tübingen, Tübingen, Germany

The present article describes VAS Generator (www.vasgenerator.net), a free Web service for creating a wide range of visual analogue scales that can be used as measurement devices in Web surveys and Web experimentation, as well as for local computerized assessment. A step-by-step example for creating and implementing a visual analogue scale with visual feedback is given. VAS Generator and the scales it generates work independently of platforms and use the underlying languages HTML and JavaScript. Results from a validation study with 355 participants are reported and show that the scales generated with VAS Generator approximate an interval-scale level. In light of previous research on visual analogue versus categorical (e.g., radio button) scales in Internet-based research, we conclude that categorical scales only reach ordinal-scale level, and thus visual analogue scales are to be preferred whenever possible.

Internet-based research has become a new procedural option in psychology and related sciences, and many researchers are looking for tools, methodologies, and services that support this type of research. If such tools for Internet-based research are offered for use via the Web (not via download and install), they are called *Web services*. Such Web services are available wherever there is an Internet connection and a Web browser.

Web services for Internet-based research include *collections of Web studies* for recruitment and archiving (e.g., the “Psychological Research on the Net” list by John Krantz, and the “Web Experiment List” at genpsylab-wexlist.unizh.ch/; Reips & Lengler, 2005), tools for *Web surveying* (e.g., surveymonkey.com or surveyWiz; see Birnbaum, 2000), *Web experiment generators* (e.g., WEXTOR at wextor.org; see Reips & Neuhaus, 2002), *Internet-based tests* for inclusion in studies (e.g., b5 at iscience.eu), *Web log analysis tools* (e.g., Scientific LogAnalyzer; Reips & Stieger, 2004), and *portals* that link to all related services (e.g., the iScience server at iscience.eu).

The present article describes a first example of a new set of tools that can be termed *Web measurement device generators*. VAS Generator (www.vasgenerator.net) is a free Web service for creating a wide range of visual analogue scales that can be used as measurement devices in Web surveying and Web experimentation, and also for local computerized assessment. VAS Generator and the scales it generates are platform-independent, meaning that the scales can be created from and used on all types and versions of operating

systems with Web browsers that support the underlying universal languages of HTML and JavaScript. The resulting scales can easily be added to surveys and experiments generated with other Web services, such as surveyWiz (Birnbaum, 2000) and WEXTOR (Reips & Neuhaus, 2002).

VISUAL ANALOGUE SCALES

Visual analogue (or analog) scales (VASs) are continuous measurement devices (see, e.g., Flynn, van Schaik, & van Wersch, 2004). In 1921, VASs were described for the first time (Hayes & Patterson, 1921). However, this type of scale was not seriously examined before 1969 (Aitken, 1969). In some respects, Aitken’s basic findings are still state of the art, since—unlike on other measuring instruments—little research has been conducted on this type of scale (for an overview, see Flynn et al., 2004). Practical concerns are the major reason for this lack of research. In paper-based VASs, a lot of time and effort are required for reading the data: The exact position of each marking has to be determined by hand. Of course, the situation changed considerably with computerization. With the rise of Internet-based research, VASs have become a measurement device that could be used widely without practical drawbacks (Couper, Tourangeau, Conrad, & Singer, 2006; Funke & Reips, 2006). The wider use of VASs may even solve some issues related to data quality with other online measurement devices (Funke & Reips, 2006; Reips, 2002a; Smyth, Dillman, Christian, & Stern, 2006).

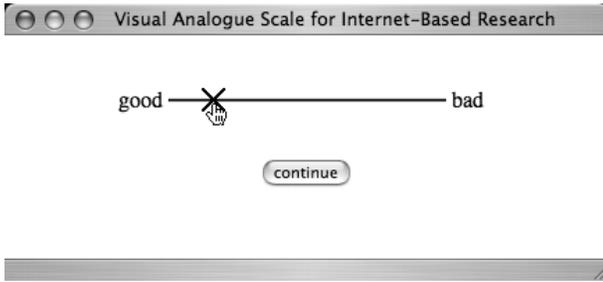


Figure 1. A visual analogue scale generated with VAS Generator.

A VAS consists of a line and two anchors, one at each end. The anchors often consist of verbal materials that mark opposite ends of a semantic dimension (e.g., *good* and *bad*). However, the anchors may also be pictures, or even sound files. Visual anchors, such as smileys, are also often used with participants who may not fully grasp the meaning of verbal materials—for example, with preschool children. See Figure 1 for an example VAS generated with VAS Generator.

In general, VASs are considered a reliable instrument for valid measurements. However, recent research has mostly been limited to comparisons between VASs and categorical scales (Likert-type questions) in a paper-and-pencil environment. The majority of studies have been conducted in the medical sector, in which reliable detection of small changes in status is appreciated a lot, especially in the fields of pain and fatigue research (e.g., Bellamy, Campbell, & Syrotuik, 1999; Brunier & Graydon, 1996). A notable exception, conducted by Gerich (2007), com-

pared VASs and 5-point categorical self-control scales not only in a paper-and-pencil format, but also with computer-assisted self-administered interviews.

VASs provide researchers with a number of advantages. In comparison with discrete scales, measurement by a VAS is more exact, and the scale needs less explanation for the research participants (e.g., smiley-face scales in studies with children). In a previous study, we were able to show that online radio button scales do not show linear correspondence to VASs (Funke & Reips, 2006). In that study, two experiments were conducted, comparing categorical scales of 4, 5, 7, 8, and 9 points with a VAS of medium length (200 pixels). The study presented here further investigated whether the VAS format diverges from the interval level and whether the length of a VAS may be a boundary condition for such an effect. First, we describe VAS Generator.

CREATING AND IMPLEMENTING A SCALE WITH VAS GENERATOR

Just five steps are required to generate an individual VAS for use on a Web page. In a simple HTML form available at www.vasgenerator.net (see Figure 2), the researcher modifies the essential graphical parameters (length, color, width, and type of marking) and the verbal anchors. Below we explain the five necessary steps.

Step 1: Defining Core Scale Parameters

Several mandatory parameters are set at default levels and can be adjusted according to one's needs. These parameters are the length, anchors, color, and marker. The length of the VAS by default is 200 pixels, a length that will

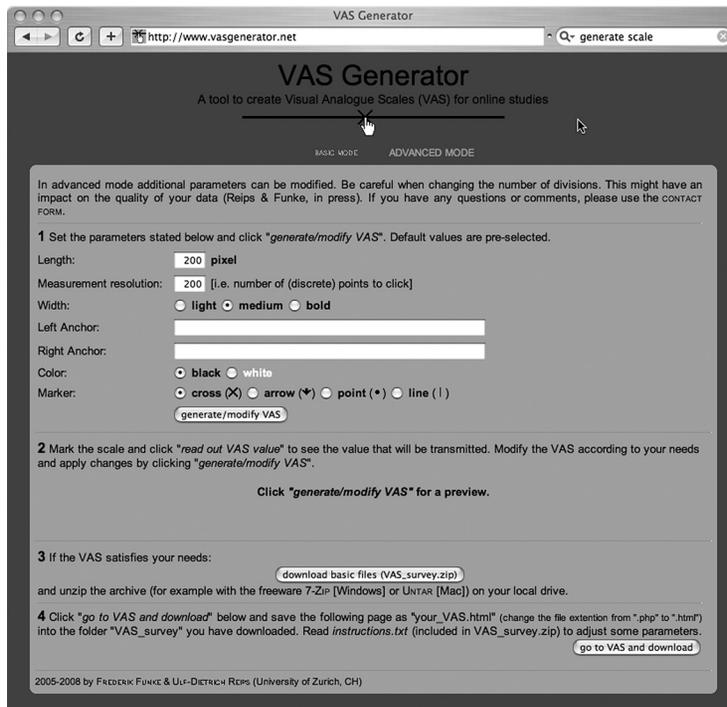


Figure 2. VAS Generator's main window.



Figure 3. Readout of rating value in feedback window.

be displayed fully without problems in most Web browsers and with most currently used monitor resolutions. The verbal anchors located at each end of the VAS have a preset maximum length of 50 characters, but this value can be adjusted when pasting the code to a Web page. The scale's color can be changed from black (which is the default value) to white, and the width from medium (3 pixels) to light (1 pixel) or boldface (5 pixels). The marker that will appear after clicking on the scale, representing the user's value, can be modified from a cross to an arrow, point, or line. Measurement resolution—that is, the number of (discrete) points to click—can also be adjusted.

After confirming these parameters by clicking on the “generate/modify VAS” button, a working preview of the scale is displayed in paragraph 2 on the same Web page.

Step 2: Testing the Scale Draft

The VAS that has been generated can be clicked on in order to display the value (i.e., the position of the marking, measured in pixels from the left-hand side) in a pop-up window.

Figure 3 shows a bold, black VAS that is 400 pixels in length. The left end is anchored with *strongly agree* and the right with *strongly disagree*. A cross acts as the indicator of the respondent's value (here, at 115 pixels).

Step 3: Downloading the Basic Scale Materials

If the VAS satisfies one's needs, all of the basic files required to include the VAS on a Web page (i.e., the JavaScript code and picture files, with an additional instruction for off-line use) can be downloaded as a compressed file in .zip format. Links to free software that can be used for decompressing (“unzipping”) the file are provided. All files in the decompressed folder need to be stored in the same folder as the Web page that will contain the newly generated VAS.

Step 4: Downloading the Customized Scale

The VAS that was customized in the first two steps can be displayed in a separate browser window by clicking on the “go to VAS and download” button. The source code for this window is downloaded by simply saving the page from the browser menu into the folder with the basic files (Step 3) and changing the file extension from .php to .html. The basic file folder should now contain an HTML document, two JavaScript files (ending in “.js”), and an “images” subfolder containing the picture files needed for displaying the scale. The JavaScript code for the rating scales was originally developed by Walter Zorn and Timo Gnamb and is distributed as open source (timo.gnamb.at/downloads/zip/ratingscales.zip).

Hence, there are two options: The VAS can either be pasted easily into an existing Web page or serve as the basis for a new survey. Pasting the VAS into an existing Web page will likely be the more frequent way of proceeding. Therefore, the related editing procedure in HTML is described in Step 5 below.

Step 5: Implementing the Scale on a Web Page (Includes the Resulting Code)

To integrate the VAS with an existing project, the source code of the downloaded page needs to be modified. Only three parameters (printed in capitals in Figure 4) have to be adjusted: The name of the current page

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 TRANSITIONAL//EN">
<html>
  <head>
    <title>YOUR_PAGE_TITLE</title>
  </head>
  <body>
<!-- "wz_dragdrop.js" and "ratingscale.js" have to be included immediately after "<body onload=next_page()" -->
    <script type="text/javascript" src="js/wz_dragdrop.js" title="Dragdrop Bibliothek (c) Walter Zorn"></script>
    <script type="text/javascript" src="js/ratingscales.js" title="Ratingscales Addon (c) Timo Gnamb"></script>
    <p>
      <table align="center" border="0">
        <tr>
          <td align="center">
            
            <form name="vas_data" action="YOUR_NEXT_PAGE.html" method="get">
              <script type="text/javascript">
                scales.scale(THIS_VAS_NAME, 'click', 400, ['pics/line20.gif', , 10], ['pics/00.gif'], ['strongly agree', 'strongly disagree']);
              </script>
            </td>
          </tr>
          <tr>
            <td align="center">
              <input type="submit" value="continue" />
            </td>
          </tr>
        </table>
      </p>
<!-- "<script type="text/javascript">scales.init('value');</script--> has to be immediately before "</body--> -->
    <script type="text/javascript">scales.init('value');</script>
  </body>
</html>
```

Figure 4. Placement of customized code for one's particular VAS in an HTML file.

("YOUR_PAGE_TITLE") that is displayed at the top of the browser window, the name of the following page ("YOUR_NEXT_PAGE.html") and—this is the most important step if one uses several VASs in one survey—the name of the current scale ("THIS_VAS_NAME").

ANALYZING DATA FROM A VAS

Data from the VAS contain the following information: a "name" that was given to this particular item (in order to separate entries from those related to other items) and a "value" that reflects a participant's rating on the scale. The value is read out automatically, with an accuracy of one pixel. This means that the greater the length of a VAS, the more diversified the range of possible values. In other words, the precision of the rating is only as high as the length of the VAS permits.

The type of data analysis depends on the method used to collect data on the Web page the VAS was embedded in. For example, data would be ready for direct download if WEXTOR's hosting feature were used. Data could also be written to a database or retrieved from the server log file, which can be analyzed with Scientific LogAnalyzer (Reips & Stieger, 2004).

EMPIRICAL TEST OF INTERVAL LEVEL MEASUREMENT

To examine whether VASs created with VAS Generator produce data on the level of an interval scale in Web studies, we conducted a Web experiment. Furthermore, we were interested in putting the scales' robustness to the test by introducing extremely short or long VASs and by varying the mathematical format (percentages vs. ratios) of the target values.

Method

Procedure. In a Web experiment, participants were randomized to one of six conditions in a 2 (mathematical format) \times 3 (length of VAS) experimental design. The Web experiment consisted of 30 pages per condition. The first page contained an introduction to the Web experiment and a seriousness check (Reips, 2002b). On the second page, participants were asked their gender, online experience (since when and frequency), speed of connection, and screen size. They were instructed to repeatedly identify 13 different values (percentages—e.g., 20%, 50%—or ratios—e.g., 1/5, 1/2—depending on the mathematical format condition) in one of three length conditions: a VAS length of 50, 200, or 800 pixels. The 13 values ranged from 5% to 95% and were displayed in two reverse orders, AB and BA (see the A and B rows in Table 1), always with 50% as the first item. We tried to keep a large difference between consecutive values. The underlying measurement on the VAS ranged from 0 to 100.

Before and after the second block, we asked participants about three usability considerations: accuracy, precision, and interest.

Participants. The participants were recruited from a student-based online panel at the University of Kassel that is maintained by the second author. Within 5 weeks, 439 persons took part in our study. Of the 405 participants (92.3%) who indicated serious participation, 205 were randomly sent to the percentage condition and 200 to the ratio condition. In the percentage condition, 17 participants (8.3%) dropped out during the study, and 31 (15.5%) dropped out from the ratio condition. Furthermore, 2 participants in the ratio condition did not provide even a single rating for 1 of the 13 values. Their data were excluded from analysis, so that 355 data sets (188 in the percentage condition, 167 in the ratio condition) remained for further analyses.

Results

For each individual, we computed the average ratings for the same values from the A and B orders.

Outliers. A standard outlier analysis revealed that overall ratings averages by 2 participants in the percentage condition and 8 in the ratio condition differed from the pack by more than three interquartile ranges (extreme outliers). Seven of these 10 cases were from the 50-pixel, 2 from the 200-pixel, and 1 from the 800-pixel length condition. All extreme outliers were excluded from further analyses.

For each individual, we computed the average ratings for the same values from the two orders A and B. Overall, the target values matched actual values well (see Figure 5). With two exceptions at the low end and two in the center, the VASs slightly underestimated the true values. This tendency was a bit stronger for the shortest VAS. Overall, the mean deviation from the target value ranged from 3.94 points for the short VAS to 2.98 for the long VAS, with 2.77 for the medium VAS (Table 2).

Interval-level measurement implies even differences between values with even distances. Therefore, we examined whether participants produced the same differences between their ratings on each portion of the scale. The mean differences for 5-percentage-point intervals (with *SDs* in parentheses) were 4.95 (8.30) for the 5%–10% interval, 4.40 (5.54) for the 20%–25% interval, 5.26 (5.07) for the 75%–80% interval, and 5.15 (7.48) for the 90%–95% interval. The mean differences for 10-percentage-point intervals were 8.47 (7.42) for the 10%–20% interval, 12.52 (6.47) for the 40%–50% interval, 10.14 (5.24) for the 50%–60% interval, and 8.92 (6.36) for the 80%–90% interval.

Analysis of potential biases when using extreme scales. We computed the differences of the average ratings from the true values. With the overall averages of these deviations from the true VAS values, we conducted

Table 1
Sample Length Values Displayed to Participants

		Percentage Condition												
A	50	75	10	33	80	95	25	67	40	5	60	90	20	
B	50	20	90	60	5	40	67	25	95	80	33	10	75	
		Ratio Condition												
A	1/2	3/4	1/10	1/3	4/5	19/20	1/4	2/3	2/5	1/20	3/5	9/10	1/5	
B	1/2	1/5	9/10	3/5	1/20	2/5	2/3	1/4	19/20	4/5	1/3	1/10	3/4	

Table 2
Mean Differences Between Target Values and Actual Values by VAS Length and Condition

VAS Length	<i>M</i>	<i>N</i>	<i>SD</i>
Combined			
50 pixels	3.94	125	2.24
200 pixels	2.77	105	1.26
800 pixels	2.98	115	1.91
Total	3.26	345	1.94
Percentage			
50 pixels	4.03	64	2.59
200 pixels	2.75	59	1.09
800 pixels	2.73	63	1.18
Total	3.18	186	1.87
Ratio			
50 pixels	3.85	61	1.82
200 pixels	2.79	46	1.47
800 pixels	3.28	52	2.50
Total	3.36	159	2.02

an ANOVA with the factors mathematical format and length of VAS. The difference between the two format conditions was small (mean deviations: 3.18, $SD = 1.87$, for percentages, and 3.36, $SD = 2.02$, for ratios) and not significant [$F(1,335) < 1$, n.s.]. Length of VAS turned out to have a highly significant effect [$F(2,335) = 17.76$, $p < .001$, $\eta^2 = .09$], but the interaction was not statistically significant [$F(2,335) < 1$, n.s.]. Table 2 shows average differences between the actual value and the average VAS rating by condition.

Planned comparisons were made between the different levels of the length factor. Both main effect contrasts between the 50-pixel length condition and the other conditions were statistically significant [$F(1,338) = 30.13$, $p < .001$ (200 pixels), and $F(1,338) = 22.21$, $p < .001$ (800 pixels)]. No significant difference was found for the 200-pixel versus 800-pixel comparison [$F(1,338) = 0.074$, n.s.].

Technically, if a participant's value is off by 1 pixel on the 50-pixel VAS, an error of two units will be recorded, whereas a 1-pixel error is only 0.125 units on the 800-pixel VAS. Therefore, the larger error in the 50-pixel VAS comes as no surprise and can be explained mostly by this lack of technical precision. Furthermore, the short VAS may have fostered a tendency to repeatedly click in the middle of the scale—a behavior that would explain the shape of the 50-pixel curve in Figure 5.

As an aside, the large difference in dropouts between the mathematical format conditions (odds ratio = 1.87) and a four-times-higher rate of extreme outliers indicates that the ratio condition is more difficult than the percentage condition. It can be assumed that the voluntary Web mode in most Web experiments is a comparatively good indicator of participants' motivation to continue with the study (see, e.g., Reips, 2002b, 2002c, 2007). Accordingly, it seems wise to avoid the ratio format with VASs, because users seem to have less motivation when using this format, and consequently either may not answer (i.e., drop out when possible) or may produce lower-quality data if they find the situation less easy to leave and feel forced to

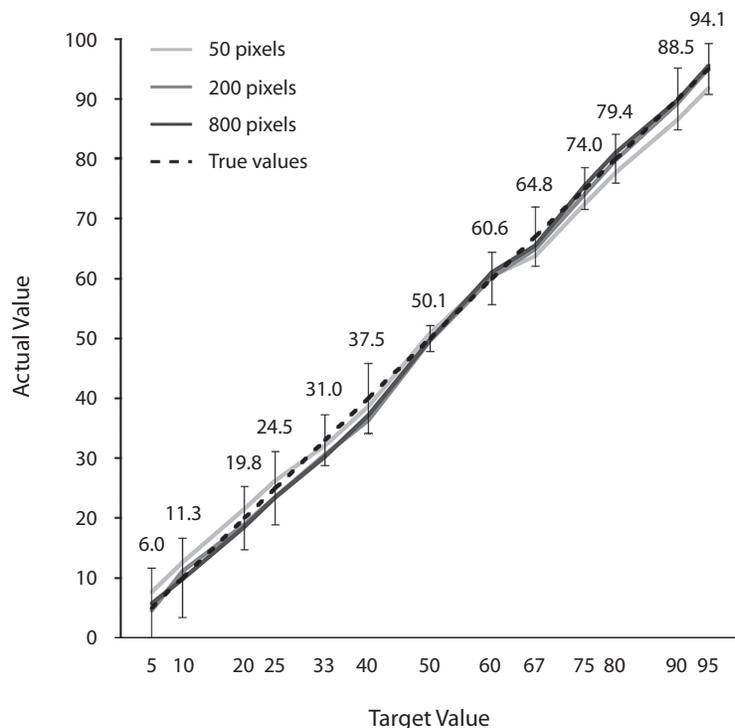


Figure 5. Estimated versus true values for VAS Generator VASs of various lengths. All ratings have been converted to percentages for the purposes of this figure.

answer, as may often be the case in laboratory studies with an experimenter present.

CONCLUSION

Because equal numerical intervals corresponded to roughly equal segments on the VASs and only minor aberrations from true values were found across the scale, even for extremely short and extremely long VASs, there is strong evidence that data collected with VASs are equidistant and are on the level of an interval scale. Therefore, a wide range of statistical procedures can be applied safely when analyzing data measured with VASs that were created with VAS Generator. In contrast, because there is a systematic difference between equally spaced radio buttons and VASs (Funke & Reips, 2006), we conclude that measurements with radio button scales differ from interval level.

Although the effects were small, we found differences for the different VAS lengths on ratings of some values. However, we used an extremely long (800-pixel) and an extremely short (50-pixel) VAS that may have looked highly unusual to most participants. If on a given screen or printed page, a 50-pixel VAS is about as wide as the word "participant" on this page, then an 800-pixel VAS is about as wide as the text on this full page would be high. Given these sizes, it is surprising how well the participants performed using the 800-pixel scale. This finding implies that measurement with VASs is robust to differences in size due to different screen sizes and resolutions. Nevertheless, care should be taken in determining a VAS that is well-suited for all displays used in a study, which most often will be a medium-sized one (150–400 pixels).

Finally, as compared with hand-coding HTML and JavaScript, VAS Generator greatly reduces the effort required in generating VASs for Web-based studies. It provides all of the files needed for the use of VASs and offers preview and pretest functions. Last but not least, VASs built by means of the VAS Generator can easily be included in existing projects (e.g., in Web experiments created with WEXTOR).

AUTHOR NOTE

This article is based on a presentation given at the 36th Annual Meeting of the Society for Computers in Psychology, Houston, TX, on November 16, 2006. All correspondence concerning this article should be addressed to U.-D. Reips, Sozial- und Wirtschaftspsychologie, Universität Zürich, Binzmühlestrasse 14, Box 13, 8050 Zürich, Switzerland (e-mail: u.reips@psychologie.uzh.ch).

REFERENCES

- AITKEN, R. C. B. (1969). Measurement of feelings using visual analogue scales. *Proceedings of the Royal Society of Medicine*, **62**, 989-993.
- BELLAMY, N., CAMPBELL, J., & SYROTUIK, J. (1999). Comparative study of self-rating pain scales in osteoarthritis patients. *Current Medical Research & Opinion*, **15**, 113-119.
- BIRNBAUM, M. H. (2000). SurveyWiz and FactorWiz: JavaScript Web pages that make HTML forms for research on the Internet. *Behavior Research Methods, Instruments, & Computers*, **32**, 339-346.
- BRUNIER, G., & GRAYDON, J. (1996). A comparison of two methods of measuring fatigue in patients on chronic haemodialysis: Visual analogue vs. Likert scale. *International Journal of Nursing Studies*, **33**, 338-348.
- COUPER, M. P., TOURANGEAU, R., CONRAD, F. G., & SINGER, E. (2006). Evaluating the effectiveness of visual analog scales: A Web experiment. *Social Science Computer Review*, **24**, 227-245.
- FLYNN, D., VAN SCHAIK, P., & VAN WERSCH, A. (2004). A comparison of multi-item Likert and visual analogue scales for the assessment of transactionally defined coping function. *European Journal of Psychological Assessment*, **20**, 49-58.
- FUNKE, F., & REIPS, U.-D. (2006, March). *Visual analogue scales in online surveys: Nonlinear data categorization by transformation with reduced extremes*. Poster presented at the General Online Research (G.O.R.) conference, Bielefeld, Germany.
- GERICH, J. (2007). Visual analogue scales for mode-independent measurement in self-administered questionnaires. *Behavior Research Methods*, **39**, 985-992.
- HAYES, M. H. S., & PATTERSON, D. G. (1921). Experimental development of the graphic rating method. *Psychological Bulletin*, **18**, 98-99.
- REIPS, U.-D. (2002a). Context effects in Web surveys. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online social sciences* (pp. 69-80). Seattle: Hogrefe & Huber.
- REIPS, U.-D. (2002b). Internet-based psychological experimenting: Five dos and five don'ts. *Social Science Computer Review*, **20**, 241-249.
- REIPS, U.-D. (2002c). Standards for Internet-based experimenting. *Experimental Psychology*, **49**, 243-256.
- REIPS, U.-D. (2007). The methodology of Internet-based experiments. In A. Joinson, K. McKenna, T. Postmes, & U.-D. Reips (Eds.), *Oxford handbook of Internet psychology* (pp. 373-390). Oxford: Oxford University Press.
- REIPS, U.-D., & LENGELER, R. (2005). The Web Experiment List: A Web service for the recruitment of participants and archiving of Internet-based experiments. *Behavior Research Methods*, **37**, 287-292.
- REIPS, U.-D., & NEUHAUS, C. (2002). WEXTOR: A Web-based tool for generating and visualizing experimental designs and procedures. *Behavior Research Methods, Instruments, & Computers*, **34**, 234-240.
- REIPS, U.-D., & STIEGER, S. (2004). Scientific LogAnalyzer: A Web-based tool for analyses of server log files in psychological research. *Behavior Research Methods, Instruments, & Computers*, **36**, 304-311.
- 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**, 6-16.

(Manuscript received March 18, 2007;
revision accepted for publication June 25, 2007.)