

## CHAPTER 11

---

# PARADATA IN WEB SURVEYS

MARIO CALLEGARO

Google London

---

### 11.1 SURVEY DATA TYPES

For a single survey instrument, it is possible to have up to four kinds of data: Substantive data (questionnaire results), Metadata, Paradata, and Auxiliary data (see Chapter 1 for more details):

*Substantive data*, also called numerical data (Mohler et al., 2008), are the core components of a survey—the results of a questionnaire that translates into a row of answer codes that populate a data file. Substantive data also contain recoded variables and recoded open-ended answers.

*Metadata* are data that describe the survey data, such as a codebook, that contains a description of the project, including the agency/firm, the dates in which the survey was in the field, and any other context information that can be relevant to interpret and use the dataset (Blank and Rasmussen, 2004).

*Paradata* (also called *process data*) are data about the process of answering the survey itself (Couper, 2000) and are collected at the respondent level (Kaczmirek, 2009, p. 79). In other words, a final dataset can contain the survey data, plus the survey paradata for each respondent. Given the fact that web surveys are self-administered, the paradata are generated by the respondents and their interaction with the survey instruments. Lastly, paradata are generally not consciously provided by the respondents (Kaczmirek, 2009, p. 79), an issue that has an effect on the privacy and ethics of online surveys and is discussed at the end of this chapter.

*Auxiliary data* are data not collected directly from the survey itself but acquired from external resources and commercial databases. Auxiliary data do not come necessarily at a respondent level of detail but can be available in aggregate forms. For example, the median income of a certain neighborhood can be appended to the original survey dataset and used for nonresponse adjustment. Auxiliary data can be used before collecting substantive data, that is, for sampling purposes or after, such as for statistical adjustments (Laaksonen, 2006). In web surveys, auxiliary data can

---

```
t=12357 : q [2] =2 £t=9544 : q [2] =1 £t=7314 : q [2] =2 £t=741 : form
submitted
```

---

**FIGURE 11.1** Example of a paradata output highlighting answer changes.

be quite useful as benchmarks to assess the quality of some variables collected in the survey.

## 11.2 COLLECTION OF PARADATA

An important technical distinction regarding the collection of paradata in web surveys is that they can be collected on the server side and/or the client side (Heerwegh, 2003, 2011). In the first case, the server collects information on a page-by-page basis, or it collects server events in the form of visits to a specific page. This can be, for example, a time stamp. Client-side paradata, as the name indicates, are data collected at the client side or at the respondent device level and can reflect events *within* a page, such as mouse clicks or change in answers. The distinction is important for two reasons. First, client-side paradata are richer in detail, precision, and amount of information that can be collected. Second, client-side paradata require placing JavaScript on the survey pages while server-side paradata require no scripting. If JavaScript is not activated, a respondent will still be able to complete the survey, but the client-side paradata will not be collected (Heerwegh, 2002).

Because server-side paradata can be collected only at the page level (Heerwegh, 2003, 2011), its value declines as more questions are placed on a single page. Most server-side paradata is not comparable to the richness of client-side paradata for the reasons delineated above. For example, server-side paradata cannot capture answer changes within a question; only the last selection on that question will be collected by the server. In contrast, on the client side, all answer changes will be collected as a sequence of events. Consider the following client-side sequence from Heerwegh (2011, p. 328).

In Figure 11.1,  $t$  stands for time elapsed (in milliseconds) from the previous mouse click, and  $q[ ]$  for question number; events are separated by the symbol £. From this sequence, we can see that the respondent first selected option 2 in question 2, then changed it to option 1, and then back again to option 2 before moving to the next page (form submitted). This example highlights the richness of paradata if collected from the client side.

In the next section, a typology of paradata is introduced. For each type of paradata, examples are provided to show the reader the kind of information that can be gathered by each and how different authors use this information.

## 11.3 TYPOLOGY OF PARADATA IN WEB SURVEYS

In web surveys, paradata can be categorized into two broad classes: device-type paradata and questionnaire navigation paradata. Device-type paradata provide information regarding the kind of device used to complete the survey, for example, a

---

```
Mozilla/5.0 (compatible; MSIE 10.0; Windows NT 6.1; WOW64;)
Trident/6.0
```

---

**FIGURE 11.2** User agent string for a PC running Windows 7<sup>1</sup>, using Internet Explorer 10.0.

---

```
Mozilla/5.0 (iPhone; CPU iPhone OS 5_0 like Mac OS X)
AppleWebKit/534.46 (KHTML, like Gecko) Version/5.1
Mobile/9A334 Safari/7534.48.3
```

---

**FIGURE 11.3** iPhone iOS 5 user agent string.

---

```
Mozilla/5.0 (Linux; U; Android 4.0.2; en-us; Galaxy Nexus
Build/ICL53F)
AppleWebKit/534.30 (KHTML, like Gecko) Version/4.0
Mobile Safari/534.30
```

---

**FIGURE 11.4** User agent string for a Samsung Galaxy Nexus S running Android Ice Cream Sandwich.

**Au:** Please provide citations for figures 11.2, 11.3, and 11.4.

desktop, a laptop, a tablet, or a smartphone. More specifically, device-type paradata indicate the browser used (user agent string); the operating system; the language of the operating system; the screen resolution; the browser window size; whether the browser has JavaScript, Adobe Flash, or other active scripting-enabled devices; the IP address of the device used to fill out the survey, the GPS coordinates, and cookies (text files placed on the visitor’s local computer to store information about that computer and the page visited). Device-type paradata are typically session-level paradata that are collected at the splash page or the first page of the survey at one point in time. Although very often respondents complete a survey in one session, multiple sessions are possible and that should be taken into account when collecting device-type paradata.

Questionnaire navigation paradata describe the entire process of filling out the questionnaire. Examples of these are authentication procedures, mouse clicks and mouse position per question/screen, change of answers, typing and keystrokes, order of answering, movements across the questionnaire (forward/backward), scrolling, number of appearances of prompts and error messages, detection of current window used, whether the survey was resumed at a later time, clicks on non-questions links (e.g., hyperlinks, F.A.Q, and help), last question answered before dropping off from the survey, and time spent per question/screen. Questionnaire navigation paradata are collected on a page-by-page or question-by-question basis and, depending on the level of detail, can completely reconstruct the survey-taking experience.

In the next sections, examples of usage for device-type and questionnaire navigation paradata are provided.

---

<sup>1</sup>Windows 7 is the marketing name for Windows 6.1.

### 11.3.1 Uses of Paradata: Device Type

Every time a browser connects to a website, it sends a sequence of text called *user agent string*. The information contained in the string is used by the website to select tailored content and formatting (e.g., the smartphone and/or tablet version of the site). User agent string is also used for statistical purposes and other reasons (Callegaro, 2010). For researchers, this kind of paradata can reveal interesting information regarding the browser used to take the survey, and by deduction, the device type and its operating system. To clarify this concept, examples of user agent strings from three devices—a PC, an iPhone, and a Samsung Galaxy Nexus S—are provided.

From a researcher’s point of view, it is important to detect the type of platform, the operating system, and the device. By looking at the bold text in the three examples, we can identify the platforms (Windows 7, iPhone, and Android), the type of device (PC, iPhone, and Galaxy Nexus), and the language of the operating system (en-us). The other information presented is more technical. If interested, the reader can consult <http://www.useragentstring.com/> for an in-depth explanation of each section of the user agent string and <http://www.texsoft.it/index.php?m=sw.php.useragent> for more details. Finally, Wikipedia (2011b) gives a broad summary of the topic.

An application of user agent strings to survey research can be found in Callegaro (2010). This publication presents three different studies analyzing user agent strings of Google advertisers answering an online customer satisfaction survey. The surveys were optimized for desktop/laptop computers only. Table 11.1 reports the results from the three studies.

Among all respondents who attempted to complete the survey, 1.2% in the first study did so from a mobile device, 2.6% in the second study, and 1.8% in the last study. As shown in the table, the breakoff rates and partial interview rates are much higher on surveys taken from a mobile device than those taken from a desktop/laptop computer (all differences are statistically significant). In the third study, more detailed paradata were collected to determine whether the respondents who started the survey via mobile phone completed it using the same device or if they switched to desktop/laptop. Among the respondents who started the survey via mobile phone, only 7.1% switched to desktop/laptop to complete it.

In a survey conducted early in 2010 on an online panel, Sood (2011) found a correlation between browser type and survey breakoff and number of items missing.

**TABLE 11.1 Summary of Breakoff Studies Using User Agent Strings**

Study	Location	Year	Breakoff Rate		Partial Interview Rate	
			Mobile	Desktop	Mobile	Desktop
1	Asia	02/10	37.4	22.0 <sup>a</sup>	18.3	12.4
2	North Amer.	06/10	24.2	8.4	6.5	3.6
3	Europe	10/10	24.7	13.4		

<sup>a</sup>  $p < 0.05$ .

---

```
Width: 1920
Height: 1080
Window width: 1391
Window height: 930
```

---

**FIGURE 11.5** Example of screen resolution and browser window size paradata.

Browsers were classified as old (Internet Explorer 7 or older and Firefox 3 or older) or new (all other browsers). The completion rate of respondents using old browsers, controlling for age, education, and race/ethnicity, was decreased by a factor of 0.75. This research shows how browser type can be used as a proxy for older computers and a potentially slower connection speed. These two characteristics, coupled with the fact that an old browser version is more likely to display a survey incorrectly, contribute to explaining the higher drop-off rates of respondents using older browsers.

The *screen resolution* and the *browser window size* are two other pieces of information that can be captured from the devices accessing the web survey. The screen resolution data results in two values: the width and the height of the screen, expressed in number of pixels. The browser window size data convey how much screen is available when excluding the browser toolbars, navigation, and scrolling bar. It tells us exactly how many pixels the respondents could have seen before doing any scrolling and whether the browser was open at full screen.<sup>2</sup> An example of screen resolution and browser window size data are reported in Figure 11.5.

The first two lines tell us that this is an HD 1080 screen with a 16:9 ratio (Wikipedia, 2011a). The other two lines show how much of the browser screen is available to view before scrolling. In this case, the browser is not open at full screen, but there is still considerable real estate shown. The browser window size is the real estate that has to be taken into account when programming web surveys (Callegaro, 2010). Baker and Couper (2007) collected browser window size paradata in a survey about energy use. They tested three versions of the web survey template: fixed at 800 by 600 pixels, fixed at 1024 by 768 pixels, and adjustable. Respondents whose monitor width was less than 1024 pixels had a higher mean completion time on the fixed template of 1024 by 768 compared to respondents in the other two conditions. Two explanations are possible: these respondents had to do horizontal scrolling to properly see the questions and that added extra time and effort, or a monitor with less than 1024 width resolution is, by definition, older, as might be the computer attached to it. This can add extra time to the navigation of the survey. They also found that respondents whose browser width was less than 800 pixels were more likely to break off in the fixed template of 1024 by 768 size compared to respondents in the two other conditions. The greater number of breakoffs could be due to the amount of horizontal scrolling necessary to properly see the questions; however, there are currently no studies to support this theory.

---

<sup>2</sup>The readers can point their browsers to the following site: <http://www.javascripter.net/faq/browserw.htm> in order to know their browser window size.

The *detection of JavaScript and Flash* is very useful to verify what a respondent can see and do in a survey. JavaScript adds functionalities to web pages. For web surveys, JavaScript is used to perform question validations, provide message feedbacks and help, and in general, to enhance the interactivity of a question screen. In current browsers, JavaScript is, by default, generally enabled. It is estimated that about 2% or less of computer users have JavaScript disabled (Zakas, 2010). Flash is used for advanced question types, such as drag and drop (rank order) question slider bars. If the respondents do not have Flash installed on their computers, they will be prompted to install it; without Flash installed, they might not be able to see that particular question. It is important to note that respondents who use iPads to answer a survey are not be able to use Flash technology, as it is not supported by Apple on iPhones and iPads. Based on a survey of Lightspeed opt-in panel members, Adobe estimates installation of Flash Player 10 and below at 98.7% in mature markets (Adobe, 2011). It is important to measure these rates in the specific survey the researcher is conducting, as they are useful statistics that can add more insight about data quality.

*IP address* information can be used in different ways. For example, in online psychological experiments, IP address paradata were used to delete duplicate submissions for the same experiment (Reips). The use of IP address to detect multiple submissions is becoming less and less effective due to higher use of dynamic IP addresses than in the past and the share of the same IP address from multiple devices. The IP address can also be used to estimate the location of the respondent while taking the survey. The collection of IP addresses needs to be carefully evaluated by the researchers because of its sensitive nature. As discussed later in this chapter, depending on what legislation develops, IP address paradata may be considered personally identifiable information; therefore, a survey would not be considered anonymous if the IP address is collected.

*GPS coordinates can also be useful information for researchers.* Some devices, such as smartphones and tablets, have a built in GPS receiver. With the user's consent, the GPS receiver can provide coordinates to the data collection agency that can be used to track the location where the survey was conducted. As an example, GPS coordinates of interviewers administering a survey with a cell-enabled iPad were used to check the location of the interviewers in the field for quality assurance reasons (Dayton and Driscoll, 2011).

*Cookies* are another source of valuable information. According to the European Society for Opinion and Market Research (ESOMAR, 2011a) Guideline for Online Research, they are defined as follows:

Cookies are small text files stored on a computer by a website that assigns a numerical user ID and stores certain information about your online browsing. Cookies are used on survey sites to help the researcher recognize the respondent as a prior user as well as for other survey control or quality functions. The data stored on cookies is [*sic*] not personalized and they can be rejected or deleted through browser settings (ESOMAR, 2011a, p. 8).

Cookies are associated with a specific domain, which can access and modify them. The information is stored with a creation and expiration time stamp. Some companies

**TABLE 11.2 Percent of Cookies Deleted per Month by Country**

Country	First party	Third party
Australia	27.7	36.6
Brazil	33.4	40.4
France	27.0	35.3
Germany	22.8	30.4
New Zealand	28.0	36.4
United Kingdom	26.8	35.0
United States	28.5	34.8
Average	27.7	35.6

**Au: Please provide citation for Table 11.2.**

use them to ensure that only one survey has been taken from the same device. The use of cookies in web surveys is, however, not very well documented. For this reason and due to the European Union’s e-Privacy Directive as well as pressure from the Center for Democracy and Technology, in September 2011, ESOMAR (2011b) launched a survey on the use of cookies and tracking technologies. Also, according to the ESOMAR guideline, “Researchers must include clear, concise, and conspicuous information about whether they use cookies and if so why” (2011a, p. 8).

Because cookies can be deleted and rejected via the browser setting, it is useful to provide an estimate of cookie deletion. In one of the few studies available on cookie deletion, Comscore (2011) recently estimated the average deletion rate in seven countries. For first-party cookies (those delivered directly by the website hosting the content), average deletion was estimated at 27.7% per month; for third-party cookies (those associated with objects, for example, advertisements that are delivered by a third party), it was 35.6% per month. Yahoo was the website used for first-party cookies, and for third-part cookies, the website was Double-Click, which uses cookies to count and serve targeted advertisements.

In a previous study conducted in the United States using the same methodology as the Comscore study, first-party cookie deletion was estimated at 31% and third party at 27% (Abraham et al., 2007). Singer and Couper (2011) make a valid point that the data presented in the Comscore study are collected on Internet users who agreed to have their Internet behavior tracked online, and for this reason, this can be an underestimation of the phenomenon. Conversely, it can also be an overestimation of the real percentages because the panel members know that they are being tracked. Regardless, these studies provide evidence that some users delete their cookies on a regular basis and that some proportion of the Internet population is concerned about them. From a survey data collection point of view, if cookies are used by the survey organization, their deletion by some of the respondents will introduce missing values in the cookies paradata.

### 11.3.2 Uses of Paradata: Questionnaire Navigation

In this section, examples of usage of questionnaire navigation paradata are presented. *Authentication procedures* capture successes and failure to log into a survey in case

```

lXNtoilre7.2|1|M677j13j1320#
M548|174|830#
M160|101|1750#
M366|192|550#
M728|4|7690#
M489|247|610#
C493|229|3301#
R110|1#
C493|280|4301#
R110|3#
C493|345|3901#
R110|5#
C521|399|3801#
SU521|399|60|undefined#

```

FIGURE 11.6 Navigation strings from Stieger and Reips (2010, p. 1490).

the respondent has to enter a login and password. One common case is when the survey invitation is given in a letter. For example in the April 2011 American Community Survey Internet test Horwitz et al. (2012) computed failed login rates where respondent had to enter a 10 digit user ID and 4 digit PIN they received in the mail.

*Mouse clicks* and *mouse position* can be captured using JavaScript. For example, Stieger and Reips (2010) were able to capture 336,262 single actions of 1046 participants during an online questionnaire. An example of navigation paradata is shown in Figure 11.6.

In Figure 11.6, `lXNtoilre7.2` is the ID assigned to the respondent. `M` stands for mouse, and the first number is the position of the mouse on the  $x$ -axis while the second number is the position on the  $y$ -axis. The third number in each row is the time spent for that action in milliseconds. `C` stands for click and `R` for radio buttons while `SU` stands for submit. Using the overall length of the mouse track for each page, the authors of this study were able to measure that 11% of questionnaires showed excessive mouse movements (defined as below and above  $\pm 2$  standard deviations from the mean).

*Change of answers* is an indicator of potential confusion with a question and paradata of these changes can be used to improve questionnaire design. Although there is research on change-of-answer analysis using paper and pencil questionnaires (Christian and Dillman, 2004), these data require tedious manual coding to be collected. Change-of-answer paradata in online surveys are easier to collect. The navigation string presented in Figure 11.6 includes change of answers. In this case, the respondent selected radio button 1 (line 8 in the code), then changed the answer to 3 (line 10) and finally to 5 (line 12) before submitting (`SU`). Stieger and Reips (2010) showed that changes of answers for opinion questions were more frequent (5.4%) than for factual questions (1.5%). They also provided evidence of how semantic differential questions are problematic to answer, from a respondent's point of view, due to their high number of answer changes. In an online survey of Washington State



students, Stern (2008) found that when a questions with a 5-point response option scale was presented with only the endpoint labeled, respondents were more likely to change their answers in a reciprocal fashion (e.g., from 1 to 5 and vice versa) compared to when the scale was fully labeled. The author of the study attributes the finding to respondents having more difficulty in interpreting the direction of the scale when there are only endpoint labels. Stern also used paradata on change in answers (checking and unchecking a box) in a question with a check-all-that-apply format to explain the phenomenon called subtraction effect (Mason et al., 1994). A subtraction effect in questionnaire design happens when important considerations are taken out of the thought process for a later question, that is, respondents avoid giving redundant information. For example, students had trouble discerning libraries from library instructions when asked about resources used at their university in a check-all-that-apply format. When “libraries” was placed before “library instructions,” 20% of students selected library instructions. When “library instructions” was placed before “libraries,” the percent endorsing it raised to 52% (Stern, 2008).

In an application of paradata to improve establishment surveys, Haraldsen (2005) created a quality index based on the *number of prompts, error messages, and data validation messages*. Subsequently, the quality of online questionnaires at Statistics Norway is now evaluated using the following formula:

$$\text{Quality index} = 1 - \frac{\sum \text{activated errors}}{\text{possible errors}} \frac{1}{\text{number of respondents}}$$

The number of possible errors is the sum of all potential prompts, all potential error messages, and all validation messages programmed in the web instrument. The number of activated errors is a count of how many prompts, error messages, and data validation messages were activated during the survey by the respondents. In the best-case scenario, this number would be 0, which means the respondent was able to fill out the web survey without generating any prompt, error, or data validation messages. The goal is to decrease the number of activated errors by improving the visual design and the clarity of the questionnaire, together with improving the wording of each question.

The *detection of the current window used* is a very insightful type of paradata that is not widely used in web surveys. It is possible to collect what window is currently activated by the respondent. This can tell us, for example, when and where the respondents are doing something else during the survey, such as reading an email. This type of paradata can enlighten the interpretation of extremely long time spent on a question screen and help cleaning the time latency data. We could not find a survey example but we did find an example in the webinar software area. Some webinar software provide a dashboard to the instructor with a measure of attentiveness. In other words the software computes the percent of attendees that have not the webinar window open as their current window used. From their point of view, the higher this number, the more attendees are distracted or doing something else.

*The last question answered before dropping off* is probably the most common type of paradata used in online surveys. This information determines if a survey

can be classified as complete (all/most questions answered), partial (key questions answered) or breakoff (not enough questions answered to be counted as complete in the final dataset). The classification of a survey as complete, partial, or breakoff is the initial component used when computing response rates (American Association for Public Opinion Research, 2011).

The number of breakoffs per question type was analyzed by Peytchev (2009). The author found that open-ended questions increased the chances of breakoff by almost 2.5 times (vs. a closed question), long questions by 3 times (vs. a short question), slider bars by almost 5 times (vs. a radio button question), and introductory screens by 2.6 times (vs. introduction screens). Sakshaug and Crawford (2010) in a national survey of college students analyzed breakoff rate by question and found that about 16% of all breakoffs happened at a specific question of a very sensitive nature, specifically when students were asked permission to access their school records.

Using a combination of paradata regarding item nonresponse and last question answered before dropping off a survey, Bosnjak (2001) was able to create a typology of seven response patterns in web surveys. The most notables are “unit nonrespondents,” subjects who break off from the survey after the welcome screen has been displayed and the “lurkers” who, if allowed, check out the entire questionnaire but have few or no responses.

*Time spent per screen or time latency* is another common type of paradata that has generated numerous publications and has provided key insights in the survey response process. Before online surveys, the collection of time latency data was confined to computer labs (Fazio, 1990), interviewers, or CATI systems (Mulligan et al., 2003). With online surveys, time latency data can be collected on the server side or on the client side. The server-side data collection contains download times and other extra components as compared to the client-side data. For example, Yan and Tourangeau (2008) estimate that response times collected on the server side are 3–4 s longer than on the client side. Kaczmirek (2009, p. 91) offers an in-depth explanation of this. From a researcher’s point of view, client-side response latency is preferred over server side because it offers a true value of response time and provides the level of precision necessary to answer the research question.” The good news is that the correlation between these two measures is very high and ranges from 0.91 to 0.99 in the Yan and Tourangeau (2008) study and between 0.94 and 0.99 in Kaczmirek (2009, study 3).

A good review on time latency paradata used to either shed light on the response process or to assist in the development of a theory is found in Couper and Kreuter (2013). Due to space considerations, only four different applications of time latency paradata with the relevant insights from each study are presented here. Other examples not discussed here include, for example, Lenzner et al. (2010) and Ranger and Ortner (2011).

Callegaro et al. (2009) used response latency to measure time spent on items of an online personality assessment. Job applicants (optimizers) took more time answering questions in the assessment tool than did job incumbents (satisficers). This result supports the perspective that deeper cognitive processing requires greater effort and takes more time.

Heerwegh (2003), using time latency collected in a web survey, reproduced an experiment conducted by Bassili and Fletcher (1991), where response time was collected by telephone interviewers. The results of the web experiment validated the previous experiment, showing that respondents with weaker attitudes take more time in answering survey questions than respondents with stronger, more stable attitudes.

Haraldsen et al. (2005) used a combination of paradata—time latency and percent of respondents changing their answers—to identify problematic questions in an online customer satisfaction survey conducted by Statistics Norway.

Lastly, Yan and Tourangeau (2008) were able to provide evidence that higher-educated respondents responded faster than lower-educated subjects, and younger respondents faster than older respondents. Survey experience also made a difference; subjects who had completed at least 15 online surveys were faster than subjects with less survey experience. Question characteristics were also of interest; demographic questions took less time to answer than factual and attitudinal questions. Finally, respondents became faster as they approached the end of the questionnaire.

Questionnaire navigation paradata can be collected on the different devices that are used to conduct surveys. Although the examples above generally refer to desktop/laptop or smartphone/tablet paradata, it is possible to collect them on other devices, such as a personal digital assistant (PDA) as shown by McClamroch (2011).

#### 11.4 USING PARADATA TO CHANGE THE SURVEY IN REAL TIME: ADAPTIVE SCRIPTING

All the examples of paradata usage presented thus far have been to explain survey processes and respondents' behaviors as well as to shed light on the quality of some answers. Another use of paradata that was pioneered in the early 2000s by Jeavons (2001) is *adaptive scripting*. Adaptive scripting refers to using paradata in real time to change the survey experience for the respondent. Jeavons explains adaptive scripting as a way to mimic what a good interviewer would do, such as adapting to the respondent's needs, encouraging a respondent who is about to give up to complete the survey, or reducing the speed of reading the questions to adjust to the respondent's pace. Because web surveys have little flexibility, paradata can be a solution to this issue.

Conrad et al. (2007) experiment 2, for example, provided clarification for questions based on the respondents' inactivity. After a certain time threshold, the system provided a clarification on the question. Another example of adaptive scripting is using time latency data to trigger specific prompts. Conrad et al. (2011) showed a prompt to respondents who were answering too quickly, encouraging them to provide accurate answers. The prompts reduced straightlining without increasing breakoffs and also improved the accuracy of answers among respondents who had some college or associate degree level of education. These two examples clarify how Jeavons' idea of adaptive scripting in web surveys can be executed using paradata as the trigger.

### 11.5 PARADATA IN ONLINE PANELS

All the examples and the studies cited above refer to use of paradata for single web surveys. There are however a separate class of paradata that are available in online panels. Because in online panels the entire history of each member can be stored, a new class of paradata emerges. Examples are number of survey invitations received, number of surveys completed, topic of survey completed, last survey completed, and many more. For example Tortora (2008) found a nonlinear relationship between number of survey invitation and attrition in the Gallup online panel. Panel members attrited at a higher level after one of the two survey requests (early attrition). Then the attrition rate stabilized up to 13–15 survey request where it increased sharply again. He also found a relationship between survey topic and attrition. By giving a score on survey topics as being poll-like/social surveys or market research surveys, the author found that attrition rate was lower for panel members who completed relatively more poll-like or social surveys than market research surveys. Paradata for online panels is definitely a less explored and published topic.

### 11.6 SOFTWARE TO COLLECT PARADATA

There are two main classes of software to collect paradata: specific paradata software and paradata collection tools embedded in commercial and non-commercial survey platforms.

Specific paradata software includes Client Side Paradata (CSP) by Dirk Heerwegh, which is probably the first freely available client-side paradata script. At the time of this publication, CPS Version 3.0. is a JavaScript component that can be added to virtually any web survey. The script detects “clicking hyperlinks,” “manipulating input elements” (radio buttons, check boxes, drop-down boxes, text fields, and text areas), and “submitting the form.” These actions are captured and recorded along with a time stamp. The script is available on Dirk Heerwegh’s personal homepage at <https://perswww.kuleuven.be/~u0034437/public/csp.htm>, and a detailed description can be found on this webpage and in Heerwegh (2002, 2003).

An extension of the CSP lies within the Universal Client Side Paradata (UCSP) project (Kaczmirek, 2009). It requires only a single code insertion per survey and collects all events on a single web page. The script is free to use and can be found on Lars Kaczmirek’s personal homepage at <http://www.kaczmirek.de/ucsp/ucsp.html> and a detailed description can be found in Kaczmirek (2009) and Kaczmirek and Neubarth (2007).

A second type of software, which comes from the online psychological experiment literature, is User-ActionTracer (UAT) (Stieger and Reips, 2010). UAT is a piece of code that tells the participant’s web browser to store information, including timing of all actions, all mouse clicks (single and double), choices in drop-down menus, radio buttons, all inserted text, key presses, and the position of the mouse pointer. An example of the output can be found in Figure 11.6. The tool is available by request from Stefan Stieger. Collecting paradata using commercial and non-commercial software

varies quite dramatically. More generally, we can say that some survey platforms allow inserting specific JavaScript in the survey and storing paradata in a hidden variable.

In a study conducted at the University of Ljubljana, Kavčič et al. (2012) reviewed 365 web survey platforms. Of 143 of those, they were able to obtain a free demo version. When looking at three basic paradata features, the evaluation was not very positive. Thirty percent of platforms provided paradata for total survey time, 4% provided page by page time latency paradata, and only 1% provided paradata on page tracking such as movement across the questionnaire. For 8% of survey platforms these three types of paradata collection was unclear.

### 11.7 ANALYSIS OF PARADATA: LEVELS OF AGGREGATION

At the post data collection stage, researchers can analyze paradata at different levels of aggregation. According to Kaczmirek (2009, p. 83), there are four levels of aggregation of paradata that have an impact on their collection and analysis.

1. At the first level of analysis, individual respondent's actions, such as mouse clicks, the location of the mouse, or any change of answers are recorded sequentially, but the number of individual actions cannot be predetermined. Paradata at this level create non-rectangular matrix datasets, making data analysis more challenging.
2. At the second level, paradata from the first level – for example, the number of mouse clicks on a page or the number of changes per answer – are aggregated across actions but within the same respondent. These data are rectangular in nature and easier to analyze. This second level is more focused and allows for better formulating of research questions.
3. At the third level, second-level paradata are aggregated across respondents or variables. For example, researchers can compute the average number of answer changes for the entire questionnaire *per respondent* or the average number of answer changes for a specific question *across all respondents*.
4. The fourth level is the highest level of aggregation. Data at this level are aggregated across respondents *and* variables. It provides a single number per survey, such as the average survey length or the survey response rate.

First-level paradata can overwhelm the researcher because of its volume and the non-rectangular data format. They do, however, provide the highest level of detail for analysis. The strategy is to decide in advance what research questions should be answered and, if possible, on what questions and or sections of the questionnaire. For example, if a survey contains a question wording experiment, focusing on first-level paradata for that specific experiment can be useful as another tool to assess the quality of the data. Second-level paradata are easier to collect and analyze, and some survey platforms provide them already precoded. For example, some survey

**274** PARADATA IN WEB SURVEYS

platforms collect and parse user agent strings automatically and tell what device was used to complete the survey. If the survey platform does not parse user agent strings automatically, there are external tools, such as the ones available at <http://user-agent-string.info/download> that can accomplish the same task. Third-level paradata are the sum of second-level paradata and provide useful summary statistics for the survey. Examples of fourth-level paradata are response rates, average time to take the survey, and percent of respondents who took the survey on a mobile phone.

**11.8 PRIVACY AND ETHICAL ISSUES IN COLLECTING WEB SURVEY PARADATA**

Some types of paradata information, such as the IP address, constitute a form of personally identifiable information that together with e-mail and/or a mailing address can be used to identify a respondent. These data need to be protected (as explained later in this chapter) and treated carefully by the survey company and the data users. The significance of collecting IP addresses is reflected in a separate section in the new ESOMAR guideline, which states the following:

An IP address might constitute personal data in combination with other identifiable data but there is no international consensus about the status of IP addresses. They can often identify a unique computer or other device, but may or may not identify a unique user. Accordingly, ESOMAR requires compliance with the relevant national and/or local law and/or regulation if it classifies IP addresses as personal data ESOMAR (2011a, p. 8).

The use of cookies by survey and online panel companies, as previously discussed, poses privacy and ethical questions. For example, the World Wide Web Consortium (W3C) is working on a document defining the mechanisms for the so-called Do Not Track (DNT) HTTP header, which is basically a way a user can easily disable any tracking from a particular website (W3C, 2012). The DNT standard is under discussion in the United States Congress and by the Federal Trade Commission. European regulators have expressed concerns about online tracking as well. The Council of American Survey Research Organizations (CASRO) and ESOMAR recently expressed their concerns about the DNT and its implication for online market research and online panel management. The two associations have argued that “regulations should be limited to tracking for online behavioral advertising purposes and not extend to legitimate research, which is distinct from advertising and marketing” (CASRO and ESOMAR, 2011), (Stark, 2011, p. 1).

The ethical issues involved in collecting paradata are emerging topics that are now under frequent discussion. Should the respondent be informed that the researcher is capturing and using paradata? If so, how should this be communicated? In a recent study, Singer and Couper (2011) asked members of the Dutch Longitudinal Internet Studies for the Social Sciences (LISS) probability-based panel if the researcher could use the members’ paradata. The authors varied three different descriptions of paradata. The request came at the end of a survey, and about 38.4% of respondents

across the three conditions agreed.<sup>3</sup> The study was repeated in the United States on the Knowledge Networks' probability-based panel (Couper and Singer, 2011). This time there were five conditions (different ways to describe paradata). About 44.8% of respondents across the five conditions agreed to do the survey and permitted paradata usage. Finally, in similar study that was conducted in a U.S. opt-in panel, the percentage of respondents willing to do the survey and allow for use of their paradata was of 63.4% across the five conditions (Couper and Singer, 2011). This is evidence that asking respondents for permission to use their paradata might make them less willing to participate in a survey. It is also possible that the manipulations used in the experimental conditions might have drawn undue attention to the paradata issue.

Ethical and communication issues are important considerations in using web survey paradata. New regulations might quickly change the way IP addresses and cookies are collected by websites and, by extension, by web survey software and online panels. This is definitely a topic of rapid change, and the readers are urged to follow carefully the new developments in this area.

## 11.9 SUMMARY AND CONCLUSIONS ON PARADATA IN WEB SURVEYS

There are numerous types of paradata that can be collected in web surveys. As technology evolves, new types of paradata will be available to the researcher. Upcoming devices and development in software code will make possible to collect different types of paradata information. This chapter has been updated until the very last proofs, but emerging types of paradata could not have been added to the discussion. We provided the state of the art of paradata for web surveys at the current time, but we invite the reader to follow the topic and experiment with types and analysis of paradata.

After placing paradata in relationship to other types of data collected and used (substantive data, metadata, and auxiliary data), this chapter proposed a taxonomy of paradata types: device-type paradata and questionnaire navigation paradata. The first type identifies the kind of device used to answer the survey, and the second tell us all other pieces of information that are useful to understand how the respondent was able to visualize and interact with the online questionnaire. Questionnaire navigation paradata allow us to reconstruct the process of filling out a questionnaire for each respondent as if we were looking at the respondent's screen when answering questions. For each type of paradata, one or more examples of usage was provided. An underlying goal of this chapter is to encourage the reader to collect and analyze online survey paradata. The examples offered here are to inspire new ways to use and analyze paradata.

---

<sup>3</sup>This number is the correct number reported in Couper and Singer (2011) and differs from what is reported in Table 6.2 in Singer and Couper (2011, p. 157) due to a coding error discovered after publication. (Couper, personal communication).

In most of the examples of device-type or questionnaire navigation paradata presented in this chapter, the information was used *after* the survey had been completed (with the exception of the detection of user agent strings, which are used to route each device to the appropriately formatted screen (e.g., desktop vs. smartphone). Paradata, however, can be used *in real time* to change the behavior of a survey with the use of adaptive scripting. Some examples of software that collects paradata were provided. Once paradata are collected, there are numerous levels of aggregation possible. The taxonomy provided by Kaczmirek (2009) is useful to clarify which level of aggregation is possible and what kind of information each level can provide. Lastly, the reader should be aware of the current debate on data privacy and what kind of information can be collected on survey respondents. Cookies and IP addresses, for example, are paradata that contain private information. At the same time, asking for permission in using the respondents' paradata creates some concerns, and in the few studies presented, a good number of respondents did not agree to share paradata with the survey organization when asked to do so (Couper and Singer, 2011; Singer and Couper, 2011).

Paradata in web surveys can give great insight into the response process. A good use of paradata is in questionnaire development and testing to supplement the pretesting of questionnaires. Although it has been suggested that the collection of paradata comes at a low cost (initial set up and extra programming), Nicolaas (2011) asserts that the cost to analyze and report them should not be underestimated. This is especially true for first-level paradata. Paradata should not be viewed as the only means to study response quality and improve question wording and data collection, but rather as one of many tools available for use in this capacity in online surveys.

Lastly, we invite the readers to ask for paradata when using a vendor, for example an online panel, or when using one of the online survey platform solutions that allow creating web surveys. The request should always be mindful of the data privacy issues we discussed above, and should be in line with the terms and conditions of panel membership in case an online panel is being used. The production of paradata should be considered best practice and one of the types of data produced in a survey (substantive data) together with metadata and, when necessary, auxiliary data.

## REFERENCES

- Abraham, M., Meierhofer, C., and Lipsman, A. (2007). The Impact of Cookie Deletion on the Accuracy of Site-server and Ad-server Metrics: An Empirical comScore Study.
- Adobe (2011). Flash Player Version Penetration. [http://www.adobe.com/products/player\\_census/flashplayer/version\\_penetration.html](http://www.adobe.com/products/player_census/flashplayer/version_penetration.html).
- American Association for Public Opinion Research (2011). *Final Dispositions of Case Codes and Outcomes Rates for Surveys*. AAPOR, 7th edition.
- Baker, R. and Couper, M. (2007). The Impact of Screen Size, Background Color, and Navigation Button Placement on Response in Web Surveys. *Paper presented at the 9th General Online Research Conference, Leipzig, Germany, March 26–28, 2007*.
- Bassili, J. and Fletcher, J. (1991). Response-Time Measurement in Survey Research a Method for CATI and a New Look at Nonattitudes. *Public Opinion Quarterly*, 55(3):331–346.



- Blank, G. and Rasmussen, K. (2004). The Data Documentation Initiative. The Value and Significance of a Worldwide Standard. *Social Science Computer Review*, 22(3):307–318.
- Bosnjak, M. (2001). Participation in Non-restricted Web Surveys: A Typology and Explanatory Model for Item Non-response. In Reips, U.-D. and Bosnjak, M., editors, *Dimensions of Internet Science*, pages 193–207. Zagreb Pabst Science Publishers.
- Callegaro, M. (2010). Do You Know Which Device Your Respondent Has Used to Take Your Online Survey? *Survey Practice*.
- Callegaro, M., Yang, Y., Bhola, D., Dillman, D., and Chin, T. (2009). Response Latency as an Indicator of Optimizing in Online Questionnaires. *Bulletin de Methodologie Sociologique*, 103(1):5–25.
- CASRO and ESOMAR (2011). ESOMAR AND CASRO Submission to the W3C Tracking Protection Working Group. Market Research Techniques That Use Cookies and Tracking Technologies.
- Christian, L. and Dillman, D. (2004). The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-administered Questions. *Public Opinion Quarterly*, 68(1):57–80.
- Comscore (2011). The Impact of Cookie Deletion on Site-server and Ad-server Metrics in Australia. An Empirical comScore Study.
- Conrad, F., Schober, M., and Coiner, T. (2007). Bringing Features of Human Dialogue to Web Surveys. *Applied Cognitive Psychology*, 21(2):165–187.
- Conrad, F.G., Tourangeau, R., Couper, M.P., and Zhang, C. (2011). Interactive Interventions in Web Surveys can Increase Response Accuracy. *Paper presented at Annual Conference of the American Association for Public Opinion Research*.
- Couper, M. (2000). Usability Evaluation of Computer-Assisted Survey Instruments. *Social Science Computer Review*, 18(4):384–396.
- Couper, M. and Kreuter, F. (2013). Using Paradata to Explore Item-level Response Times in Surveys. *Journal of the Royal Statistical Society, Series A*, 176(1).
- Couper, M. and Singer, E. (2011). Ethical Dilemmas in Dealing with Web Survey Paradata.
- Dayton, J. and Driscoll, H. (2011). The Next CAPI Evolution - Completing Web Surveys on Cell-enabled iPads. In *Proceedings of the 66th AAPOR Annual Conference*. [http://www.aapor.org/source/AMProceedings/files/2011/05-13-11\\_4F\\_Dayton.pdf](http://www.aapor.org/source/AMProceedings/files/2011/05-13-11_4F_Dayton.pdf).
- ESOMAR (2011a). ESOMAR Guideline for Online Research.
- ESOMAR (2011b). New ESOMAR Survey on Use of Cookies and Tracking Technologies.
- Fazio, R.H. (1990). A Practical Guide to the Use of Response Latency in Social Psychological Research. In Hendrick, C. and Clark, M.S., editors, *Review of Personality and Social Psychology, Research Methods in Personality and Social Psychology*, volume 11, pages 74–97. Sage Publications.
- Haraldsen, G. (2005). *Using Client Side Paradata as Process Quality Indicators in Web Surveys*.
- Haraldsen, G., Kleven, Ø., and Sundvoll, A. (2005). Big Scale Observations Gathered with the Help of Client Side Paradata. In Netherlands, S., editor, *Proceedings of the 5th Quest Workshop*, pages 27–40. Statistics Netherlands, Heerlen, NL.
- Heerwegh, D. (2002). Describing Response Behavior in Websurveys Using Client Side Paradata. *Paper presented at the International Workshop on Web Surveys held at ZUMA, Mannheim, Germany, October 25, 2002*.
- Heerwegh, D. (2003). Explaining Response Latencies and Changing Answers Using Client-Side Paradata from a Web Survey. *Social Science Computer Review*, 21(3):360–373.

## 278 PARADATA IN WEB SURVEYS

- Heerwegh, D. (2011). Internet Survey Paradata. In Das, M., Ester, P., and Kaczmirek, L., editors, *Social and Behavioral Research and the Internet. Advances in Applied Methods and Research Strategies*, pages 325–348. Taylor and Francis.
- Horwitz, R., Guarino Tancreto, J., Zelenak, M., and Davis, M. (2012). *Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument*. United States Census Bureau, Washington D.C.
- Jeavons, A. (2001). Paradata: Concepts and Applications. In *ESOMAR Net Effects 4 conference, Barcelona, Spain*.
- Kaczmirek, L. (2009). *Human Survey-interaction: Usability and Nonresponse in Online Surveys*. Herbert Von Halem Verlag.
- Kaczmirek, L. and Neubarth, W. (2007). *Nicht-reaktive Datenerhebung: Teinahmeverhalten bei Befragungen mit Paradata evaluieren. [Non Reactive Data Collection. Evaluating Response Behavior with Paradata in Surveys]*, pages 293–311. Herbert von Halem Verlag.
- Kavčič, L., Lenar, J., and Vehovar, V. (2012). *Survey Software Features Review. A WebSM Study*.
- Laaksonen, S. (2006). Need for High Quality Auxiliary Data Service for Improving the Quality of Editing and Imputation. In United Nation Statistical Commission, 3 Impact on data quality, editor, *Statistical Data Editing*, pages 334–344. United Nations.
- Lenzner, T., Kaczmirek, L., and Lenzner, A. (2010). Cognitive Burden of Survey Questions and Response Times: A Psycholinguistic Experiment. *Applied Cognitive Psychology*, 24(7):1003–1020.
- Mason, R., Carlson, J.E., and Tourangeau, R. (1994). Contrast Effects and Subtraction in Part-whole Questions. *Public Opinion Quarterly*, 58(4):569–578.
- McClamroch, K.J. (2011). Evaluating the Usability of Personal Digital Assistants to Collect Behavioral Data on Adolescents with Paradata. *Field Methods*, 23(3):219–242.
- Mohler, P.P., Pennell, B.-E., and Hubbard, F. (2008). Survey Documentation: Towards Professional Knowledge Management in Sample Surveys. In De Leeuw, E., Hox, J., and Dillman, D., editors, *International Handbook of Survey Methodology*, pages 403–420. Lawrence Erlbaum.
- Mulligan, K., Grant, T., Monson, Q., and Mockabee, S. (2003). Response Latency Methodology for Survey Research: Measurement and Modeling Strategies. *Political Analysis*, 11(3): 289–301.
- Nicolaas, G. (2011). Survey Paradata: A Review. *National Centre for Research Methods*.
- Peytchev, A. (2009). Survey Breakoff. *Public Opinion Quarterly*, 73(1):74–97.
- Ranger, J. and Ortner, T.M. (2011). Assessing Personality Traits Through Response Latencies Using Item Response Theory. *Educational and Psychological Measurement*, 71(2): 389–406.
- Reips, U. Theory and Techniques of Web Experiments. In Batanic, B., Reips, U., and Bosnjak, M., editors, *Online Social Sciences*, pages 229–250. Hogrefe and Huber.
- Sakshaug, J. and Crawford, S.D. (2010). The Impact of Textual Messages of Encouragement on Web Survey Breakoffs: An Experiment. *International Journal of Internet Science*, 4(1): 50–60.
- Singer, E. and Couper, M.P. (2011). Ethical Considerations in Internet Surveys. In Marcel Das, P.E. and Kaczmirek, L., editors, *Social and Behavioral Research and the Internet. Advances in Applied Methods and Research Strategies*, pages 133–162. Routledge.

**Au: Please provide pubyear for Ref. "Reips".**

- Sood, G. (2011). *Poor Browsers' and Internet Surveys*.
- Stark, D. (2011). Do Not Track Gathers Momentum. *Research World*, pages 40–41.
- Stern, M. (2008). The Use of Client-Side Paradata in Analyzing the Effects of Visual Layout on Changing Responses in Web Surveys. *Field Methods*, 20(4):377–398.
- Stieger, S. and Reips, U. (2010). What Are Participants Doing While Filling in an Online Questionnaire: A Paradata Collection Tool and an Empirical Study. *Computers in Human Behavior*, 26(6):1488–1495.
- Tortora, R. (2008). Recruitment and Retention for a Consumer Panel. In Lynn, P., editor, *Methodology of Longitudinal Surveys*, pages 235–249. Wiley and Sons, Inc.
- W3C (2012). Tracking Preference Expression (DNT).
- Wikipedia (2011a). Display Resolution.
- Wikipedia (2011b). User Agent.
- Yan, T. and Tourangeau, R. (2008). Fast Times and Easy Questions: The Effects of Age, Experience and Question Complexity on Web Survey Response Times. *Applied Cognitive Psychology*, 22(1):51–68.
- Zakas, N.C. (2010). How Many Users Have JavaScript Disabled?