11

Božidar Kliček, Ph.D.
Sandro Gerić, M.Sc.
Nina Begičević, MA.
University of Zagreb
Faculty of Organization and Informatics

DATA COLLECTION IN THE GREEN PATROL PROJECT

Zbiranje podatkov v projektu zelene patrulje

Abstract

UDC: 001.891:504.06

This article describes the methodology of a complex survey which uses multimodal delivery methods for data collection. Its general principle is how to reach all the needed profiles in a prompt manner and in different circumstances. This methodology is illustrated with the Green Patrol, an ecological project organized as a part of E!2584 Ulixes – Intelligent Tourist Organization project. This pilot project researches the present state of the environment and some environmental incidents as well. The Green Patrol project was performed in a local community of the town of Varazdin during a four week period at the end of the year 2004. Satisfaction with the present state and environmental protection was surveyed on-line, via SMS, and a paper survey questionnaire.

Key words: multimodal survey methodology, multimodal delivery methods, paper survey, web survey, SMS survey

Izvleček

UDK: 001.891:504.06

Članek predstavlja metodologijo kompleksne ankete, ki uporablja več različnih načinov zbiranja podatkov. Njeno splošno načelo je, najti vse potrebne profile hitro in v različnih okoliščinah. Metodo ilustriramo z Zeleno patruljo, ekološkim projektom, ki so ga organizirali v okviru E!2584 Ulixes – Projekta Inteligentne turistične organizacije. Ta pilotski projekt raziskuje trenutno stanje okolja in tudi nekaj ekoloških incidentov. Projekt Zelene Patrulje so izvajali štiri tedne ob koncu leta 2004 v lokalni skupnosti v Varaždinu. Zadovoljstvo s sedanjim stanjem in z varovanjem okolja so preverjali preko interneta, preko mobilnega omrežja in z klasičnim vprašalnikom.

Ključne besede: multimodalna metodologija anketiranja, multimodalne metode zbiranja podatkov, klasična anketa, internetna anketa, SMS-anketa

JEL: C 81

1 Introduction

This article describes the methodology of a complex survey which uses multimodal delivery methods for data collection in the field of marketing and social sciences.

The main reason for using it is the possibility of reaching all the needed profiles in a prompt manner and in different circumstances. The effects of these methods can be seen through their effect on the respondent. We have analyzed this aspect because we presume that respondents of different ages and different educational backgrounds have different preferences for different types of survey technologies (e.g. paper, web and SMS questionnaires). The impact of these methods on data quality as well as others (e.g. cost efficiency) is not within the main scope of this paper.

Our multimodal survey methodology will be illustrated with the Green Patrol, an ecological project with large public participation, organized as a part of E!2584 Ulixes – Intelligent Tourist Organization project. This pilot project researches the present state of the environment and environmental protection. Satisfaction with the present state of the environment and environmental protection is surveyed on-line (web), via SMS, and with a paper survey questionnaire. Environmental incidents (red points) and positive examples (green points) are collected through e-mail, MMS, interactive ecological maps, and »Green telephone«.

The focus of this article will be the survey about satisfaction with the present state of the environment and environmental protection, and the reasons for using multimodal delivery methods (on-line/web, SMS and a paper survey questionnaire) for data collection in the scope of a single survey. Details about the ecological project Green Patrol can be found on the web site: http://www.zelena-patrola.com.

There are not many examples of research that combine delivery methods such as a web survey, SMS, and a paper survey. Some of this research will be described in the chapter that follows.

2 Review of research

Computer assisted data collection methods are increasingly replacing paper-and-pen methods of survey data collection. Most professional research organizations – academic, governmental, and commercial—now employ these new methods for the majority if not for all of their survey data collection. Computer assisted telephone interviewing (CATI) is most prevalent and computer assisted personal interviewing (CAPI) is rapidly gaining popularity. Also, new interesting forms of computerized data collection, for instance automatic speech recognition and surveys through the Internet, are emerging. Characteristic of all forms of computer assisted interviewing is that questions are read from the computer screen, and responses are entered directly in the computer, either by an interviewer or by a respondent. An interactive programme presents the questions in the proper order, which may be different for different (groups of) respondents.

If computers are used for data collection, there are several models that can be used. »Those models are known under abbreviations: CATI (Computer Assisted

Telephone Interviewing), CAPI (Computer Assisted Personal Interviewing) and CASI (Computer Assisted Self Interviewing)« (De Leeuw, E. & Nichols, W., 1996). Besides these new models for data collection, in our survey we used some more traditional methods, e.g. paper questionnaires and new and modern method interviewing via SMS (SMS survey). There aren't many examples of surveys that combined these three methods. Characteristics and comparison of CASI (a method used in our survey), paper survey and SMS survey are described as follows.

Computer assisted self–administered questionnaires are a relatively new invention. CASI differs clearly from both CAPI and CATI by employing a different interviewing situation. The computer has taken the role of the interviewer. Theoretically, this combines the advantages of traditional self-administered questionnaires, such as more openness with sensitive questions, with the possibility of using complex question structures.

»Respondents generally like CASI; they find it interesting, easy to use, and amusing« (Zandan & Frost, 1989; Witt & Bernstein, 1992).

The general positive appreciation of CASI is also shown in the relatively high response ratio with Disk By Mail (DBM) surveys. »DBM response ratios vary between 25% and 70%, and it is not unusual to have response ratios of 40 to 50 percent without using any reminders« (Saltzman, 1992). »Assuming that this is a special population interested in the research topic, an ordinary well conducted mail survey using no reminders may be expected to yield about 35% response« (Dillman, 1978).

The respondents will experience a higher degree of privacy and anonymity, which should lead to more self-disclosure and less social desirability bias. Strong support for this hypothesis is given by Weisband and Kiesler (1996). »In a meta-analysis of 39 studies they found a strong significant effect in favor of computer forms« (Weisband and Kiesler, 1996). This effect was stronger for comparisons between CASI and face-to-face interviews, but even when CASI was compared with self-administered paper-and-pen questionnaires, self–disclosure was significantly higher in the computer condition. The effect reported was larger when more sensitive information was asked.

A similar picture emerges in studies of electronic mail questionnaires. »Sproull and Kiesler (1991) report about five experiments on decision making in small groups.« Using an electronic network for communication leads to more open communication, more ideas and general participation in the discussion. In the face-to-face situation the discussion tended to be dominated by one or two high-status individuals. This may also be the result of differences in the social interaction. »However, in a direct comparison of a mail questionnaire and an electronic mail health-questionnaire Kiesler and Sproull (1986) also found fewer socially desirable answers in the electronic version.« They also investigated other aspects of data quality in this study. Both the item non-response and the number of errors were lower with CASI.

There are no systematic cost comparisons for CASI. The literature on disk-by-mail reports that DBM is generally more expensive than a comparable paper-and-pen mail survey. However, the gain in response in a single mailing is thought to be worth the extra costs.

The use of mobile phones for surveys can be seen as a modification of the CASI method, so in some articles this type of survey technique is defined as Mobile CASI or MCASI. MCASI has many different forms, some of which are: surveys via web pages, surveys via GPRS, or surveys via MMS and SMS messages.

In our survey we used this last method and the problems that emerged were: how to contact and motivate respondents to become survey participants; and secondly, what data collection method should be used? Besides these, what are the factors influencing response rates and the willingness or ability to complete a mobile survey, and what are the human and technical problems with regard to this new method? A target group for this type of survey is owners of mobile phones. Mobile phones have a large scale of penetration among customers, which means that there is a large group of respondents available through mobile phones.

One of the first surveys on this topic was made in 2003 by Tjostheim and Thalberg. »They conducted two parallel surveys under code name MCASI1 and MCASI2« (Tjostheim and Thalberg, 2003). MCASI1 was a simple survey containing mostly questions with single answers and only a few questions with multiple answers. The MCASI2 was more advanced and respondents received an MMS and had to look at multimedia entertainment content. The mobile survey contained more questions with multiple answers and open questions. Results of those surveys are as follows.

In MCASI1 only 15% of the participants experienced technical problems, and in MCASI2 almost 56% did. After the MCASI1 survey, 75% of participants were willing to participate in future surveys, and after MCASI2 49% were willing to participate in future surveys.

In a comparison of preferred survey methods, 65% of participants preferred a web-based survey instead of a mobile phone survey. The conclusion of this survey is that mobile phone surveys are good for short and simple surveys, and in more complex surveys the preferred method is a web-based survey.

3 The Green Patrol survey

The Green Patrol survey was conducted with the purpose of exploring the level of satisfaction with the environmental situation in the town of Varazdin. It was conducted in cooperation with Vecernji list newspaper, project E!2584 Ulixes (Faculty of Organization and Informatics) and the Town of Varazdin, and it represents a new and technologically advanced form of public survey.

The survey had two basic parts. First, participants were encouraged to detect and make reports about so-called »red« and »green« spots in the area. »Red« spots symbolized negative examples of environmental architecture, horticulture or environmental pollution; and »green« spots

symbolized positive examples in our environment. We collected the information about those spots by different methods: by paper questionnaires, web questionnaires, SMS and MMS messages, and by an interactive web map where we enabled our users to mark the exact location of »red« or »green« spots on a very detailed digital map of Varazdin. In this part of the survey, a total of 21 »red« spots were detected by participants, mainly by using the interactive web map. We explain that for this type of survey e-mail messages and interactive web maps are the most appropriate because they offer interactive communication with users and the possibility of uploading files (e.g. photos of »red« and »green« spots that are later connected with the interactive web map).

In the second part of the survey, the participants could, again by using different delivery methods, examine and express their satisfaction with the environmental surroundings of the town of Varazdin. The focus of this article will be on the second part of the survey. In the survey three different types of questionnaires were used: paper questionnaires, SMS questionnaires, and web questionnaires. All questionnaires had the same basic structure and they were all focused on the same problem area with the same questions. We used multiple types of questionnaire because we wanted to make this survey more attractive and approachable to a larger scope of respondents.

The survey was conducted over a period of four weeks, and during that time paper and SMS questionnaires were distributed and a web questionnaire was published on the Internet.

The used questionnaires had two main parts. The purpose of the first part was to collect general information about participants, particularly information about gender, age, level of education, marital status, employment, etc. For this purpose we used a set of eleven questions of the closed type with the possibility of adding additional information. The questions were derived from classical statistical surveys performed by the Croatian Central Bureau of Statistics (DZS).

This way we achieved the compatibility and comparability of our participant groups with other surveys performed by the Croatian Institute for Statistical Research.

The second part of the questionnaire differed between the »red and green points survey« and the »environmental surroundings satisfaction survey.«

The questionnaire used for »red« and »green« point detection was simpler than the other. It consisted of six open-ended questions through which the survey participants indicated which type of ecological point he/she spotted, the exact location of this point, the date and time when the points were noticed, and a short comment on the spotted point. These sets of questions were used in »red« and »green« point questionnaires regardless of the delivery method used.

The questionnaires used for the environmental surroundings satisfaction survey were a bit more complicated. The reason for this is the nature of the information that we were trying to get. There were 9 critical factors that were evaluated: satisfaction with the general state of environmental surroundings and environmental protection in the town; satisfaction with the care of local authorities for the environment (protection); satisfaction with surroundings from the horticultural and architectural points of view; waste management; quality of water and air; satisfaction with citizens' attitudes towards environmental

Figure 1: *The structure of the »Green Patrol« survey*

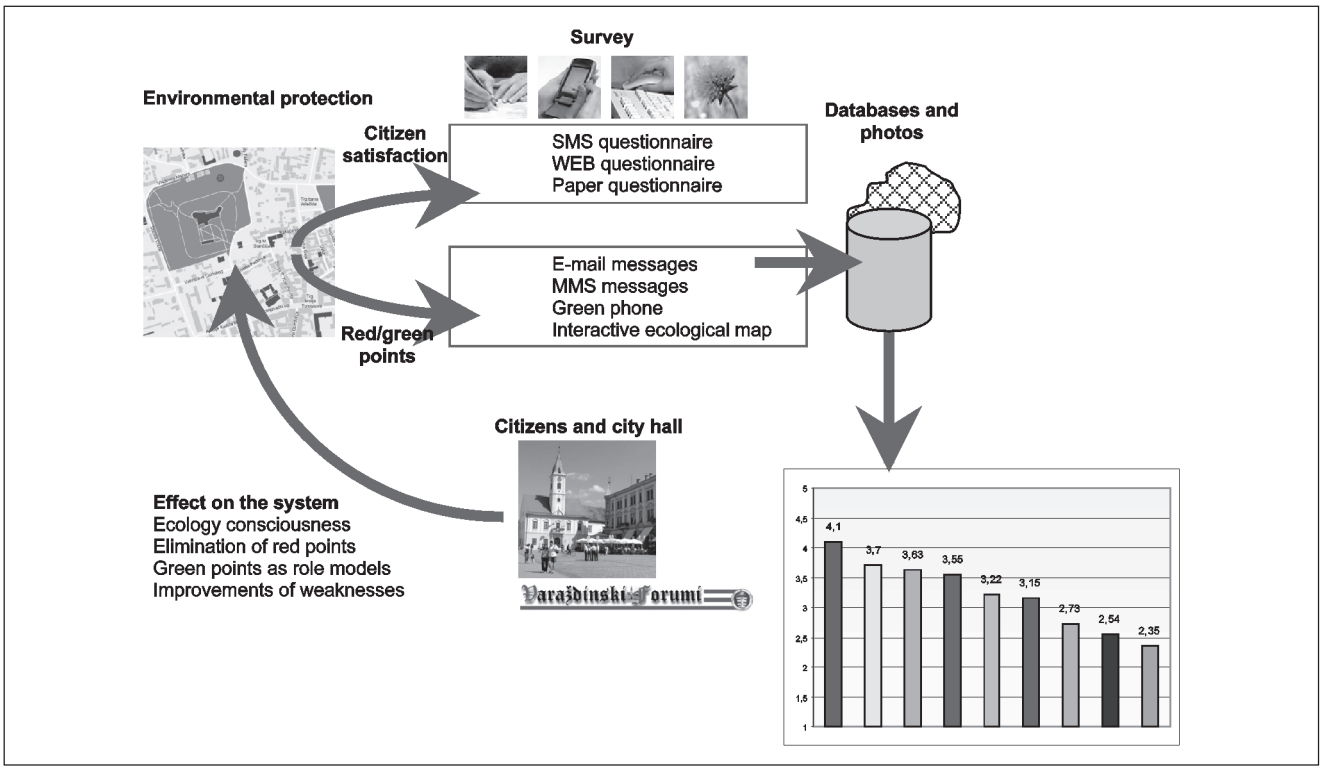
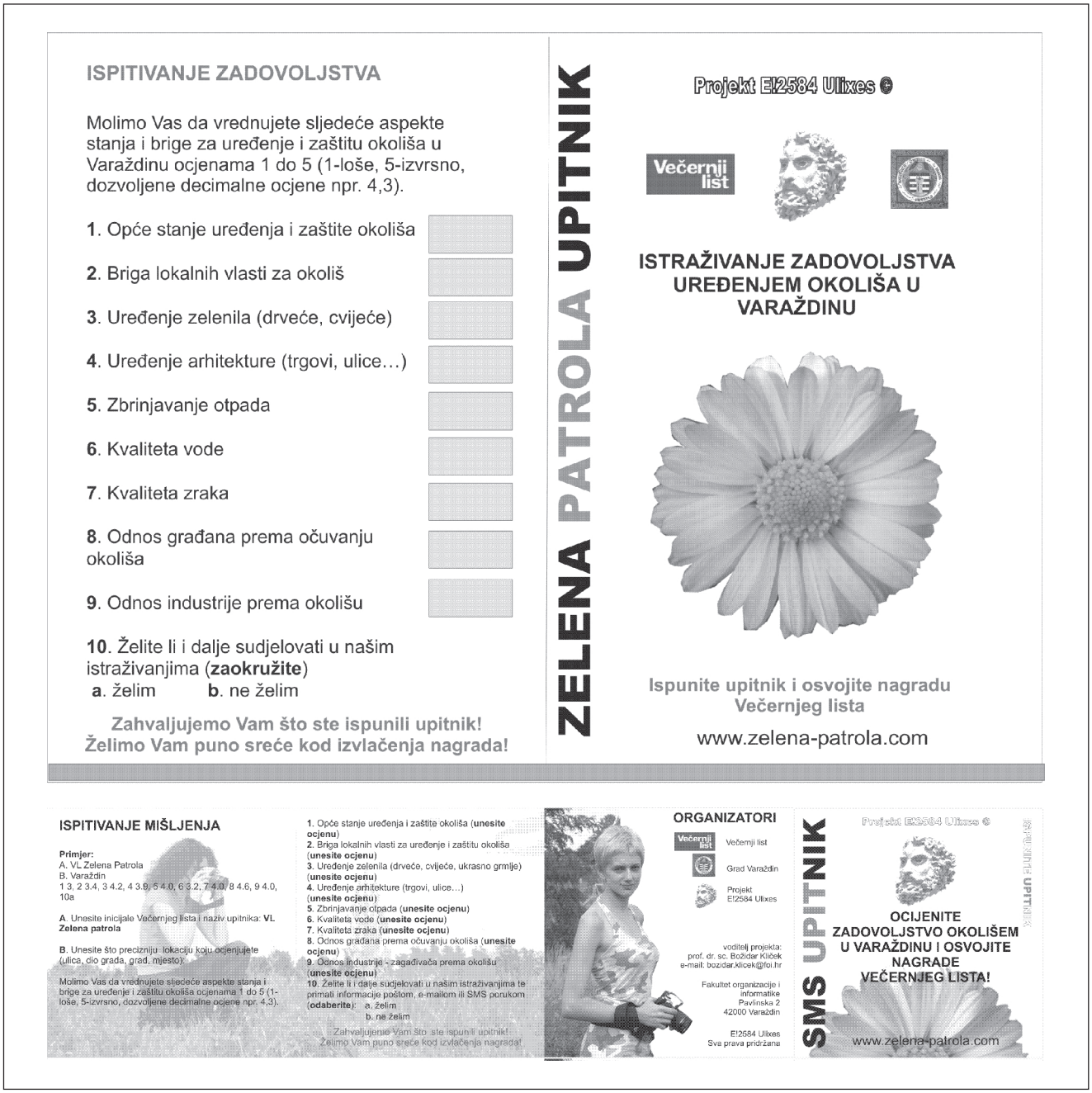


Figure 2: *Example of questionnaires used (paper and SMS questionnaires) (Kliček, 2004)*



protection; and satisfaction with industry's attitude towards environment protection. The participants expressed their satisfaction with marks from 1 to 5 (where 1 stands for »not satisfied at all« and 5 stands for »completely satisfied«). All questionnaires had the same questions regardless of the delivery methods used. These factors were selected as critical based on some international surveys from the field of environmental protection.

The randomness of the sample is not analyzed in this paper. The most important focus of the survey was to motivate as many citizens as possible to participate with their opinions.

With descriptive statistics our information was organized and summarized (using graphs, charts, tables and the calculation of various statistical measures). The complete analysis and results of the Green Patrol survey can be found at the following address: http://git.inet.hr/projects/tzgvz/zelenapatrola/?FlashID=4403&LanguageID=-1.

The total number of respondents in the part »environmental surroundings satisfaction survey« was about 600. Of these 600 participants, 470 used a paper questionnaire, 70 used a web-questionnaire, and only 37 of them used a SMS questionnaire.

We think that the reason for this lies in the ease of use and the adequateness of some survey methods. For example, this type of survey was most appropriate for paper questionnaires and less appropriate for SMS questionnaires (because of the characteristics of the methods and the types

of questions used in the questionnaire). A more detailed description of the benefits of these methods can be found in the next chapter.

The structure of the entire survey is shown in Figure 1.

The final results of this survey were presented in a press conference together with local authorities of the Town of Varazdin. Conclusions were that this is a good way of monitoring citizen satisfaction with environmental surroundings and for detecting negative points in our environment. It was concluded that the model of this survey can be used in the future for similar projects regarding environmental protection.

4 Reasons for and advantages of using multimodal delivery methods

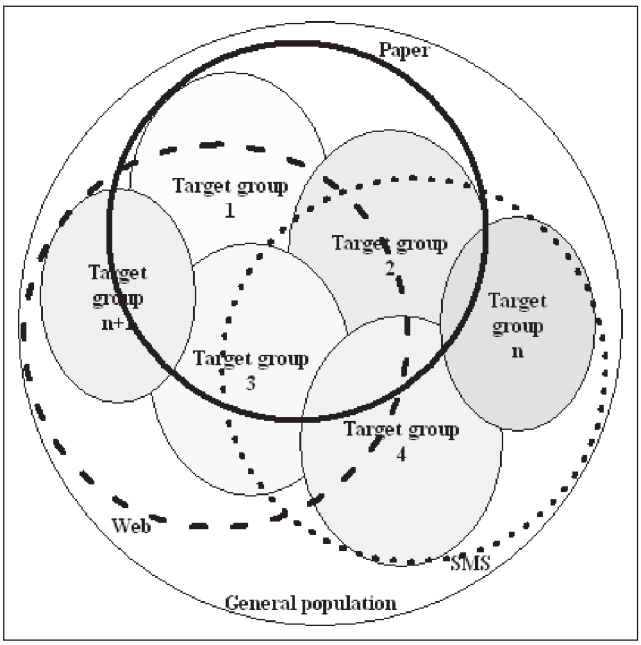
The »Green Patrol« survey of satisfaction with state and environmental protection was conducted in parallel with all three types of questionnaires described: web, SMS, and classic paper survey questionnaire.

The reason for using combined survey methods is their characteristics (*Table 1*). Different characteristics of methods enabled us to use different methods for respondents of different profiles.

The main reason for using these delivery methods was to create a large scale survey consisting of different groups of respondents (Young and Ross, 2000). The hypothesis was that younger participants would predominantly use mobile phones surveys (SMS), working people (middle age) would predominantly use web-based and paper surveys, and older participants would use paper surveys. No matter which method is used, all questionnaires were the same: the same structure and questions were defined under the same criteria and were focused on the same evaluated characteristics.

The use of multimodal methods for data collection helped us to create and use a larger sample of survey participants. The idea is shown in *Figure 3*. If the general population (GP) is represented as a set of different target groups (TG) (subsets of the general population), then it can be represented as

Figure 3: *The use of multimodal survey methodology - different target groups are preferred with different methods*



$GP = \{TG_1, TG_2, TG_3, TG_4, \ldots, TG_n, TG_{n+1}\}$. Different target groups prefer different methods of data collection. The specific data collection method (DCM) can be used for data collection in specific target groups, or $DCM_i \sim TG_n$. Therefore, if we use different data collection methods we are able to create a subset of survey participants that is more or less equal to the general population, and that is the desirable structure of survey participants in every scientific survey.

There are some theoretical and mathematical models which explain this, but it is not possible to build an exact model because of its incompatibility with practical solutions.

Age was a very significant determinant factor. Older age groups (45 - 70 and over 71 yrs) used mostly paper surveys and rejected SMS and web surveys much more than younger age groups. Teen-agers and people in their twenties (15 –

Figure 4: *Number of responses by age to different delivery methods*

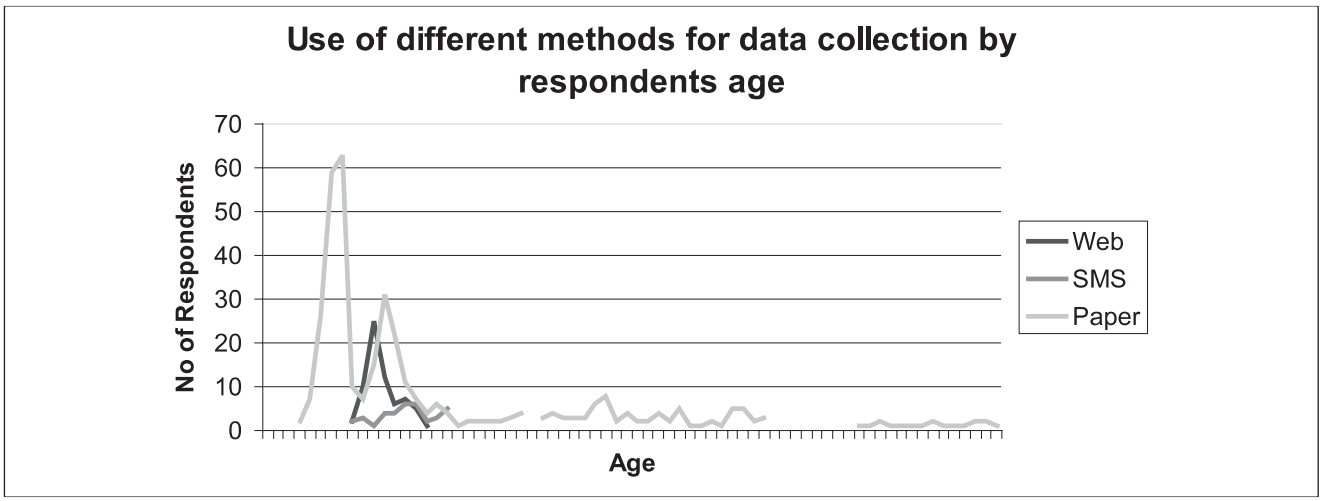
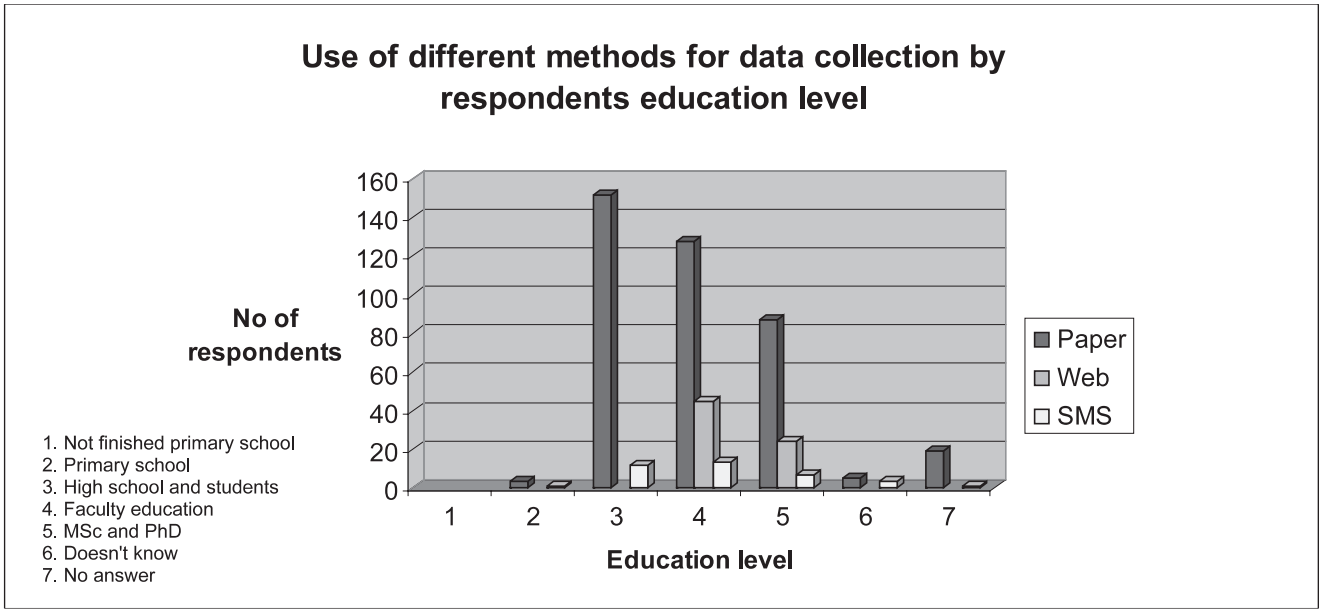


Figure 5: *Probability of response by education level to different delivery methods*



30) used mostly SMS surveys, and respondents in the age group 30 - 45 used mostly web surveys (*Figure 4*). The older population has not adopted to SMS and web usage as well as they have to paper surveys with classic paper survey questionnaires. The number of responses by age to different questionnaire methods is shown in *Figure 4*.

The survey has shown that participants of different ages prefer different survey techniques. The »modern« techniques, such as web and SMS questionnaires, are most popular among participants between 15 and 35 years of age. More traditional survey forms, such as paper questionnaires, are used mostly by people middle aged and older. The survey has shown that level of education of the respondents was also a very significant determinant factor. Students and respondents who had finished high school used mostly the paper and SMS method. They didn't use the web method. Respondents with a university education used mostly the paper method, but they also used the web method a great deal. The number of respondents that used the web method increases with level of education. The reasons are the necessary technical requirements and the basic computer skills of Internet users. There are still a large number of people who do not have access to the Internet (because of the cost of equipment and time online) or choose not to use the Internet because of the lack of familiarity with the Internet. Further reasons of this effect were not analyzed within the scope of this research. The number of responses by level of education to different delivery methods is shown in *Figure 5*.

By subjective estimation we define the basic characteristics of each survey method used in the »Green Patrol« survey: advantages and disadvantages, type of questions, target group and cost of the surveys. The characteristics of the different delivery methods are presented in *Table 1*.

5 Conclusion

The focus of this article was a survey on satisfaction with the present state of the environment and environmental protection, and the benefits of using multimodal delivery methods (web, SMS and classic paper survey questionnaires) for data collection.

The main reason for using multimodal delivery methods is to create a large scale survey that consists of different groups of respondents. Namely, younger participants used predominantly mobile phone surveys (SMS), working people (middle age) used predominantly web-based and paper surveys, and older participant used paper surveys.

We confirmed that the age of respondents is a very important factor. The older age groups (45 - 70 and over 71 yrs) used mostly paper surveys and rejected SMS and web surveys more than younger age groups and teen-agers. People in their twenties (15 – 30) used mostly SMS surveys, and respondents in the age group of 30 - 45 used mostly web surveys. The older population has not adopted SMS and web usage as well as paper surveys with classic paper survey questionnaires.

In addition, we established that level of education is also a very important factor. The number of respondents using the web method increases with level of education. It is likely that the reason for this is the necessary computer skills that Internet users possess. The SMS method was used mostly by students. It seems that the characteristics of this type of delivery method are the most appropriate for them.

The basic characteristics of each survey method—advantages and disadvantages, type of questions, costs of the surveys, and links between these characteristics and target groups—are also presented in the article.

The use of various methods for data collection helped us to create and use a large sample of survey participants. A

Table 1: *Characteristics of different delivery methods based on the »Green Patrol« survey*

Type of survey	Advantages	Disadvantages	Type of questions	The profile of respondents (target group)	Costs
Paper	➢ the profile of potential respondents is the biggest ➢ the sample consists of the general population ➢ the biggest percentage of returned questionnaires ➢ method with no technical constraints	➢ *a long time needed to distribute, collect and encode data* ➢ *the problem of collecting paper questionnaires* ➢ time to fill out (writing is slower than clicking) ➢ bigger possibility of data entry errors and irregular questionnaires ➢ problem of data entry in data base (costs and time)	➢ multiple choice and open- ended questions	➢ 45 - 70 and over 71 yrs ➢ the respondents must be able to read and write	➢ it is more expensive than web surveys (cost of printing paper questionnaires) ➢ costs for dissemination to respondents - postage
WEB	➢ faster speed of responses (increased response rates) ➢ familiarity and comfort of the respondents (flexibility) ➢ larger effects when more sensitive information was asked ➢ minimal data entry errors ➢ good for complex surveys ➢ no data entry costs (eliminates data entry from paper since all data is already in a database) ➢ web-based questionnaires can also be automatically validated ➢ easy to correct problems during survey administration since ➢ web-survey forms can be easily modified ➢ reduces time to fill out - clicking is faster than writing ➢ the progress report of the survey can be updated daily ➢ the forms can be designed so that only “legal” answers are accepted	➢ they typically do not reflect the general population (the lack of familiarity with the Internet and/or lack of access to the Internet) ➢ possibility of technical problems ➢ security (name and other identifying information) can be a big issue ➢ possibility of multiple copies of responses (mistakenly or purposefully) ➢ internet-based survey method needs to be used with caution (one click with mouse pointer can make a big change) ➢ no sample control (random respondents may reply if survey is on web)	➢ multiple choice and open- ended questions	➢ 30 - 45 yrs ➢ only Internet users with basic computer skills	➢ low costs - only software, no printing, envelopes or postage needed ➢ costs for the respondents – time on the Internet
SMS	➢ responses provided via SMS are more accurate than responses from web or paper surveys; there were surprisingly few empty or inadequate responses; ➢ SMS surveys allow the survey recipients to respond when it is convenient for them ➢ accessibility - SMS survey can be accessed from any mobile phone ➢ speed in gathering and handling the data	➢ not suitable for open ended questions ➢ SMS questionnaires were somewhat time-consuming to answer (around 5-15 min) ➢ possibility of technical problems ➢ SMS survey may be considered aggressive by respondents (detection of telephone number)	➢ multiple choice questions ➢ good for short and simple surveys	➢ 15 - 30 yrs ➢ owners of mobile phones	➢ it is more expensive than web surveys (the reason: printing SMS questionnaires) ➢ costs for the respondents - the price of SMS

combination of several methods enabled us to reach survey participants of different age groups and different levels of educational.

The Green Patrol survey was conducted with the purpose of exploring the level of satisfaction with the environmental situation in the town of Varazdin. The combination of on-line, SMS and paper questionnaires was used in the data collection process. It proved that this is useful to use when opinions of different profiles of people are needed. Nevertheless, the research also opened further questions (data quality, sampling frame, random sample, etc.) that should be addressed in further research.

Literature

1. Beckenbach, A. (1995). Computer Assisted Questioning: The New Survey Methods in the Perception of the Respondents. *BMS,* vol. 48, pp. 82 - 100.
2. Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method.* New York: Wiley.
3. De Leeuw, E. & Nichols, W. (1996). *Technological Innovations in Data Collection: Acceptance, Data Quality and Costs.* Sociological Research Online vol. 1, no. 4. http://www.socresonline.org.uk/socresonline/1/4/leeuw.html <12.03.2005.>
4. Griffis, S. E., Goldsby, T. J., Cooper, M. (2003). Web-based and mail surveys: a comparison of response, data, and cost, Journal of Business Logistics, 2003, http://www.findarticles.com/p/articles/mi_qa3705/is_200301/ai_n9201654/, <10.11.2005.>
5. Kiesler, S. & Sproull, L. S. (1986). Response Effects in Electronic Surveys. *Public Opinion Quarteryl* no. 50, pp. 402 - 413.
6. Kliček, B., Begičević, N. & Gerić, S. (2004). *Sažetak istraživanja zadovoljstva brigom za okoliš – »Zelena patrola«.* http://git.inet.hr/projects/tzgvz/zelenapatrola/ZP_Izvjestaj.pdf <02.11.2005.>
7. Medin, C., Roy, S. & Ann, T. (1999). *World Wide Web versus mail surveys: A comparison and report.* Paper presentation at ANZMAC99 Conference. Marketing in the Third Millennium Sydney Australia. http://www.anzmac99.unsw.edu.au/anzmacfiles/papers.htm <20.09.2005.>
8. Saltzman, A. (1992). Improving Response Rates in Disk-By-Mail Surveys. *Sawtooth Software Conference Proceedings*. Evanston: Sawtooth Software.
9. Solomon, D. J. (2001). Conducting web-based surveys. *Practical Assessment, Research & Evaluation* 7(19). http://PAREonline.net/getvn.asp?v=7&n=19 <20.10.2005.>
10. Sproull, L. & Kiesler, S. (1991). Computers, Networks and Work. *Scientific American* pp. 84 - 91.
11. Tjostheim, I. & Thalberg, S. *Are the mobile phone users ready for MCASI* http://www.program.forskningsradet.no/puls <15.04.2005.>
12. Weisband, S. & Kiesler, S. (1996). *Self- Disclosure on Computer Forms: Meta-Analysis and Implications.* Tucson: University of Arizona. http://www.al.arizona.edu/~weisband/chi/chi96.html <19.04.2005.>
13. Witt, K. J. & Bernstein, S. (1992). Best Practices in Disk-By-Mail Surveys. *Sawtooth Software Conference Proceedings*. Evanston: Sawtooth Software.
14. Young, S. J., Ross, C. M. (2000). Web Questionnaires: A glimpse of Survey Research in the Future, Parks & Recreation, June, 2000, http://www.findarticles.com/p/articles/mi_m1145/is_6_35/ai_63257084, <12.12.2005.>
15. Zandan, P. & Frost, L. (1989). Customer Satisfaction Research Using Disk-By-Mail. *Sawtooth Software Conference Proceedings*. Evanston: Sawtooth Software.
16. ***: Survey Solution Allows Wide Range Of Delivery Methods (2005). Online Product News, 2005., http://www.findarticles.com/p/articles/mi_m0BNP/is_10_24/ai_n15395986, <02.12.2005.>