ONLINE RESEARCH ABOUT ONLINE RESEARCH:
ADVANTAGES AND DISADVANTAGES

MARIJA TOPUZOVSKA LATKOVIJK
Institute for sociological, political and juridical research, University “Ss. Cyril and Methodius” in Skopje, Karposh, Skopje, 1000, Macedonia
E-mail address: marija_t@isppi.ukim.edu.mk
ORCID number: https://orcid.org/0000-0001-6069-5690

MIRJANA BOROTA POPOVSKA,
Institute for sociological, political and juridical research, University “Ss. Cyril and Methodius” in Skopje, Karposh, Skopje, 1000, Macedonia
E-mail address: mborota@isppi.ukim.edu.mk
ORCID number: https://orcid.org/0000-0002-7272-7554

Abstract

Aim. Many people believe that the Internet may revolutionise several disciplines by allowing easier and faster data collection, larger and geographically diverse samples, and therefore collecting more representative data. However, others are skeptical of its usability as well as its practical value. The paper’s aim is to highlight both positive and negative outcomes in the online research.

Methods. The paper is based on the e-methodological approach by using exclusively online data and online research methods: secondary data/articles with an open access about the topic in the literature review section; and online forum as an online research method for the data collection.

Results. As the most common advantages are recognised: quick access to information, wide geographical scope, a larger sample size leading to greater statistical power, reduced cost of conducting empirical research, support from online search engines and data bases, open data sources, no need for face-to-face interaction, no paper wasted. The most common disadvantages appointed by the respondents are: the need of evaluation for the accuracy of the online sources, the need for Internet access, limited access or high costs of some data bases, the contrast between fresh and redundant data available, irrelevant data, the need and knowledge for precise formulation of the key phrases/questions, users’ multiple accounts, AI involvement, low motivation of the targeted respondents, unreliable answers, etc.

Conclusions. It is interesting that the narratives confirm the already highlighted positive and negative outcomes in the online research process as they are already elaborated in the theoretical frame. This research proves that it is possible to use the online forums for gathering data, especially in the kinds of networks, where the most important attribute of the respondents is known solely by the membership itself. Because the internet is widespread, easily available, low costing, it is important to open a wide discussion about overcoming shortcomings on online research approach, and come to the useful and improved solutions.

Key words: online research methods, online data, online sampling, Internet.
Introduction

Online research involves the gathering of novel, original data (via the Internet) to be subjected to analysis in order to provide new evidence in relation to a particular research question (Hewson et al., 2003). Secondary Internet research involves techniques and procedures for locating and accessing bibliographic materials available online, such as journals, newspapers, official documents, laws and reports, databases, and so on. The use of the Internet as a medium for conducting both primary and secondary research is relatively new, but the topic demands increasing attention across a diverse range of academic disciplines, as more and more researchers, students and professionals are beginning to make use of the approach. The scope for quick, cheap access to a massive wealth of bibliographic information available online, as well as for gathering primary research data for social and behavioral research (e.g. using techniques such as web-based surveys, email interviews, analysis of online documents, and so on), is promising and exciting. However, the use of the Internet as a research medium remains relatively new territory, and the range of techniques, procedures, tools, and issues which emerge are still only just beginning to be developed and explored. Surveys, questionnaires and experiments are generally associated with quantitative approaches, interviews more so with qualitative approaches. A key principle is that online research studies, like any other research study, require careful planning, design and piloting.

Why would it be better to implement a research study online, rather than offline? The most obvious advantage of internet mediated research (IRM) procedures, and one which has attracted many researchers, is the cost- and time-efficiency often involved. Another special feature of the Internet is the ready access it provides to a potentially vast, geographically diverse participant pool. This may confer several advantages to an online research design, since it may facilitate cross-cultural research (Pohl, Bender, Lachmann, 2002), generate larger sample sizes than would otherwise be possible, thus leading to greater statistical power, and also help gain access to select, specialist populations, for example via online special interest or support group discussion forums (Coomber, 1997).

Another potential benefit of online approaches relates to the nature of the online interactional medium – in particular, that interactions can emerge which are fairly elaborate in terms of the richness of communication exchange, but where perceived (and actual) anonymity levels, and levels of perceived privacy, can be high. This feature is not something easily achieved in offline contexts. Hewson (1996) suggested that this may serve to reduce biases resulting from the perception of biosocial attributes (both of the participant and researcher), a point perhaps most relevant in quantitative research designs, where objectivity is often a prime goal.

One major issue, especially problematic for quantitative research approaches, is the reduced levels of control typically inherent in online research methods, compared with offline methods. This arises due to technical issues, such as different hardware and software configurations, and network traffic performance. Never-
theless, software and hardware failures may lead to unpredicted effects which may cause problems for the running of a study, and the lack of direct researcher presence has potential implications in terms of the extent to which researchers can gauge participants’ intentions, and levels of sincerity and honesty during a study, which may potentially hamper qualitative research goals (as in an online interview setting).

The lack of extra-linguistic cues available (e.g. tone of voice, facial expression, body language) could lead to ambiguities in the communication process (Bowker, Tuffin, 2004), for example, or difficulties in establishing good levels of rapport with participants (Strickland, Moloney, Dietrich, Myerburg, Cotsonis, & Johnson, 2003). The lack of extra-linguistic cues in online communication, compared with traditional face to face communication, however, is also associated with the heightened levels of anonymity and perceived privacy.

Document analysis in online research may be seen as essentially the same as non-real time observation, since it will also involve accessing online archives. Clearly, there is a wealth of online documentary data available, including web pages, scientific articles, news articles, data bases, reports, bibliographies, etc. The open access to large volumes of data, and the cost-effectiveness of obtaining this in a form ready for analysis, is a key benefit of document analysis techniques in online research.

**Online research methods**

Surveys have been identified as probably the most widely used technique within the social sciences. Surveys and questionnaires also appear to have been the most commonly implemented technique in web-based research (Hewson et al., 2003).

Online surveys have numerous strengths and potential weaknesses. Scholl (2002) states that when most of a society has internet access and savvy, the basic drawback for the use of online survey research, the lack of representativeness, disappears. The internet will then be an even more valued tool to obtain information from respondents living in different parts of a country or around the world, simply and at a low cost.

Online surveys are quite flexible. They can be conducted in several formats: e-mail with embedded survey; e-mail with a link to a survey URL; visit to a web site by an internet surfer who is then invited to participate in a survey; etc. In addition, they can easily be tailored to customer demographics, language, purchase experience, etc. by having multiple versions of a questionnaire. Each respondent sees only the pertinent questions. Online surveys can be administered in a time-efficient manner, minimising the period it takes to get a survey into the field and for data collection. Kannan (1998) concludes that the speed and global reach of the internet allow real-time access for interactions with geographically diverse respondent groups and information servers. Online surveys have come a long way from the simple, text-based, e-mail surveys of the 1980s
to the technologies available today. Respondents can click on a URL sent by e-mail and be transported to a feature-rich web survey tool that is directive and powerful, or reply directly to an e-mail survey by inputting answers as instructed (Dommeyer, Moriarty, 2000).

Online surveys provide convenience in several ways. Respondents can answer at a convenient time for themselves. They may take as much time as they need to answer individual questions. Some online surveys let respondents start and then return later to the question where they left off earlier. Instead of being annoyed at an inconvenient time with a telephone survey, a respondent can take an online survey whenever he or she feels it is convenient. It is relatively simple for respondents to complete online surveys and for their responses to be tabulated and analysed. For companies conducting online surveys, much of the administrative burden of sending and receiving questionnaires, as well as inputting data, is considerably reduced. In addition, once the last questionnaire for a study is submitted, the researcher instantaneously has all the data stored in a data base (Wilson, & Laskey, 2003).

Online surveys are capable of including dichotomous questions, multiple-choice questions, scales, questions in a multimedia format, both single-response and multiple-response questions, and even open-ended questions. Survey costs can be divided into two categories: preparation and administration. With regard to preparation costs, until recently, online surveys could be costly to construct because of the technological and programming requirements. Today, with the availability of advanced survey software and specialised online questionnaire development firms, preparation costs are much lower. In terms of survey administration, online surveys are automatically placed into the database, and then tabulated and analysed in a coordinated, integrated manner that greatly reduces costs. And because surveys are self-administered and do not require postage or interviewers, costs are also kept down.

Unlike with mail surveys, online surveys can require the respondent to answer questions in the order intended by the study designer, as well as prohibit the respondent from looking ahead to later questions. Online surveys can be constructed so that the respondent must answer a question before advancing to the next question or completing the survey, and so that instructions are followed properly (such as providing only one answer to a question). This eliminates item non-response and the necessity to throw out answers that have been entered improperly. Studies indicate that online surveys have a much higher item completion rate than mail surveys; and answers to open-ended questions tend to be longer with online surveys than with mail surveys (Ilieva et al., 2002). Online surveys can be constructed to ensure that respondents answer only the questions that pertain specifically to them, thus, tailoring the survey (by logic control). This eliminates respondent confusion, because complicated instructions are not needed. In addition, the perceived questionnaire length is reduced. Schonlau (2001) states that this means the software program, rather than the respondent, manages skip patterns. This reduces errors and, from the respondent’s perspective, makes the process simpler for taking a survey.
While email may feasibly be considered the most simple and accessible approach, in that most Internet users these days will be familiar with email and probably own and use an email account, and email can be easily used with relatively low levels of computer literacy, it can suffer the drawbacks of imposing lower levels of data security and researcher control. Contacting individuals directly by email may give the closest approximation of the sampling frame, and who had the opportunity to take part in a study, but has been more controversial in terms of whether this approach should be considered an invasion of privacy. In many research contexts, keeping participants responses confidential is important.

As widely reported, mail surveys generally have a number of major strengths, including the ability to use a large sample, the geographic coverage, the lack of interviewer bias, less respondent time pressure, the variety of questions that may be asked, possible respondent anonymity, and the low cost per respondent relative to personal surveys. Mail surveys also have these major potential weaknesses: the time needed for a company to receive all responses, the high non-response rate, unclear instructions, the tendency for some item non-responses – where answers are left blank, incomplete answers, brief answers to open-ended questions, an impersonal approach, and respondent ability to control the order in which questions are answered (Spitzer, & Alpar, 1989; Cavusgil, & Elvey-Kirk, 1998; Gendall, Menelaou, Brennan, 1996; Jussaume, Yamada, 1990).

Archer (2003) does a good job of summing up many of the advantages and potential disadvantages of online surveys versus mail surveys. Online advantages include: the virtual elimination of paper, postage, mail out, and data entry costs; reduced implementation time; reduced surveying costs once an electronic data collection system is developed is in place; data display and tabulation simultaneous with completion of surveys; availability of data in graphic and numerical format; easy to send reminders to non-respondents; and simplicity of importing data into data analysis programs. Potential online disadvantages include: everyone not connected, so online surveys will not work with some populations; limited computer literacy among some possible respondents; different screen configurations; difficulty of sampling e-mail addresses; and the quicker, sometimes instantaneous, decision not to respond.

The meaning of the word panel in online panel is different from the traditional meaning of that word in the survey research world (Göritz, Reinhold, & Batinic, 2002). According the traditional definition “panel surveys measure the same variables with identical individuals at several points in time” (Hansen, 2008, p. 330). The main goal of a panel in this usage is to study change over time in what would be called a “longitudinal panel.” In contrast, an online panel is a form of access panel, defined in the international standard, ISO 20252, as a “sample database of potential respondents who declare that they will cooperate for future data collection if selected” (International Organisation for Standardisation, 2019). These panels sometimes include a very large number of people (often one million or more) who are sampled on numerous occasions and asked to complete a questionnaire or answer certain questions.
Originally, these panels were called discontinuous access panels (Wansink, & Sudman, 2002, p.303). Göritz (2002) defined an online panel as a “pool of registered persons who have agreed to take part in online studies on a regular basis” (Göritz, Reinhold, & Batinic, 2002, p. 27). The attraction of online panels is threefold: (1) fast data collection; (2) promised lower cost per interview than most other methods; and (3) sampling efficiency due to extensive profiling. The recruitment methods for non-probability panels are numerous and varied, but with virtually all of them, anyone who is aware of an open invitation to join can volunteer to become a panel member. That is, people select themselves into the panel, rather than the researcher selecting specific individuals from a sampling frame that contains all members of a target population.

An internet forum or message board is an online discussion site. Internet forums have a treelike structure: usually, different topics are discussed within different thematic sections and sometimes sub-sections. Within the sections or sub-sections, users can start a discussion – a so-called thread – with a starter posting. Other users can reply to the starter posting or to other users’ comments. These messages are called a post or a posting. In many forums, threads and postings can be read by every internet user, but in order to achieve the right to post or to start a thread, users will have to register and log in. There are other forums where postings can only be read by registered users. Besides commercially oriented discussion boards, there are also many forums run by religious, political, or other societal interest groups. Members of these groups may know each other in real life, and the groups may also be linked to an nongovernmental organisation (NGO), church, or party. Alternatively, they may just be virtual communities (Rheingold, 1993), without the members knowing each other in person. Usually, these forums are used almost exclusively by members and supporters of the organisation or community for discussing matters of concern of the respective interest group. Hence, such forums allow for an analysis of typical discourses taking place within such communities.

An obvious advantage of internet forums is the almost unlimited amount of material for analysis. Some forums, with thousands of users, feature millions of postings in hundreds of thousands of threads. Even small forums usually contain more than enough text material for any kind of social scientific analysis. Because the material exists already in digital format, labour-intensive procedures like the transcription of audio material are not necessary. In a sense, forums constitute a kind of not moderated virtual focus group (Moloney, Dietrich, Strickland, Myerburg, 2003), in which members of a community discuss topics without a researcher interfering and possibly influencing the expression of thoughts. Hence, material from internet forums can be considered as relatively authentic natural data.

When considering using material from internet forums as a data source, the researcher will most likely be interested in discourses within certain (online) communities. Data from internet forums can be used for almost all kinds of qualitative analysis. Using postings as units of analysis in content analytic approaches is most adequate.
When posting to discussion groups, or on web pages, certain procedures are to be recommended. Selection of which discussion groups to post requests to will depend on the research question, and goals. Some authors have reported targeting particular discussion groups in order to obtain samples with certain characteristics (Birnbaum, 2001). Asynchronous approaches allow the participant to take part at a time convenient to themselves, and to potentially engage in greater levels of reflectivity, reflexivity, and also to consult external documents or sources. However, a possible disadvantage of asynchronous approaches is that they may suffer from a breakdown in conversational “flow”, due to the lack of continuity in the discussion over an extended timescale, and some authors have reported experiencing this effect (Bowker, Tuffin, 2004; Murray, & Sixsmith, 1998).

**Online sampling**

Using the Internet to obtain representative samples has been considered particularly problematic, especially due to the lack of a central register of all Internet users. Another alternative is to sample offline and ask participants to access and complete a study online, by online survey tools available on the Internet, or to develop a tailored online application and conduct a CAPI (computer assisted personal interviewing). Considering interviews, these can be conducted online asynchronously using email (including mailing lists, for larger focus group interviews) and asynchronous discussion forums.

Samples for online surveys can be generated in several different ways. Through their own databases, companies can develop opt-in mailing lists of their customers. Companies can also work with research firms and gain access to demographically balanced panels. If not properly addressed, online surveys also have these potential weaknesses. Perception as junk mail/spam is a big problem. It is important if the e-mail comes from a trusted source. Although the internet population is becoming more representative, there may still be survey difficulties due to the lack of familiarity of possible respondents with internet protocols. Online surveys are affected by both the type of internet connection and the configuration of the user’s computer. Because online surveys are self-administered, answering instructions must be extremely clear. If not, as Ray and Tabor (2003) report, some people may be frustrated and exit a survey without finishing the entire questionnaire. As with mail surveys there is usually no human contact in online surveys. This can limit the ability to probe in-depth as a skilled interviewer could do (Scholl et al., 2002).

The recruitment methods for non-probability panels are numerous and varied, but with virtually all of them, anyone who is aware of an open invitation to join can volunteer to become a panel member. That is, people select themselves into the panel, rather than the researcher selecting specific individuals from a sampling frame that contains all members of a target population.

Whereas the relative anonymity of the internet is an advantage as it reduces social constraints, it also complicates analysis in so far as there is normally only little socio-demographic information available about the users, especially in
online forums. Although on social networks it is possible to get information of various types: a) personal data such as name, surname, gender, date of birth, city, residence, etc.; b) contacts such as phone number, mobile phone number, e-mail address, personal website, instant messaging; c) educational path: academic path, qualification, diploma, achieved specialisations; d) further info: political orientation, general interests, aggregations, groups (Di Sia, 2018), most often, in online forums users participate under a fictitious nickname, even though the age and sex is often shown as part of their profile, there seems to be no way to verify this information. Another issue can be the users’ tendency to make more extreme statements on the internet (Williams et al., 2002, 2005) than they would in face-to-face situations, due to de-individuation effects (Lea, & Spears, 1991; Reicher, Spears, & Postmes, 1995). It must be acknowledged that in most cases it will not be possible to make claims regarding representativeness for a certain population. Not every member of a given social group may have access to the internet and only a few of those who have will engage in discussions within such forums. It is important to define and select relevant forums.

**Methodology**

The paper itself is based on the e-methodological approach by using exclusively online data and research methods: secondary data/existing articles with an open access about the topic in the literature review section; and online forum as an online research method for the data collection. The research target population is the ResearchGate community (members), a professional network (and interactive online platform) for scientists and researchers from all over the world: “ResearchGate is the professional network for scientists and researchers. Over 15 million members from all over the world use it to share, discover, and discuss research. We’re guided by our mission to connect the world of science and make research open to all (p. https://www.researchgate.net/about).

The online discussion was initiated by the authors, from the personal RG profile for the purpose of this paper, with the following question as a starter posting: “Online research: advantages and disadvantages? What is your opinion/experience on this topic? Online research for this particular discussion is defined as a usage of online research methods for collecting primary data, and as well research sample design by usage of available online data or online tools”.

Total number of N=14 respondents have joined the discussion (started one month ago) by sharing personal opinions/answers, while the discussion counts 134 reads.

The data analysis follows the qualitative research strategy, and the data/posts are narratively analysed and illustrated in the text by (anonymous) quotes/anecdotes.
There are some research limitations and implications which require special attention: 1) there is no information available about the respondents’ demographic characteristics; 2) research findings cannot be generalised at the whole target population level as there is no representativeness of the sample; 3) user profiles with a smaller group of followers have lower potential to attract high response rate in the open discussion; 4) the proactive behaviour in the open discussion depends from the field of expertise, knowledge, motivation and personal interest of the potential respondents and the nature of the question asked/starter post.

Results and discussion

The most popular and recommended answers are analysed and presented below. It is interesting that the narratives confirm the already highlighted positive and negative outcomes in the online research process as they are elaborated in the theoretical frame.

The most common advantages are recognised as the following: quick access to information, wide geographical scope, a larger sample size leading to greater statistical power, reduced cost of conducting empirical research, support from online search engines and data bases, open data sources, no need for face-to-face interaction, no paper wasted.

The most common disadvantages appointed by the respondents are: the need of evaluation for the accuracy/trustworthiness of the online sources, the need for Internet access, limited access of some data bases or high cost data access, the contrast between very fresh and very redundant data available, irrelevant data, the need and knowledge for precise formulation of the key phrases or key questions, respondents as users with multiple accounts, AI involvement, low motivation of the targeted respondents, unreliable answers, exclusion of the individuals with limited or no computer access as a part of the target population.

We can agree for sure that the listed advantages add a high value for the research process, but it is concerning that the respondents were more focused on the disadvantages of the online research in their statements. This gives an implication that the trust/confidence in the online research process is still questionable (and the validity of the data), which represents a significant barrier/resistance, and therefore from the academic members can be expected to avoid to rely on the usage of Internet in their research projects, whenever it is possible.

But if the research budget is sufficient, and the time framework is tight, there are big chances that the online empirical research will be the first choice of the researchers, in order money and time to be saved in the field research, while getting work done.

Examples:

“Advantages: quick access, access to data from distant countries, enhanced support from online search engines and data bases, can process large data sets, Disadvantages: access skewed to more recent data, need to screen/evaluate online sources for accuracy/authenticity, must have
internet access, precludes conducting fieldwork, cost of data access can be high, many data bases have limited access”.

“Advantages: finding several and different unprocessed and raw data through different sources in the the event they are open data sources; possibility of finding the right person and/or team during your research that there would be the chance to ask directly from the very right one no matter to have a necessarily face-to-face interaction. Disadvantages: facing redundant and/or partially redundant data; coming from different sources that may even seem different at their first appearance. Challenges: how to express what exactly you are looking for by applying the right words or phrases, through searching engines or links to get the satisfactory answer; how to find and track the information you need through tons of irrelevant data where only a few parts could be probably beneficial for your purpose.”

“Disadvantages depend on the sorts of questions asked and where one looks for answers. Here are a few more disadvantages: Population is not clear, users with multiple accounts may provide multiple answers, accounts might not all be human, unequal use patterns”.

“As we all know, people are busy with their social media apps so online research has to face following problems: the engaged population is low, the answers are not reliable as people hurry to finish the survey without truly understanding the questions – they just put tick mark in answer boxes, and finally, no one really cares about your research”.

“I think that the online data analysis depends on the trustworthiness of the source where you search”.

“As the person interviewed I would prefer online CAPI survey (than paper survey). It wastes less paper and I can often do it in less time than a paper survey. If in person I do not have to deal with the person conducting the interview and if by mail I do not have to mail it back. As the scientist, the answer is less clear. In many ways the paper survey provides more information. I know who the survey was sent to, and I can try and get a reluctant individual to take the survey even if their first inclination is to discard the survey. I also don’t have to worry about individuals with limited or no computer access. My results could be weaker if I have to include “computer use” as part of the definition in describing my target population. That said, it is not all that different from sitting at the corner of a street and asking people to take a survey, or mailing it to them. You just need to match the survey method to the target audience and be aware that the method may exclude some subpopulations”.

“The ability to obtain a larger sample, which increases statistical power”.

“Advantage: easy access to a lot of information. Disadvantage: not all information can be verified”.

“Among the advantages of conducting research on the Internet: reduced cost of conducting empirical research”.
This research also proves that it is possible to use the online forums for gathering data and analysis, especially in those kinds of online networks/platforms where the most important attribute(s) of the respondents (as the professional background) is known solely by the membership itself. But the research limitations described in the methodology section must be considered as very important.

**Conclusions**

Research results obtained in this paper by using “online forum techniques” targeting respondents from ResearchGate community, confirmed already highlighted positive and negative outcomes in the online research process as they are elaborated in the theoretical frame. They can be briefly summarized:

**Advantages:** elimination of paper and data entry costs; reduced implementation time; (geographically) larger samples; reduced surveying costs; easy data display and tabulation; availability of data in graphic and numerical format; easy to send reminders to non-respondents; and simplicity of importing data into data analysis programs.

**Disadvantages:** everyone is not connected on the net; limited computer literacy among some possible respondents; different software configurations; difficulty of sampling; low motivation of potential participants; and easier decision of potential respondent not to respond.

For sure, there are more (or more significant) advantages than disadvantages in using online research. Beside the larger number of advantages, there is still one essential and very serious limitation that lies in correct sampling and accuracy/authenticity of respondents. This is the new task for researchers to find solution in overcoming this barrier. Future studies have to address two main questions:

“Is it possible to find a way to develop precise and inclusive database of potential respondents who will cooperate for future data collection?” Then, this database can be used for developing accurate sampling.

“Can we use some other tools like “finding leads tools”, already used in marketing?” These tools are used for “data mining” and “data scraping”, and by using them they easily develop a list of customer details including company names, employees, social media links, even telephone numbers. However, these tools have to be adapted and improved for online research sampling.

Even if the online internet research process is relatively new and has its own disadvantages it cannot be neglected and rejected for future research endeavours. Because the internet is widespread, easily available, low cost, it is important to open a wide discussion about overcoming shortcomings with online research approach, and come to useful and improved solutions. For example, a recent study (Brosnan, Kemperman, & Dolnicar, 2019) has confirmed ten drivers for targeted respondents’ motivation: incentive payments, speed of completion, ease of completion, topic interest, software functionality, benefit to others, topic know-
ledge, impact, relationship with brand/organisation, and respondent’s opinions being valued (these drivers are not independent; they can reinforce or compensate for one another).

References


Mann, C., & Stewart, F. (2000). *Internet communication and qualitative research: A handbook for researching online*. Sage


