

# Studying Migrants with the Help of the Internet: Methods from Psychology

Ulf-Dietrich Reips and Laura E. Buffardi

*The article describes how the Internet can be used for research both with migrants and on the psychological impressions others have of migrants. Internet-based research methods are currently one of the hot fields in methodology. Though Internet-based research has only a brief history—around 15 years—the field has seen a massive increase in the number of studies conducted on the Internet, and the collection of tools with which to conduct Internet-based research. This development demonstrates a grass-roots change in how research in the behavioural and social sciences is often conducted. In our paper, we use the example of examining social networking Web pages (in terms both of its quantitative and its qualitative aspects) for markers of biculturalism in migrants to illustrate how this type of investigation can be a fruitful avenue for researchers in the area.*

**Keywords:** *Internet-Based Research; Methodology; Migration; Social Networking; Social Media*

## Introduction

Many features of the Internet make it a desirable venue for conducting behavioural research with migrant samples. First, earlier work has shown that much Internet-based research produces results similar to those of studies conducted in the laboratory (Dandurand *et al.* 2008; Krantz and Dalal 2000; for exceptions see e.g. Birnbaum 2001; Buchanan *et al.* 2005a, 2005b). Second, in recent years, methodologists have created an array of free tools that are available to help researchers conduct many types of study on the Internet (e.g. Freelon 2010; Reips and Funke 2008; Reips and Garaizar 2011; Reips and Stieger 2004). Third, the Internet allows researchers to

---

Ulf-Dietrich Reips and Laura E. Buffardi are respectively Ikerbasque Research Professor and Post-Doctoral Researcher at the University of Deusto. Correspondence to: Prof. U.-D. Reips/Dr L.E. Buffardi, Faculty of Engineering, iScience Group, University of Deusto, Bilbao, Spain. E-mails: reips@deusto.es; lbuffardi@gmail.com.

readily access samples with specific qualities of interest (e.g. migrants from a certain culture currently living in another culture of interest; see the articles by Conversi, Kissau and Oiarzabal, this issue).

Fourth, and finally, the rise in the popularity of social media can provide migration researchers with a unique insight into migrants' thoughts and behaviours that are occurring naturally in their social networks. To illustrate this point take, for example, the construct 'biculturalism'. For centuries, migration has contributed to the development of bicultural identities. More recently, improved communication and transportation technologies in conjunction with increasing globalisation also foster bi- (or multi-)cultural identities (e.g. Arnett 2002; Phinney 1996). Thus, researchers have become increasingly interested in the ways in which individuals negotiate their cultural identities and how this affects psychological processing (see Berry 1990; Harker 2001; Lay and Safdar 2003). Examining social media activity may provide a window into both how biculturalism is manifested in the communication and self-presentation of migrants and how the integration of two (or more) cultural identities affects their psychological well-being. Similarly, using social media methods, researchers can more easily study social responses and attitudes towards migration and migrants that would complement traditional approaches in migration research (e.g. Pardos-Prado 2011). Before further discussing how the Internet can be used to study migrants and attitudes towards migrants, we first review some information about the history of the Web, social media, and Internet-based social and behavioural research.

The history of social media can be seen as the time before and after the introduction of the World Wide Web (WWW). After Tim Berners-Lee developed and added the graphical, interlinking layer to the Internet at CERN in 1991, and particularly when forms—and thus interactivity—were introduced in HTML Standard 2.0 in 1995 (Musch and Reips 2000), the ground was prepared for modern social media. Tim Berners-Lee (1998) describes this development as the second part of the dream behind the WWW, calling it 'intuitive hypertext creation tools', or what later became the 'Web 2.0' or 'modern social media':

The dream behind the Web is of a common information space in which we communicate by sharing information. Its universality is essential: the fact that a hypertext link can point to anything, be it personal, local or global, be it draft or highly polished. There was a second part of the dream, too, dependent on the Web being so generally used that it became a realistic mirror (or in fact the primary embodiment) of the ways in which we work and play and socialize. That was that once the state of our interactions was on line, we could then use computers to help us analyse it, make sense of what we are doing, where we individually fit in, and how we can better work together. (...) With the dramatic flood of rich material of all kinds onto the Web in the 1990s, the first part of the dream is largely realized.

When the interactive WWW became available, a few pioneers of Internet-based research discovered that this technology can be used very well in social and behavioural research. It is even clear that some types of scientific inquiry can *only* be

conducted on the Internet. There is a ‘...*fundamental asymmetry of accessibility*: What is programmed to be accessible from any Internet-connected place in the world, will surely also be accessible in a university laboratory, but what is programmed to work locally may most likely not be accessible anywhere else. A laboratory experiment, for instance, cannot simply be turned into a Web experiment by connecting the host computer to the Internet. But any Web experiment can also be used in the laboratory’ (Reips 2006: 76). For social media, with their complex and dynamic mix of technology and social interaction which depends on high accessibility, wide reach and high usability, the situation is even more asymmetrical.

The first Web-based research questionnaires appeared on the Internet in 1994. The first Internet-based experiments were conducted in 1995, by Krantz *et al.* (1997) and Reips (1997). The number of studies conducted via the World Wide Web has grown ever since—Reips and Krantz (2010) report more than 2,000 archived Internet-based research studies in their archives alone.<sup>1</sup>

A portal for researchers who are considering conducting research on the Internet is the *iScience Server* at <http://www.iscience.eu/>. It contains tools that can be used to create Internet-based studies or response items for Web questionnaires, to include an Internet-based Big Five personality test with automatic scoring in one’s own study, to recruit participants, to analyse log files and to learn about Internet-based research.

With the technological realisation of Tim Berners-Lee’s dream of a rich, worldwide communication network approaching universality,<sup>2</sup> modern social media provide many new opportunities for individuals, particularly for migrants. Maintaining connections in a fuller and more flexible way via the Internet is likely to provide psychological and financial relief for migrants, offer organisational advantages, and many other practical achievements over previous situations. Maintaining connections via information and communication technologies (ICTs) like the Web is particularly prevalent and often emotionally essential for migrants to maintain ties with their families and friends (see Nedelcu, this issue). Indeed, with new media, the transnationalisation of the immigrant family experience seems to become the norm (Bacigalupe and Cámara, this issue). Because the human desire for connections with family members can be achieved better in the age of the Internet than before, it seems likely that migration can be more successful psychologically despite the limitations of spatial distance and thus provide fruitful ground for true global citizenship. Nedelcu writes: ‘ICTs allow migrants to form multiple belongings, to capture cosmopolitan values, to develop deterritorialised identities and biographies and to act at a distance in real time; on the other hand, while accelerating integration and incorporation paths in host societies, ICTs also enable migrants to defend particularistic values and to claim a particular belonging while living as global citizens’.

In this article, we present our findings on methods, techniques and tools in Internet-based research and tailor this discussion to show how they can be useful to those who conduct studies with migrants.

## Internet-Based Research Methods: Types and Issues

In this section we provide an overview of the basic methods often used in Internet-based research and argue that the often semi-legal or illegal status of migrants favours Internet-based research methods, particularly with social media, over other methods. Boundaries between methods are sometimes blurry, and many successful research projects combine different methods to profit from methodological triangulation. Thus, provided researchers bear in mind that it is possible to combine different methods, as we show later in the paper, the following categorisation hopefully provides a rich overview of the many options available to researchers who would like to use the Internet in data collection.

### *Internet-Based Interviews and Surveys*

Until recently, most Internet-based interviews were conducted using text-based technology (Hewson 2007). Interviewer and interviewee(s) would type via an interface, most often within a browser window. Naturally, the texts from such interviews could easily be subjected to content analysis, with software to support the analysis itself (Amsterdam Content Analysis Toolkit [AmCAT] Software: <http://content-analysis.org/>) or the calculation of intercoder reliability (Freelon 2010). Interviews can be assisted by dialogue bots (Janetzko 2002) or even be automatically performed by scripts. Stieger and Reips (2008) developed the Dynamic Interviewing Program (DIP), software that conducts automatic online interviews via the instant messenger ICQ. With it, Paine *et al.* (2007) were able to automatically select and contact more than 70,000 ICQ users and also automatically interview all those who were willing to be interviewed.

Hewson (2007) includes enhanced candour and self-disclosure as chief advantages of online interviews, which also offer more-balanced levels of power between the interviewer and interviewee(s) and better participation options for those who may otherwise not be able to participate in interviews (e.g. new mothers or family members of migrants living in remote areas). Hewson (2007) also provides a set of useful distinctions and guidelines for those using the Internet to conduct interviews. Among them is the important distinction between *synchronous* (presence of interviewer and interviewee at same time) and *asynchronous* approaches. In particular, asynchronous interviews, e.g. via e-mail, allow for depth and reflexivity.

The most frequently used Internet-based research method is the Web survey or Web questionnaire (in a way, a Web questionnaire resembles a very highly structured text-based interview). The frequent use of questionnaires and polls on the Internet may come from the ease with which Web questionnaires can be constructed and conducted. However, true Web survey methodology is a complicated matter—it aims to generalise results from a sample to a particular population. Work by Dillman *et al.* (2010), among others, has shown that many Web surveys are plagued by problems of usability, display, coverage, sampling, non-response or technology. Data quality can be influenced by many factors—for example, the degree of anonymity and information about incentives (Frick *et al.* 2001; Göritz 2006), the degree of

personalisation and the power attributable to the sender of an invitation to participate in the survey (Joinson and Reips 2007) and whether a 'one screen, one question' procedure is applied or not (Reips 2002, 2010). Aspects of a Web study that may annoy participants, such as forcing a response, lead to an increase in random answering behaviour and non-response (Stieger *et al.* 2007). Despite these findings, converging evidence shows that Web-based survey methods result in qualitatively comparable results to traditional surveys (Dillman *et al.* 2010). Web survey researchers have begun to explore options of mixed-mode surveys (De Leeuw 2005; Shih and Fan 2007). A good resource is the Web survey methodology website at <http://websm.org>.

A question that arises often in reactive Internet-based research methods such as interviews and surveys is the question of verifiable identities. For much Internet-based research with voluntary interviews, surveys, tests and experiments, it has been shown that participants rarely provide incorrect demographic information (<3%, according to Voracek *et al.* 2001). In research on sensitive issues, the researcher can employ verification techniques like randomised response and unmatched count (Coutts and Jann 2011; Musch *et al.* 2001) or instead choose to work with non-reactive research methodology.

### *Non-Reactive Methods*

Non-reactive data collection refers to the use and analysis of existing traces, databases and collections of materials. Unobtrusive observation and data mining fall into this category. 'Non-reactive' is the appropriate term because it is (almost) impossible for the data to be in any way biased from reactions to knowledge that these data will be used in research. Most often, at the time when the data are generated, it is not even planned that they will later be used for research purposes. Examples for the possible application of non-reactive data collection methods on the Internet are social media profiles, personal homepages, forum contributions, server log files or picture repositories. The vast size of such stores of data on the Internet multiplies the strength of non-reactive methods: behaviours like navigating, chatting, reading, searching, selecting, etc. can be studied as they occur naturally when using the Internet. This facilitates the examination of rare behavioural patterns and their timing. Many of these behaviour patterns are stored in server databases or log files. Thus, log-file analysis is an important example of a non-reactive Internet-based method (Reips and Stieger 2004).

By the definition of non-reactive data collection, it is necessary that those who produce the data do not know or suspect that their data may be used for research purposes. Also, it assumes that the researcher is not interfering in any substantial way with the setting (e.g. joining a small group and talking in unusual ways would change what the group would have done if the researcher had not joined the group). In social media this problem is somewhat reduced due to the limited physical presence and the large numbers of users. However, care must be taken to avoid accidentally switching from non-reactive to reactive data collection, e.g. if the researcher joins groups and befriends or otherwise gains access to conversations and data in Facebook.

Search engines are among the most frequently used services on the Internet. With billions of searches performed every day, it is natural that these searches contain much information about many aspects of human life. Simple estimates for relative importance are *search engine count estimates* (SECEs; e.g. 'googling X returns 1.4 million links'). Janetzko (2008) has shown good objectivity and validity for SECEs. Some search-engine providers have recently moved beyond frequencies and subjected their data to higher order algorithms for mining, as the following examples show. Google Trends (<http://www.google.com/trends>) and Google Insights (<http://www.google.com/insights/search/>) have been used to successfully predict local epidemiology in influenza (Ginsberg *et al.* 2009) and local unemployment rates (Askatas and Zimmermann 2009). Migration researchers and professionals working with migrants may benefit directly from these new intelligent data-mining tools—for example such tools may enable them to better prepare for predicted important events in a particular area. A very simple search, for 'migration', without setting any filters, suggests that media attention focused on migration may actually be hype and not a reflection of true increases in public interest in migration. The frequency with which the term appears in the media is negatively correlated with the frequency with which the term is being searched for by the general public (see Figure 1). Figure 1 further reveals that most searches are performed in countries that have recently been affected by new migration laws (the UK, India), internal migration issues (the USA, Malaysia), popular debates (Germany) or popular votes (Switzerland). Data from websites for researchers may also be mined to discover mega trends, for example changes in the topics studied in research (Reips and Lengler 2005).

### *Internet-Based Tests*

Assessments or tests are validated measurement instruments. Some of the most frequently used tests are personality assessments, because personality traits are quite stable and often correlate with behaviours. Thus, knowing a person's personality helps when predicting his/her behaviour. For example, there is evidence that personality predicts migration within and between US states and within Finland (Jokela 2009; Jokela *et al.* 2008). A 'typical' migrant personality may even exist (Boneva and Frieze 2001).

Tests have, in recent years, been ported to the Internet, mostly with success. Buchanan and Smith (1999) found that an Internet-based measure of self-monitoring (a personality trait associated with the extent to which individuals self-present differently in various social situations) showed similar psychometric properties to its paper-and-pencil equivalent (Snyder 1974) and compared favourably as a measure of self-monitoring. Buchanan *et al.* (2005b) observed acceptable levels of internal reliability for a modified International Personality Item Pool (IPIP) inventory and significant correlations with the relevant criterion variables across two studies using different recruiting techniques. The issue of the psychometric equivalence of paper-and-pencil versions of questionnaires with their Web-based counterparts will remain subject to research; issues like computer anxiety obviously account for differences.



**Figure 1.** An illustration of the use of Google Trends (<http://www.google.com/trends>) for migration research. Search volume for the term “migration” (upper line) correlates negatively with the term’s appearance in the media (news reference volume, lower line).  
With kind permission from Google.

### Internet-Based Experiments

Experimental research is rare in the field of migration. The definition of ‘experiment’ requires the systematic manipulation of independent variable(s), repeatability and random assignment to conditions. These criteria are hard to meet in migration research, and often an experimental variation of conditions is not suited to answer research questions in this field. A quasi-experiment would involve the non-random assignment of subjects to conditions (Campbell and Stanley 1963), but even this ‘softer’ version of the experimental method is not common in migration research. There is, however, much basic experimental research in areas related to migration, e.g. social psychological research on factors important in the *impression of and attitudes towards migrants*, and in the stereotyping of migrant groups. Regarding these questions, it is very easy to set up experiments, and we report below an example from our own research at the University of Zurich (Reips and Welter 2009).

This experiment was conducted in Switzerland, where a large proportion of migrants had come from South-Eastern Europe during the Balkan wars. We sought to create different impressions of a migration background by varying just three items on

fabricated Facebook profiles: name, hometown, and name of football club in 'Activities'. The South-East European name was *Ermal Ahmeti*, the central European *Christoph Hefti*. Hometowns were *Prishtë, Albania* and *Zürich, Switzerland*. The football clubs were *FC Albani HWZ* and *FC HWZ*. All other information, including schools and job positions, was identical. Figure 2 shows a typical Facebook profile layout at the time.

The interactive Web-based experiment generator WEXTOR at <http://wextor.org> was used to create the experiment. The participants in all the conditions were assigned randomly by WEXTOR to one of the two levels of the migration variable (among other variables, reported below). Invitations were sent to potential participants from multiple backgrounds (not only university students). Those who responded were randomly distributed to the experimental conditions.

The variation of the Facebook profiles was very successful—45 out of 50 participants identified the central European non-migrant profile as 'West/Central', and 36 out of 51 identified the 'South-Eastern European' migrant profile as 'South-East/Balkan'. Thus, the apparent migration background of owners of Facebook profiles can be easily manipulated in Internet-based research studies by changing just a few criteria.

### *Ethics, Sensitive Issues and Semi-Legal Status*

Because the Internet is a relatively new environment for conducting studies, researchers from different fields are still debating the details of ethical conduct in setting up such studies. While it has become clear and undisputed that the fundamental principles of research ethics (e.g. 'Do no harm to research participants', 'Only do research that improves the body of knowledge', etc.) apply to Internet research (Dzcyk 2001; Ess 2007; Reips 1997), there are conflicts in the details of *how* such principles can be applied in an international transdisciplinary Internet, in which space and time have become unimportant, while human-computer interaction and technological changes have become highly important for the participant experience.

A typical case of conflict is the valuation of authorship versus anonymity by participants in research. Via search engines, a portion of a sentence may suffice to identify who has written something, so the traditional techniques of protecting the identities of research participants by using nicknames etc. work less well on the Internet. On the other hand, many researchers are surprised to see that their participants actually want to be identified as authors of their words, even in the sense of owning the copyright (which may become an issue if the researcher wants to cite lengthy portions from interviews).

Ethics codes and legal ramifications vary by country, field and research tradition. In an attempt to carefully limit the scope of guidelines, Barchard and Williams thus address their advice on 'designing online studies that meet accepted standards for informed consent, deception, debriefing, the right to withdraw, security of test materials, copyright of participants' materials, confidentiality and anonymity, and





**Figure 2.** Example of a Facebook profile info tab and data collected from it in the case study example. With kind permission from Unai G.

*Note:* This example does not include all of the data collected in the case study example.

avoiding harm' (2008: 1111) to psychologists working in the USA only. Other sources focus on particular fields, e.g. ethnographic research (Ess 2007) or issues like the sensitivity of medical information (Buchanan *et al.* 2007). It is questionable whether it will be possible to develop a detailed ethics code for Internet-based research in the foreseeable future. Thus, the general principles of ethics may actually be strengthened, because each agent in Internet-based research will have to think through them when designing and evaluating his or her research.

The sensitivity of information is an issue that applies to a large proportion of migration research using ICTs. Some target groups may be easier to study via the Internet, because persons belonging to these groups will only reveal critical information under the protection of anonymity. In many studies with migrants, participants may not want to reveal any identifying information, particularly if they have an illegal or semi-legal inhabitant status. The Internet is an ideal environment for data collection in this situation. Mangan and Reips (2007) describe two Web surveys on the sensitive and rare condition of *sexsomnia* that reached more than five times as many participants from the target population than all nine previous studies from 20 years of research combined. Other groups that have been accessed for research more easily on the Internet include drug-dealers (Coomber 1997) and Ecstasy-users (Rodgers *et al.* 2001; Scholey *et al.* 2010).

### *Social Media*

Social media technologies are rapidly growing (e.g. Facebook now has some 750 million active users—see Facebook 2011) and offer researchers a number of exciting opportunities for recruiting and learning about migrants. First, social media provide emerging methodological means of recruitment for participation in Web-based surveys or experiments. For example, using Facebook groups can be an efficient and fruitful way to target specific samples of interest (e.g. Athletic Bilbao football fans living in the US versus Athletic Bilbao football fans who have not migrated or have migrated to another country). Second, the pervasiveness of social media use in everyday life has made social media interfaces a recognisable, reputable and ecologically valid way to study impressions of and attitudes towards migrants from particular groups. As discussed in an example study described above, creating and manipulating content on bogus Facebook profiles proved a successful tool for assessing attitudes towards migrants from the Balkans in Switzerland.

The third feature of social media—including blogs, social networking profiles and micro-blogs (e.g. Twitter)—that makes it useful for migration research is that it is a convenient place to find rich stores of non-reactive data. Non-reactive data mined from these sources can also be combined with survey or experimental data to gain insight into more complex research questions. Next, we discuss a step-by-step methodological example of how social networking profiles can be examined and combined with survey data. Then, we specifically demonstrate how the methodological steps can be carried out to study a question related to migrant psychology.

## Four Steps for Using Social Networking Web Pages to Study Migrant Psychology

### *Step 1: Data Collection From Social Networking Profile-Owners*

*Obtaining permission.* When doing research using social networking Web pages, permission must first be obtained from the owner of the Web page. This can be accomplished by asking volunteer Web page-owners to sign a waiver stating that they agree to allow the researchers to electronically save their Web page and use the information displayed on it in future research (Buffardi and Campbell 2008). Web page-owners should also have the option to decline permission.

*Saving the Web pages.* If the owner agrees to allow the researchers to include his or her social networking page in the study, he or she should be asked to log in and open the pages which the researchers have decided to utilise. It is important, at this stage, to account for the structure of social networking websites. There are usually a number of different Web pages that an owner accesses when he or she logs in. For example, there is the profile (see Figure 2), which displays personal information about the owner to others in the profile-owner's network (e.g. birthday, photo, short self-description, etc.), the owner's news feed—which displays the recent actions of those in his or her network to the owner alone (e.g. Friend X has posted a link to a news story, Friend Y has posted new photographs)—the message board, where the owner and those in his/her network can post public messages, and other separate components each with their own unique URL. The researcher should then save each page to be used in the study to a separate file. This is necessary because accessing different components of the owner's social networking Web page will not be possible once offline if only one component has been saved.

*Collecting other information from the Web page-owners.* So far data collection in this example has been non-reactive. However, researchers might also need to obtain other information about the Web page-owner that cannot be collected from his or her social networking Web page. For example, demographics or information about the owner's background might be important (e.g. languages spoken, first-, second- or third-generation immigrant, etc.). Personality measures, information about the owner's social networking website use, or outcome measures of interest, such as psychological well-being, might also be vital to collect in order to test the research hypothesis. Such self-reports may be obtained by interview or questionnaire.

### *Step 2: Recording Quantitative Aspects of the Profiles*

Social networking Web pages contain a variety of numerical social indicators as well as countable information that might be of interest to researchers. Therefore, after gaining access to owners' pages, researchers might want to gather some of these quantitative data points. Social networking pages, for example, provide the number of friends a user has, a number of photographs the user has been 'tagged' in, the number of groups the user belongs to, and many tallies. Other information from social networking pages can be counted, including, for instance, the number of

favourite TV shows a user has listed, the number of lines of text a user includes in his or her biography, the number of wallposts that have been posted on a user's page in the last 24 hours or seven days, etc. Counting information is a very flexible method that can provide an array of information to researchers with varying interests.

### *Step 3: Coding Subjective Aspects of the Profiles*

In addition to gathering numerical data from social networking Web pages, researchers can also collect subjective data in a systematic way by using coding schemes. Independent coders can rate aspects of a social networking Web page on a numerical scale to obtain a measurement of the extent to which subjective features are present. For example, in previous research, the owner's main profile photograph was coded for how sexy the owner appeared in it by obtaining five independent coders' ratings of how sexy, modest (reversed), and clothed (reversed) the owner appeared on a scale of 1 to 7 (Buffardi and Campbell 2008). By adapting this method, the components of social networking Web pages (e.g. main photograph, biographical information, messages on the owner's page, owner's friends' activity) can be coded to measure a limitless number of variables of interest.

The key to creating valid and reliable qualitative coding schemes involves two steps. First, researchers should use multiple items to assess each variable of interest because multiple-item in comparison to single-item measures increase the reliability of measurements (Hinkin and Schriesheim 1989). For example, if one was interested in assessing how humorous the quotes were on a social networking Web page profile, coders should read the quotations and rate them on how funny, humorous, witty and hilarious they seem to be, as opposed to only how humorous they are. Second, achieving appropriate levels of agreement amongst coders is essential. Researchers should be aware that achieving adequate levels of intercoder reliability might involve training the coders to agree on the definition of codes, make important distinction between codes, etc. It is important to allow time for practice coding, discussion amongst coders, and detailed note-taking during the process of preparing to code. Finally, sufficient levels of intercoder or interrater reliability (e.g. Cronbach's alphas  $\geq .70$ ; see Cronbach 1951) should be achieved before composite scores are created and the main analyses conducted. Free web-based software is available from Freelon (2010) to calculate intercoder agreement.

There is one remaining methodological note to mention with respect to qualitative coding. The components that are being coded should be isolated from the rest of the social networking page as much as possible. For example, when coding for themes in the Bio section of a Facebook profile, the content of the Bio section should be copied and pasted into a word-processing document (or something similar) in order to avoid 'contaminating' coders' judgments with the other textual information and pictures surrounding it on the profile page.

#### *Step 4: Data Analysis*

The final methodological step is data analysis. When using data from social networking Web pages, usually correlation/regression analysis is employed.

### **Applying the Four Steps to Study Migrants' Psychology: A Case Study of Biculturalism**

Now, we demonstrate how the four steps to using social networking Web pages in research can be carried out in a study of migrant psychology. In this example, we discuss markers of bi- (or multi-)culturalism on social networking Web pages. We define biculturalism as a measure of individual difference (e.g. Phinney 1992; Ryder *et al.* 2000; Tsai *et al.* 2000). That is to say, we assume that individuals vary in the extent to which they are bi- or (multi-)cultural and this individual difference is measured on a continuum. Those who identify as being highly bicultural may be immigrants or first-generation citizens. Others who identify as bicultural may have residences, family members or jobs in different parts of the world. In this example, we examine whether individuals' generational immigration/migration status relates to the content of their social networking Web pages, specifically the inclusion of markers of biculturalism.

Previous research has shown differences between immigrants, the children of immigrants, the grandchildren of immigrants and so on in their degrees of biculturalism (Edmonston and Passel 1994). Thus, we expect that a negative linear relationship exists between generational proximity to immigration and the inclusion of markers of biculturalism on a person's social networking Web page. Specifically, we hypothesise that those who are immigrants/migrants themselves or who have closer generational contact to immigrants/migrants will display more markers of biculturalism. We will use Facebook.com, currently the most popular social network worldwide, in this example.

#### *Step 1: Data Collection*

*Obtaining permission.* First, volunteer Facebook page-owners will sign a waiver stating that they agree to allow us to electronically save their Web pages and use the information displayed in this research.

*Saving the Web pages.* Second, if the owner agrees to allow the researchers to include his or her social networking page in the study, he or she will be asked to log into Facebook.

Third, in this example study, we decided to collect the following components of Facebook pages: the Wall, the Info Tab, the Photos Tab and the News Feed. Therefore, the researcher will save each of these components separately. We suggest saving the individual components of interest using a labelling system that fits the needs of the researchers. In this study, for example, we will save a file under the

title '1234\_Wall'. This file name would refer to participant number 1234's Wall tab file.

*Collecting other information from the Web page-owner.* Fourth, all Facebook page-owners (whether or not they agreed to allow the researchers to save their social networking Web page) will provide demographic information, including their generational immigration/migration status. Fifth and finally, all participants will be thanked for participating and dismissed.

A note on sample size: in this type of correlational study, a sample size of 100 Facebook profiles is adequate. Collecting more than 100 (say 125–250) would be even better to increase statistical power. Collecting fewer than 100 might lead to problems detecting effects.

### *Step 2: Quantitative Measures*

Once the collection of Facebook pages is complete, we will gather quantitative data from the saved Web pages. First, we will begin with the Info tab that lists demographic, familial and biographic information. We will record whether the owner's hometown and current city are in different countries using a dichotomous yes/no format. We will also determine and record the distance (in miles or kilometres) between the hometown and the current city of residence using an Internet mapping website (e.g. Google Maps). On the Info tab, we will count and record the number of countries in which the owner has attended school. Finally, on the Info tab, we will count and record the number of countries in which the owner's current and former employers are located.

Second, we will turn to the Wall tab. On Facebook, the Wall is an electronic bulletin board on which the owner of the page and his friends can post public messages. The Wall also lists the owner's most recent Facebook activities publicly. On the Wall, we will count and record the number of languages that messages are written in.

Third, on the News Feed, we will also count and record the number of languages that appear. Importantly, if the owner of the Facebook page has excluded any of the information we have decided to collect (e.g. does not list employers), we will record the data as missing or as intentionally excluded.

### *Step 3: Subjective Coding*

In this step, we will code subjective information included on the Facebook pages. First, on the Info tab, we will examine three different sections: 'Bio', 'Quotations' and 'Likes and Interests'. The Bio generally contains the Facebook page-owner's brief self-description. The Quotation section includes a selection of his or her favourite quotes. The Likes and Interests section is a list of the page-owner's favourite sports teams, books, movies, music, television shows, etc. Because the information in these sections varies greatly between users and because the extent to which this information can be

considered indicative of a bi- (or multi-)cultural identity is subjective, the most systematic way to extract information is by coding.

The text in the Bio sections will first be copied and pasted into a separate word-processing document. Then, five independent coders will rate the 'Bio' section on the following adjectives: multicultural, international and multinational, using a Likert-type scale of 1 (Not at all) to 7 (Very much). A similar procedure will be followed for the Quotations and Likes and Interests. They will also be copied and pasted into separate word-processing documents, and five coders will rate each on the following adjectives: multicultural, multinational, diverse, homogeneous (reversed) and culturally varied, using the 1–7 rating scale.

Next, we will turn to the Photos tab. On this tab, the most recent 15 photographs that the owner has been 'tagged' in appear, as well as the first photograph in each of the five photo albums that the owner has the most recently created. Five coders will rate the content of the visible photographs on the multicultural, multinational and culturally varied, again using a scale of 1 (Not at all) to 7 (Very much). Finally, the same coding will be used on the News Feed.

Importantly, we must then check to ensure that sufficient levels of inter-rater agreement have been achieved. To do this, Cronbach's alphas will be used to quantify the agreement between the five coders on each adjective (i.e. separate alphas will be calculated for the Bio section's multicultural ratings, international ratings and so on, up to the News Feed adjective ratings). If the majority of the alphas are  $\geq .70$ , we will have sufficient agreement to continue with the main analyses (Cronbach 1951).

#### *Step 4: Data Analysis*

To prepare the data for hypothesis testing, we will create indices from the coders' subjective ratings. Free software that is available as a Web service will allow us to do this easily (Freelon 2010), but for the sake of understanding what is happening in each step we will do it 'by hand' here. First, we will measure agreement within coders for each Facebook component. For example, we will obtain a Cronbach's alpha for the three adjectives that the Bio sections were rated on for each coder (this will result in five separate alphas measuring the internal agreement of each coders' ratings of the level of multiculturalism displayed in the Bio sections). If the mean of the resulting alpha statistic (most correctly found by first transforming the five alphas into Fisher's z-scores, taking the mean of them, and transforming the mean back from a 'z' to an 'r' score) is sufficiently high (again,  $\geq .70$ ), we will create a Bio\_Multicultural index by taking the mean of all the coders' ratings on the three adjectives. The same steps will be repeated for the other Facebook-page components to ultimately create an index of the extent to which each component was perceived as multicultural by the five coders.

At this point, we will be prepared to conduct the main correlation analyses. We will predict the quantitative information collected and the subjective code indices with generational immigration/migration status. If our hypothesis is supported,

significant inverse relationships will be found between generational proximity to immigration/migration and, at least, some of the markers of biculturalism that we have gathered from Facebook pages. Those who are generationally closer to immigration/migration should display greater quantities of information linked with biculturalism on their social networking Web pages and higher subjective multicultural content scores.

## Conclusion and Outlook

What methods are available to study migrant psychology with the help of the Internet and how will they evolve in the future? In this paper we have presented Internet-based research methods and examples that may be useful to those who study migrants, attitudes towards migrants, and migration. We have shown how different types of method can be combined, and explained in a step-by-step procedure how profiles in social media can be coded and thus made accessible to migration research.

Internet-based methods provide researchers with numerous advantages, ranging from easy access to large numbers of participants with a diverse or very specific background, to increased anonymity that may be of particular benefit in research on illegal migration. There is little doubt that Internet-based research methods will continue to develop quickly, and they will grow with and be adapted to the speedy changes constantly occurring in social media. Facebook alone has released more than a dozen waves of fundamental changes, in particular regarding the options and interface available for modifying one's privacy settings (Reips 2011). Widespread use of HTML5, an imminent change in Web technology, will allow input on Web pages to be more easily validated. It will also bring about better use options for researchers, like identifying geo-location and drag-and-drop. The use of social media will further spread to smartphones and other handheld devices in the future and this will create new sampling options.

As the use of information and communication technologies like social media becomes even more widespread and affordable, the possibilities for using it in migration research will increase. The modern migrant is likely to be a user of such technologies. Thus, researchers who embrace new technologies and adapt their research accordingly will greatly benefit.

## Notes

- [1] These archives and some other designated Web sites with examples for Internet-based social and behavioural studies currently in progress are listed here:

*Web experiment list* (Reips and Lengler 2005): <http://wexlist.net/>

*Web survey list*: <http://wexlist.net/browse.cfm?action=browse&modus=survey>

*Web Experimental Psychology Lab* (Reips 2001): <http://wexlab.eu/>

Krantz' *Psychological Research on the Net*: <http://psych.hanover.edu/research/exponnet.html>

*Online Social Psychology Studies* by Plous: <http://www.socialpsychology.org/expts.htm>



- [2] Despite the increase in universal access to the Internet, there also is a widening digital divide that becomes more important the more essential aspects of daily life move to the Internet and are likely to remain for decades.

## References

- Arnett, J.J. (2002) 'The psychology of globalization', *American Psychologist*, 57(10): 774–83.
- Askatas, N. and Zimmermann, K.F. (2009) 'Google econometrics and unemployment forecasting', *Applied Economics Quarterly*, 55(2): 107–20.
- Bacigalupe, G. and Cámara, M. (2012) 'Transnational families and social technologies: reassessing immigration psychology', *Journal of Ethnic and Migration Studies*, 38(9). doi:10.1080/1369183X.2012.698211.
- Barchard, K.A. and Williams, J. (2008) 'Practical advice for conducting ethical online experiments and surveys for United States psychologists', *Behavior Research Methods*, 40(4): 1111–28.
- Berners-Lee, T. (1998) *The World Wide Web: A Very Short Personal History*. Retrieved 14 September 2010 from <http://www.w3.org/People/Berners-Lee/ShortHistory.html>.
- Berry, J.W. (1990) 'Psychology of acculturation: understanding individuals moving between cultures', in Brislin, R.W. (ed.) *Applied Cross-Cultural Psychology*. Newbury Park, CA: Sage, 232–53.
- Birnbaum, M.H. (2001) 'A Web-based program of research on decision making', in Reips, U.-D. and Bosnjak, M. (eds) *Dimensions of Internet Science*. Lengerich: Pabst Science, 23–55.
- Boneva, B.S. and Frieze, I.H. (2001) 'Toward a concept of migrant personality', *Journal of Social Issues*, 57(3): 477–91.
- Buchanan, T. and Smith, J.L. (1999) 'Using the Internet for psychological research: personality testing on the World Wide Web', *British Journal of Psychology*, 90(1): 125–44.
- Buchanan, T., Ali, T., Heffernan, T.M., Ling, J., Parrott, A.C., Rodgers, J. and Scholey, A.B. (2005a) 'Nonequivalence of on-line and paper-and-pencil psychological tests: the case of the prospective memory questionnaire', *Behavior Research Methods*, 37(1): 148–54.
- Buchanan, T., Johnson, J.A. and Goldberg, L. (2005b) 'Implementing a five-factor personality inventory for use on the internet', *European Journal of Psychological Assessment*, 21(2): 115–27.
- Buchanan, T., Joinson, A.N., Paine, C. and Reips, U.-D. (2007) 'Looking for medical information on the Internet: self-disclosure, privacy and trust', *He@lth Information on the Internet*, 58(1): 8–9.
- Buffardi, L.E. and Campbell, W.K. (2008) 'Narcissism and social networking Web sites', *Personality and Social Psychology Bulletin*, 34(10): 1303–14.
- Campbell, D. and Stanley, J. (1963) *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton, Mifflin.
- Conversi, D. (2012) 'Irresponsible radicalisation: diasporas, globalisation and long-distance nationalism in the digital age', *Journal of Ethnic and Migration Studies*, 38(9). doi:10.1080/1369183X.2012.698204.
- Coomber, R. (1997) 'Using the Internet for survey research', *Sociological Research Online*, 2(2): retrieved 16 June 2002 from <http://www.socresonline.org.uk/2/2/2.html>.
- Coutts, E. and Jann, B. (2011) 'Sensitive questions in online surveys: experimental results for the randomized response technique (RRT) and the unmatched count technique (UCT)', *Sociological Methods and Research*, 40(1): 169–93.
- Cronbach, L.J. (1951) 'Coefficient alpha and the internal structure of tests', *Psychometrika*, 16(3): 297–334.
- Dandurand, F., Shultz, T.R. and Onishi, K.H. (2008) 'Comparing online and lab methods in a problem-solving experiment', *Behavior Research Methods*, 40(2): 428–34.

- De Leeuw, E.D. (2005) 'To mix or not to mix data collection modes in surveys', *Journal of Official Statistics*, 21(2): 233–55.
- Dillman, D.A., Reips, U.-D. and Matzat, U. (2010) 'Advice in surveying the general public over the Internet', *International Journal of Internet Science*, 5(1): 1–4.
- Dzeyk, W. (2001) 'Ethische Dimensionen der Online-Forschung', *Kölner Psychologische Studien*, 6(1): 1–30. Online at: <<http://www.allg-psych.uni-koeln.de/dzeyk/home/docs/ethdimon.pdf>>.
- Edmonston, B. and Passel, J.S. (1994) 'Ethnic demography: US immigration and ethnic variations', in Edmonston, B. and Passel, J.S. (eds) *Immigration and Ethnicity*. Washington, DC: The Urban Institute Press, 1–30.
- Ess, C. (2007) 'Internet research ethics', in Joinson, A., McKenna, K., Postmes, T. and Reips, U. (eds) *The Oxford Handbook of Internet Psychology*. Oxford: Oxford University Press, 487–502.
- Facebook (2011) *Statistics*. Retrieved 21 August 2011 from <http://www.facebook.com/press/info.php?statistics>.
- Freelon, D.G. (2010) 'ReCal: intercoder reliability calculation as a Web service', *International Journal of Internet Science*, 5(1): 20–33.
- Frick, A., Bächtiger, M.T. and Reips, U.-D. (2001) 'Financial incentives, personal information, and drop-out in online studies', in Reips, U.-D. and Bosnjak, M. (eds) *Dimensions of Internet Science*. Lengerich: Pabst Science, 209–19.
- Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S. and Brilliant, L. (2009) 'Detecting influenza epidemics using search engine query data', *Nature*, 457: 1012–15.
- Görzitz, A.S. (2006) 'Incentives in Web studies: methodological issues and a review', *International Journal of Internet Science*, 1(1): 58–70.
- Harker, K. (2001) 'Immigrant generation, assimilation, and adolescent psychological well-being', *Social Forces*, 79(3): 969–1004.
- Hewson, C. (2007) 'Gathering data on the Internet: qualitative approaches and possibilities for mixed methods and research', in Joinson, A., McKenna, K., Postmes, T. and Reips, U.-D. (eds) *The Oxford Handbook of Internet Psychology*. Oxford: Oxford University Press, 405–28.
- Hinkin, T.R. and Schriesheim, C.A. (1989) 'Development and application of new scales to measure the French and Raven (1959) bases of social power', *Journal of Applied Psychology*, 74(4): 561–7.
- Janetzko, D. (2002) 'Artificial dialogues: dialogue and interview bots for the World Wide Web', in Batinic, B., Reips, U.-D. and Bosnjak, M. (eds) *Online Social Sciences*. Toronto: Hogrefe and Huber, 357–64.
- Janetzko, D. (2008) 'Objectivity, reliability, and validity of search engine count estimates', *International Journal of Internet Science*, 3(1): 7–33.
- Joinson, A.N. and Reips, U.-D. (2007) 'Personalized salutation, power of sender and response rates to Web-based surveys', *Computers in Human Behavior*, 23(3): 1372–83.
- Jokela, M. (2009) 'Personality predicts migration within and between US states', *Journal of Research in Personality*, 43(1): 79–83.
- Jokela, M., Elovainio, M., Kivimäki, M. and Keltikangas-Järvinen, L. (2008) 'Temperament and migration patterns in Finland', *Psychological Science*, 19(8): 831–7.
- Kissau, K. (2012) 'Structuring migrants' political activities on the Internet: a two-dimensional approach', *Journal of Ethnic and Migration Studies*, 38(9). doi:10.1080/1369183X.2012.698207.
- Krantz, J.H. and Dalal, R. (2000) 'Validity of Web-based psychological research', in Birnbaum, M.H. (ed.) *Psychological Experiments on the Internet*. San Diego: Academic Press, 35–60.
- Krantz, J.H., Ballard, J. and Scher, J. (1997) 'Comparing the results of laboratory and World Wide Web samples on the determinants of female attractiveness', *Behavior Research Methods, Instruments, and Computers*, 29(2): 264–9.

- Lay, C.H. and Safdar, F.S. (2003) 'Daily hassles and distress among college students in relation to immigrant and minority status', *Current Psychology: Developmental, Learning, Personality, Social*, 22(1): 3–22.
- Mangan, M. and Reips, U.-D. (2007) 'Sleep, sex, and the Web: surveying the difficult-to-reach clinical population suffering from sexsomnia', *Behavior Research Methods*, 39(2): 233–6.
- Musch, J. and Reips, U.-D. (2000) 'A brief history of Web experimenting', in Birnbaum, M.H. (ed.) *Psychological Experiments on the Internet*. San Diego: Academic Press, 61–88.
- Musch, J., Bröder, A. and Klauer, K.C. (2001) 'Improving survey research on the World-Wide Web using the randomized response technique', in Reips, U.-D. and Bosnjak, M. (eds) *Dimensions of Internet Science*. Lengerich: Pabst Science, 179–92.
- Nedelcu, M. (2012) 'Migrants' new transnational habitus: rethinking migration through a cosmopolitan lens in the digital age', *Journal of Ethnic and Migration Studies*, 38(9). doi:10.1080/1369183X.2012.698203.
- Oiarzabal, P.J. (2012) 'Diaspora Basques and online social networks: an analysis of users of Basque institutional diaspora groups on Facebook', *Journal of Ethnic and Migration Studies*, 38(9). doi:10.1080/1369183X.2012.698216.
- Paine, C., Reips, U.-D., Stieger, S., Joinson, A. and Buchanan, T. (2007) 'Internet users' perceptions of "privacy concerns" and "privacy actions"', *International Journal of Human-Computer Studies*, 65(6): 526–36.
- Pardos-Prado, S. (2011) 'Framing attitudes towards immigrants in Europe: when competition does not matter', *Journal of Ethnic and Migration Studies*, 37(7): 999–1015.
- Phinney, J.S. (1992) 'The multigroup ethnic identity measure: a new scale for use with diverse groups', *Journal of Adolescent Research*, 7(2): 156–76.
- Phinney, J.S. (1996) 'When we talk about American ethnic groups, what do we mean?', *American Psychologist*, 51(9): 918–27.
- Reips, U.-D. (1997) 'Das psychologische Experimentieren im Internet', in Batinic, B. (ed.) *Internet für Psychologen*. Göttingen: Hogrefe, 245–65.
- Reips, U.-D. (2001) 'Merging field and institution: running a Web laboratory', in Reips, U.-D. and Bosnjak, M. (eds) *Dimensions of Internet Science*. Lengerich: Pabst Science, 1–22.
- Reips, U.-D. (2002) 'Context effects in Web surveys', in Batinic, B., Reips, U.-D. and Bosnjak, M. (eds) *Online Social Sciences*. Göttingen: Hogrefe and Huber, 69–79.
- Reips, U.-D. (2006) 'Web-based methods', in Eid, M. and Diener, E. (eds) *Handbook of Multimethod Measurement in Psychology*. Washington DC: American Psychological Association, 73–85.
- Reips, U.-D. (2010) 'Design and formatting in Internet-based research', in Gosling, S. and Johnson, J. (eds) *Advanced Internet Methods in the Behavioral Sciences*. Washington, DC: American Psychological Association, 29–43.
- Reips, U.-D. (2011) 'Privacy and the disclosure of information on the Internet: issues and measurement', in Blachnio, A., Przepiórka, A. and Rowinski, T. (eds) *Internet in Psychological Research*. Warsaw: UKSW, 71–104.
- Reips, U.-D. and Funke, F. (2008) 'Interval level measurement with visual analogue scales in Internet-based research: VAS Generator', *Behavior Research Methods*, 40(3): 699–704.
- Reips, U.-D. and Garaizar, P. (2011) 'Mining Twitter: microblogging as a source for psychological wisdom of the crowds', *Behavior Research Methods*, 43(3): 635–42.
- Reips, U.-D. and Krantz, J. (2010) 'Conducting true experiments on the Web', in Gosling, S. and Johnson, J. (eds) *Advanced Internet Methods in the Behavioral Sciences*. Washington, DC: American Psychological Association, 193–216.
- Reips, U.-D. and Lengler, R. (2005) 'The Web Experiment List: a Web service for the recruitment of participants and archiving of Internet-based experiments', *Behavior Research Methods*, 37(2): 287–92.

- Reips, U.-D. and Stieger, S. (2004) 'Scientific LogAnalyzer: a Web-based tool for analyses of server log files in psychological research', *Behavior Research Methods, Instruments, and Computers*, 36(2): 304–11.
- Reips, U.-D. and Welter, M. (2009) 'Experimenting with Facebook profiles: experimental setting, presence of an authority, and migrant background'. Zurich: University of Zurich, unpublished manuscript.
- Rodgers, J., Buchanan, T., Scholey, A.B., Heffernan, T.M., Ling, J. and Parrott, A. (2001) 'Differential effects of Ecstasy and cannabis on self-reports of memory ability: a web-based study', *Human Psychopharmacology: Clinical and Experimental*, 16(8): 619–25.
- Ryder, A.G., Alden, L.E. and Paulhus, D.L. (2000) 'Is acculturation unidimensional or bidimensional? A head-to-head comparison in the prediction of personality, self-identity, and adjustment', *Journal of Personality and Social Psychology*, 79(1): 49–65.
- Scholey, A.B., Owen, L., Gates, J., Rodgers, J., Buchanan, T., Ling, J., Heffernan, T.M., Swan, P., Stough, C. and Parrott, A.C. (2010) 'Hair MDMA samples are consistent with reported Ecstasy use: findings from an internet study investigating effects of Ecstasy on mood and memory', *Neuropsychobiology*, 63(1): 15–21, <http://dx.doi.org/10.1159/000321833>.
- Shih, T.-H. and Fan, X. (2007) 'Response rates and mode preferences in Web-mail mixed-mode surveys: a meta-analysis', *International Journal of Internet Science*, 2(1): 59–82.
- Snyder, M. (1974) 'Self monitoring of expressive behavior', *Journal of Personality and Social Psychology*, 30(4): 526–37.
- Stieger, S. and Reips, U.-D. (2008) 'Dynamic Interviewing Program (DIP): automatic online interviews via the instant messenger ICQ', *CyberPsychology and Behavior*, 11(1): 201–7.
- Stieger, S., Reips, U.-D. and Voracek, M. (2007) 'Forced-response in online surveys: bias from reactance and an increase in sex-specific dropout', *Journal of the American Society for Information Science and Technology*, 58(11): 1653–60.
- Tsai, J.L., Ying, Y.-W. and Lee, P.A. (2000) 'The meaning of "being Chinese" and "being American": variation among Chinese American young adults', *Journal of Cross-Cultural Psychology*, 31(3): 302–32.
- Voracek, M., Stieger, S. and Gindl, A. (2001) 'Online replication of evolutionary psychology evidence: sex differences in sexual jealousy in imagined scenarios of mate's sexual versus emotional infidelity', in Reips, U.-D. and Bosnjak, M. (eds) *Dimensions of Internet Science*. Lengerich: Pabst Science, 91–112.