

myPersonality project: Example of successful utilization of online social networks for large-scale social research

David J. Stillwell, Michal Kosinski
The Psychometrics Centre, University of Cambridge, UK
ds617@cam.ac.uk, mk583@cam.ac.uk

1. BACKGROUND

The Internet has made it possible for social scientists to inexpensively reach large samples of research participants, sometimes hundreds of thousands of people or more [12]. These samples are more diverse and representative than the WEIRD (Western, Educated, Industrialized, Rich and Democratic) [7] samples traditionally used by researchers [5, 6], and the quality of data collected is as good or better than in traditional pencil and paper methods [11] or even national surveys acquired via telephone polling [1].

In recent years there has been a remarkable shift towards more social and less anonymous Internet use. Interactions between people using anonymous nicknames, email addresses, or avatars are increasingly replaced by interactions within Online Social Networks (OSN) that are based on real identities and connections that largely mirror offline social links [2, 3, 10]. This presents a great opportunity in terms of research design, solves problems related to anonymity that were characteristic to previous web studies, and makes the data available on OSNs both far richer and on a larger scale than has been available to social scientists in the past [8].

The most popular OSN today, Facebook, boasts over 800 million active users (as of December 2011) that spend on average 41 minutes daily interacting on the platform.

2. MYPERSONALITY APPLICATION

The myPersonality application was released in June 2007. It offers Facebook users a set of genuine personality and ability measures, and then gives them personalized feedback on their results.

To date, over 6 million users completed the most popular questionnaire, the IPIP version of the NEO Personality Inventory [4], and nearly half of them have allowed us to anonymously record the information stored on their Facebook profiles. The number of measures provided by myPersonality is constantly expanding; as of January 2012 there were more than 25 scales available that ranged from the 336 item long IPIP NEO personality facets questionnaire to our own adaptive IQ test.

The great majority of myPersonality's data were made freely available to other academic researchers on the myPersonality project website [13], including records of online behaviour, rich demographics, friendship networks, and scores on other psychometric measures.

Users do not receive a payment or course credit for using myPersonality. This provides better quality data, since unmotivated or careless participants may answer randomly, skip, or misread the test items [9]. The myPersonality sample is also more representative than those obtained in traditional studies:

1. Gender

According to [6], 71% of participants in 510 traditional samples from 156 articles published in the *Journal of Personality and Social Psychology* (JPSP) were female. In comparison, in the sample of web studies only 57% of participants were female, and 63% of ours are female (Figure 1).

2. Age

Facebook population is still predominantly young, and so are myPersonality users. The average participant is 23.5 years old and nearly half of them (47%) are between 18 and 24 years old (Figure 1). Nevertheless, the age composition of the myPersonality sample seems to be very similar to that of traditional studies. As estimated by [6], the average age in the JPSP samples was 23 years.

3. Geographic region

There are few international barriers in web-surfing and thus online based studies are inexpensively available to geographically detached populations. myPersonality attracts more than 42% of respondents from outside the US. Additionally, 44 countries are represented by more than 1000 respondents. Consequently, the geographical diversity of our respondents (note that myPersonality is currently available exclusively in English) is 2.5 times higher than in the published samples reviewed by [6], which had 17% non-US participants.

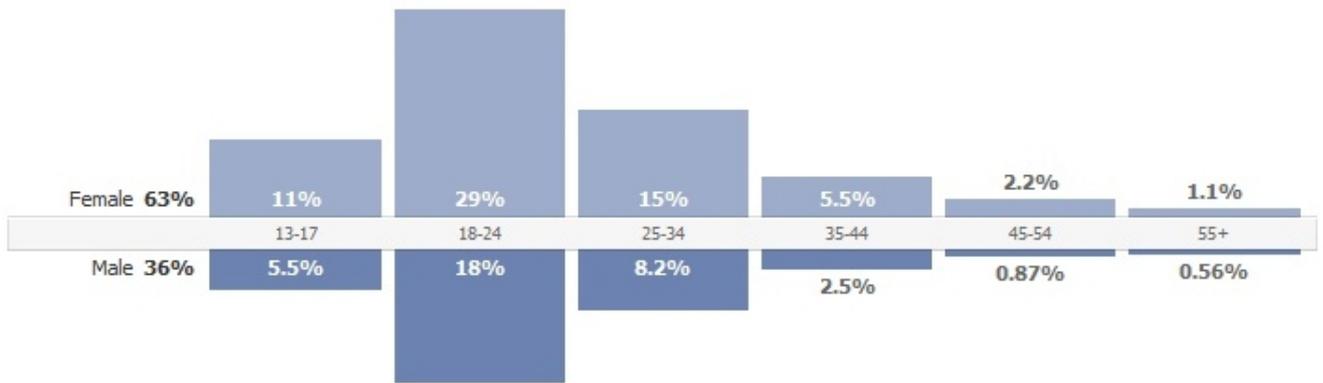


Figure 1: Gender and age distribution of myPersonality users

3. COMPLETED AND ONGOING RESEARCH

myPersonality has been used across social sciences to research topics such as music personality, addictive substance use and delay discounting, associations between personality and web browsing, links between personality and social network shape, the personality of popular FB and Twitter users, the privacy settings used by certain types of people, test-retest reliabilities of questionnaires, individuals' "sensational interests", and whether well-being can be measured with status updates. There are also major ongoing research programs on aggregate geographic personalities, the words different types of people use in status updates, and the structure of FB Likes and their relationship to individuals' behaviour. See: mypersonality.org/wiki

4. REFERENCES

- [1] CHANG, L., AND KROSCHICK, J. A. National surveys via rdd telephone interviewing versus the internet. *Public Opinion Quarterly* 73, 4 (2009), 641–678.
- [2] CLOUSTON, S., AMIN, S., VERDERY, A., AND GAUTHIER, G. R. The structure of undergraduate association networks: A quantitative ethnography. *Connections* 29, 2 (2009), 18–31.
- [3] ELLISON, N. B., STEINFELD, C., AND LAMPE, C. The benefits of facebook "friends": Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication* 12, 4 (2007), 28.
- [4] GOLDBERG, L. R., JOHNSON, J. A., EBER, H. W., HOGAN, R., ASHTON, M. C., CLONINGER, C. R., AND GOUGH, H. G. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality* 40, 1 (2006), 84–96.
- [5] GOSLING, S. D., SANDY, C. J., JOHN, O. P., AND POTTER, J. Wired but not weird: The promise of the internet in reaching more diverse samples. *Behavioral and Brain Sciences* 33, 2-3 (2010), 94.
- [6] GOSLING, S. D., VAZIRE, S., SRIVASTAVA, S., AND JOHN, O. P. Should we trust web-based studies? a comparative analysis of six preconceptions about internet questionnaires. *American Psychologist* 59, 2 (2004), 93–104.
- [7] HENRICH, J., HEINE, S. J., AND NORENZAYAN, A. The weirdest people in the world? *Behavioral and Brain Sciences* 33, 2-3 (2010), 61.
- [8] KLEINBERG, J. The convergence of social and technological networks. *Communications of the Acm* 51, 11 (2008), 66–72.
- [9] KURTZ, J. E., AND PARRISH, C. L. Semantic response consistency and protocol validity in structured personality assessment: The case of the neo-pi-r. *Journal of Personality Assessment* 76, 2 (2001), 315–332.
- [10] MAYER, A., AND PULLER, S. L. The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics* 92, 1-2 (2008), 329–347.
- [11] PETTIT, F. A. A comparison of world-wide web and paper-and-pencil personality questionnaires. *Behavior Research Methods Instruments & Computers* 34, 1 (2002), 50–54.
- [12] RENTFROW, P. J., GOSLING, S. D., AND POTTER, J. A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics. *Perspectives on Psychological Science* 3, 5 (2008), 339–369.
- [13] STILLWELL, D. J., AND KOSINSKI, M. mypersonality research wiki, 2011.