Studying Response Rate and Quality of Response in the Context of Social Networks

Dominique Joye[1], Marlène Sapin[2] and Alexandre Pollien[2]

[1]Institute of Social Sciences, University of Lausanne and NCCR Lives
[2]NCCR Lives and FORS Lausanne
[3]FORS, Lausanne

August 17, 2015

1   Introduction

Surveying social networks has always been an object of debate, also because some types of tools are seen as tedious for the respondents and complicated for the interviewers. Furthermore, the length of the questionnaire is sometimes depending of the number of ties, that left a great freedom for the interviewer to control the length of the task, just by playing with this parameter. By contrast, this characteristic is also a strong incentive to use self-administered surveys, even if the complexity of some tasks is clearly an argument against.

In this short study, we will discuss the differences between 3 variants of a study, realised in Switzerland at the end of the 2014, and testing each one specific kind of network generator. Furthermore, after the main survey, a supplementary questionnaire was send later, largely the same as the one leave as a drop-off of the last MOSAiCH survey (spring 2013). In this sense we will be able to control the effect to begin with a mail survey rather than with a face to face survey. More generally, this will allow us to discuss the strategy of using a drop-off after a main survey.

Furthermore we have adopted a design that allow to take into account the local context in the structure of answer by choosing a priori three localities of similar size in the French speaking Switzerland.

As far as we know, there are not so many studies that try to test the effect of differences in design on the response rate and structure of answers. Furthermore, it allows to go back to the discussion between length and burden of surveys from one side and response rate and quality of answer, probably heavily mediated through mode and interest for the study.

2   Questionnaires and Methods

For a mix of reasons between cost and efficacy, we have opted for a mail questionnaire, with a standard unconditional prepaid incentive of 10.- Sfr. With 4 contacts, we expected more than 50% of response rate.

In a second wave, we expected 50% of response, of course among the respondent of the first wave.

We have used 2500 addresses bought to a company called "AZ Direct" – from which 200 were no more valid –, and made a random and equal repartition between the 3 communities – exactly 835 individual addresses in each.

2.1   3 questionnaires and a matrix design

In the literature about networks, there is a discussion about three ways to inquire about their structure and importance in the daily life: resources, position and name generators.[1]

In summary, these three ways to survey networks can be summarized in the next paragraphs.

- Resource generators ask for a set of resources characterised by some functions. For example "Do you know somebody that..." A) can help you repare a bike? B) Can give you some support when you

---

[1]For an extensive discussion about the characteristics of these modes, see in particular the PhD thesis of Martin Van der Gaag. [Van05]

feel sad C) can help if you need to fill some administrative tasks... and so on, with some response categories like 1) "Yes, in my family", 2) "Yes, among my friends" ... until, "No, Nobody".

- Position generators ([Le08] ask for a set of occupations, chosen to describe the global social space, in term of status but perhaps also taking into account some form of horizontal differentiation. The formulation could be very similar to the Resource generator format, with the same answering scale, but asking if people know some "medical doctor", "peasant" or "computer specialist".

- Name generators ([Fis82] function with a different logic, first asking a list of person with which the respondent interacts in a meaningful way, before to detail the characteristics of these persons and eventually their interrelations.[2] Of course the time taken to answer is function of the number of relations mentioned: if surveys like the GSS has used a maximum of 5 persons that could be mentioned, dedicated surveys can use until 20 persons in the network.

In order to have a questionnaire not too long, we opted for a matrix design,[3] with each of the type of generator forming the core of one variant. The other items, were repeated in each part of the design, at least as long as the total length was kept to a reasonable value. The first two variants were limited to less than 150 ticks while the latest was more difficult to assess because a length heavily depending on the number of ties created by the name generator.

2.2 A second wave

If the three types of tools were distributed in three different questionnaire, we needed some common instrument in order to measure their impact. Outside the common variables in the questionnaire, we have added a supplement questionnaire, as drop-off distributed in a second wave. In order to maximise the possibility of comparison, this second wave, putting more accent on family networks, was largely the same as the one used in a second step after the MOSAiCH 2013 survey.

For preserving the goodwill of the respondents, even more in a case where we have not announced before this second questionnaire, we have made only one recall, with the risk that the final response rate at the end will be lower than it could have been with a more aggressive strategy.

2.3 Three local communities

We have chosen to limit the survey to a small part of Switzerland for many reasons:

- To avoid translation and stick to a one language solution: more reasonable choice in an experimental setting;

- To use explicitly three different contexts in order to contrast the results according different situations

- to choose medium sized communities, not too small in order to avoid to have to interview all inhabitants but at the same time small enough in order to survey a non-negligible part of the population.

Consequently, we have surveyed 3 Swiss French-speaking local communities, sending a questionnaire to 835 addresses (individual named sample) in each of these three communities.

- Bulle, a city of 17'000 inhabitants situated in the South part of the canton of Fribourg and capital of a region well known for his cheese: the Gruyères region. If thirty years earlier it was clearly a smalél center of a peripheral region, it is now also a city heavily linked to Fribourg and Bern on one side, Lausanne and Geneva on the other on, making it part of a metropolitan region.

- Morges, a city of 15'000 inhabitants on the shoe of the lake of Geneva, now heavily connected to Lausanne and Geneva but keeping more middle class character than the 2 other choices;

- Vernier, more than 30'000 inhabitants, a typical suburban community near Geneva, having a very important growth in the sixties (4'000 inhabitants in 1950, 8'000 in 1960 and 24'000 in 1970)

The characteristics of these three communities can be summarized by figures on nationality and education, see table 1.

---

[2]An example of the structure of multiple names generator as used in the Swiss survey is given in Annex.

[3]The detailed schema is given in annex

Table 1: Characteristics of the three local communities

Swiss nationality	City	Low education	Bachelor, Master or PhD
71.4	Bulle	41.7	10.7
70.6	Morges	29.6	20.2
67.4	Vernier	38.2	17.5

Source: "Swiss Structural Survey" 2012, Swiss Statistical Office

3 Results

Three types of results will be considered here: response rates, but also differences on answers as well as overview of qualitative appreciation of the survey by interviewers.
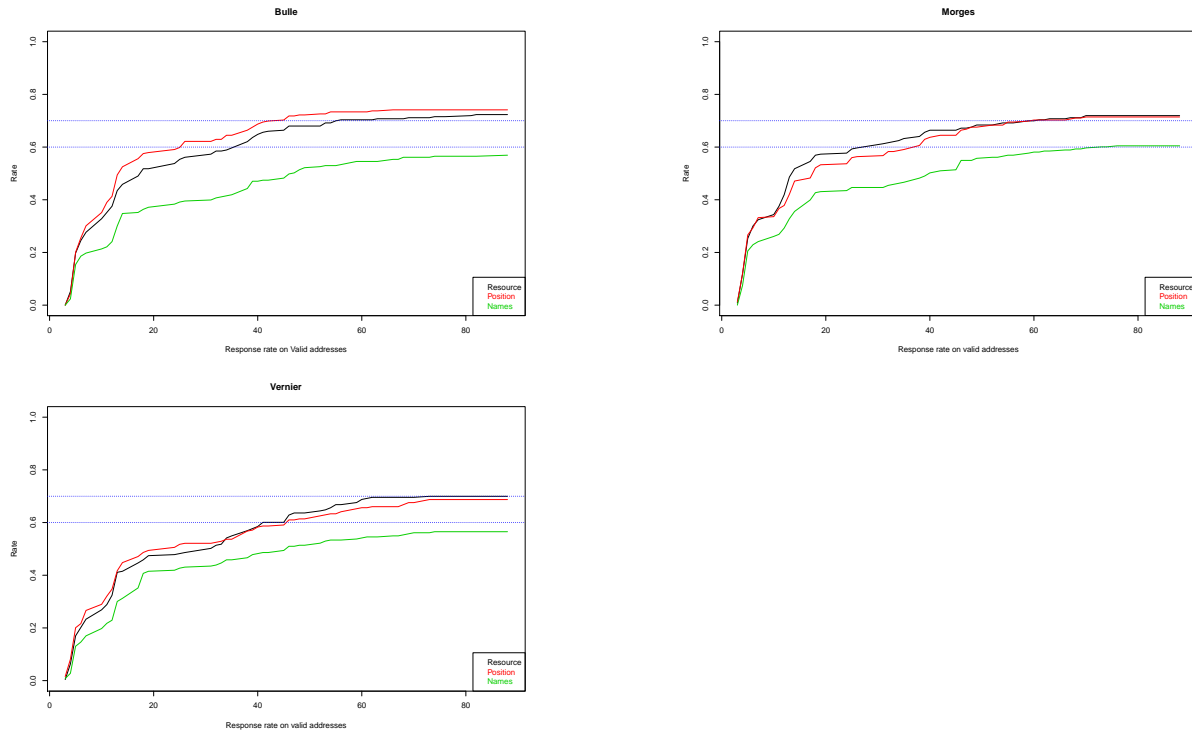
3.1 Response rate

The response rate is to be considered for the first step as well as for the second one.

3.1.1 First wave

The first question is of course the distribution of response rates according method and space. A couple of remarks here:

- In all the three localities, the final response rate is rather good. When in Switzerland, face to face survey are credited of a little more than 50% at most, we observe nearly 70% in the most "classical" forms of survey. Of course, these studies were formatted for around half an hour rather the the full hour of ESS.

- The "Name generator version" has a lower response rate from the beginning. The recall are efficient, but more or less at the same rate as for the other version. In any case, we wend up with around 10% deficit of response rate in this version, staying nevertheless well upon a 50% limit!

Figure 1: Response rate by locality and version

Generally, Morges, the most "middle class city" has had a better response rate in all situations. Bulle, more peripheral, had had a good response rate in the standard situation, a little bit lower for the most

demanding ïames generatorẅhile the suburban city of Vernier was the lowest in terms of response rate in any case.

The differences between position and resources generators are not significant but it is worth mentioning the tendency to insist more on position, a hierarchical criteria, in Bulle, the lowest locality in terms of education, and more on position in Morges, the highest one on this criteria.

3.1.2 Second wave

As mentioned, the response rate at the second wave was much lower, roughly between 40 and 50% of the first wave respondents, also because of a design limited to one recall. This was decided in order not tu put too much burden on the respondent that have already accepted to answer to the fist round. If the

Table 2: Response rate at the second wave in function of the first

Place	Resource	Position	Name
Bulle	39	42	48
found Morges	46	44	47
Vernier	39	31	48

response rate is then a little bit higher in Morges, the most middle class locality of the three, this is not the most salient point: the response rate is far higher in the cases where "Name generators" were used... and this second wave was precisely based on a "name generator". That means that probably the first wave, in this case, was also perceived as a way to familiarize itself to such a tool.

Of course, a final analysis as to take into account a composition effect of the two waves and, then, the strong differences appearing after the first phase could be strongly diminished in considering the complete process.

3.2 Substantive change: Comparison between situations

The question is of course how far the differences in response rate will induce some differences in answers. We expect a contrast between, on one side, resource and position generator, very similar in form and response rate, and, on the other side, multiple names generators.

Table 3: Responses to some questions, 3 conditions

Generator		Resource	Position	Names
Sexe	Man	46	46	46
	Woman	54	54	54
Nationality	Swiss	75	78	82
	Foreign	25	22	18
Education	Low	26	32	22
	Middle	46	43	51
	High	27	24	28
Happiness	High	41	38	37
	Middle	55	60	61
	Low	4	2	2
Health	Very good	26	27	22
	Good	55	48	54
	Fair	14	20	19
	Bad	4	4	3
	Very bad	0	1	2

In fact the results here (table 3) are not simply in line with such hypotheses. First, there is no difference between male and women in any condition, which is good. If we have less foreigners and low educated people in the third and most demanding condition, names generator, the variation are more or less as important between resource and position generators which seem to be perfectly comparable in terms of format and response rate. Perhaps we can argue that the hierarchical dimension, more present in the position generator, is seen as more important by lower educated people, the higher educated ones being more sensible to the differences in resources.

For more substantive variables, like happiness or health, there are apparently no systematic or meaningful differences. That means that the change of situation and lower response rate do not change the distribution of these variables.

3.3   Comments by interviewees

The comments given at the end by interviewees are not a systematic source of information. The same can be said of the letters explaining refusal.[4] Nevertheless, the reactions are interesting to take into account.

- A small number of potential respondent had mentioned, particularly in the case of the "Name generator" format, that such a survey was too intrusive. This was even more the case when they received the second wave questionnaire. Interestingly this was not only refusers but also people first filling the questionnaire before to mention the point when sending back the form.

- A larger number of respondents mentioned the interest of the survey, and even thanking us for that,also because they find at this occasion a time to think to their own lives and relations with others. In fact a lot of comments included very intimate remarks about life events and how the relationships can change during the life-course.

That means probably that the way a survey is presented, and the interest to answer, is probably a factor as important as the length. It is surely of prime importance, even if until know we do not have seen so many studies trying to integrate explicitly this parameter and even less trying to quantify it, for example in meta-analysis. Furthermore, this will probably have some selection effects to take into account.

Once again, this is perhaps part of the discussion between two positions: one insisting on the fact that answer to survey have to be universal and that there is no particular selection effect due to the subject of the survey or framing of question, the only limitation being errors in the cognitive process, or a second one insisting more on the social condition of production of the data and arguing that the interest for the topic surveyed has not to be the same for the different respondents according their situation in the social structure.[5] But of course, this could even lead to specific adaptation of the instruments to some social groups, a point difficult to implement without a careful set of validation.[6]

4   Discussion and conclusions

The first point to underline is the quality of the data obtained by comparison to the price. Even if the length of an self-administered questionnaire was the half of a face to face survey like the ESS, the cost was more in a ten times ratio, and this with higher response rate! Of course this is not to say that ESS has to be done by mail in Switzerland but more to say that a careful examination of the most acceptable form of survey in our societies has to be done and perhaps that self-administered forms are better adapted now to the majority of the population in a country like Switzerland.

The second point is that the load perceived by the respondents has an important impact. In the case, just the form of presentation of the name generator, nevertheless optimised after focus group meeting and a careful evaluation of a set of successive surveys, is clearly a factor affecting response rate. However, such a 10 % of decrease in response rate does not seem to change so much the social composition of the respondents and even less their substantive answers. Even if this is perfectly in line with results on the importance of response rate as quality check, this has to be explained.

The third point is that we have to investigate more the way to implement and interpret multi-steps surveys. Of course we know the general tendency of selection according social integration, at least after a long first step. But a special attention to length and implementation of long surveys in a context of mixed-mode could drives survey practitioners to use more this kind of design and to integrate it in a Total Survey Error perspective.

The forth point is the need to theorize a little bit more the relation between the interest taken to answer by the respondents and the results of the survey. Most of the literature insist on the length of the

---

[4] And they are relatively numerous in case of prepaid incentive, a number of contacted persons preferring to send back the incentive rather than to keep it

[5] It is interesting in this context to consider the following quote of Albaum and Smith in [Gid01], p. 183 about mail survey: "Since people responding to mail questionnaire tend to do so because they have stronger feelings about the subject than the non-respondents, biases results are to be expected. To measure this bias, it is necessary to contact a sample of non-respondents by other means, usually telephone interviews. This is a type of non-response validation. The non-response validation may indicate that population subgroups have not been omitted and that results may not be biased."

[6] In fact we have the same debate between adaptation and standardisation in the translation problem in comparative study.

questionnaire[7] and some thumb's rules have been developed: very often around 20-25 minutes have been seen as an optimum for self-administered surveys – and for CATI[8] – and one hour in the case of face to face.[9] However, the interest in the questionnaire is also probably an important variable. In other words, it is difficult to disentangle length and interest, even if the first is far more easier to measure than the second.[10]

And finally, the interest, even more than the length, is probably variable according the position and the characteristics of the respondents. Even in this small experiment we have seen what is a probably an apprenticeship effect to the kind of questionnaire proposed. That means in our opinion, that we have probably to give more attention to the condition of production of the data in the survey process. This could be perfectly in line with the agenda proposed in [Cou13].

5 References

References

[Bog96] Karen Bogen. "The effect of questionnaire length on response rates: a review of the litterature". In: *Proceedings of the Section on Survey Research Methods*. American Statistical Association. Alexandria: Bureau of the Census, 1996, pp. 1020–1025. URL: http://www.census.gov/srd/papers/pdf/kb9601.pdf.

[Cou13] National Research Council. *Nonresponse in social science surveys: a research agenda*. The national academies press. 00000. Washington, D.C., 2013. URL: http://www.russellsage.org/research/reports/nonresponse-survey-data-collection.

[Fis82] Claude Fisher. *To dwell among friends*. 00000. Chicago: Chicago University Press, 1982.

[Gid01] Lior Gideon. *Handbook of Survey Methodology for the Social Sciences*. Vol. 1. Springer New York, New York, NY, Jan. 101.

[Lav08] Paul J. Lavrakas. *Encyclopedia of Survey Research Methods*. 1st ed. Sage Publications, Inc, Sept. 12, 2008. 1072 pp. ISBN: 1-4129-1808-1.

[LDH07] Edith De Leeuw, Don A. Dillman, and Joop J. Hox. *International Handbook of Survey Methodology*. 1st ed. Lawrence Erlbaum Associates Inc, Nov. 30, 2007. ISBN: 0-8058-5753-2.

[Le08] Nan Lin and Bonnie Erikson (eds). *Social Capital: An International Research Program*. Vol. 1. Sept. 2008.

[RB08] Debra J. Rog and Leonard Bickman. *The SAGE Handbook of Applied Social Research Methods*. SAGE Publications, Aug. 27, 2008. 681 pp. ISBN: 978-1-4129-5031-2.

[Van05] Martin Van Der Gaag. *The Measurment of Individual Social Capital*. PhD Thesis Groningen, 2005.

---

[7]On this see [Bog96] but the is a lot of discussion on other confounding factors, for example Mangione and Van Ness, in [RB08], p. 494.

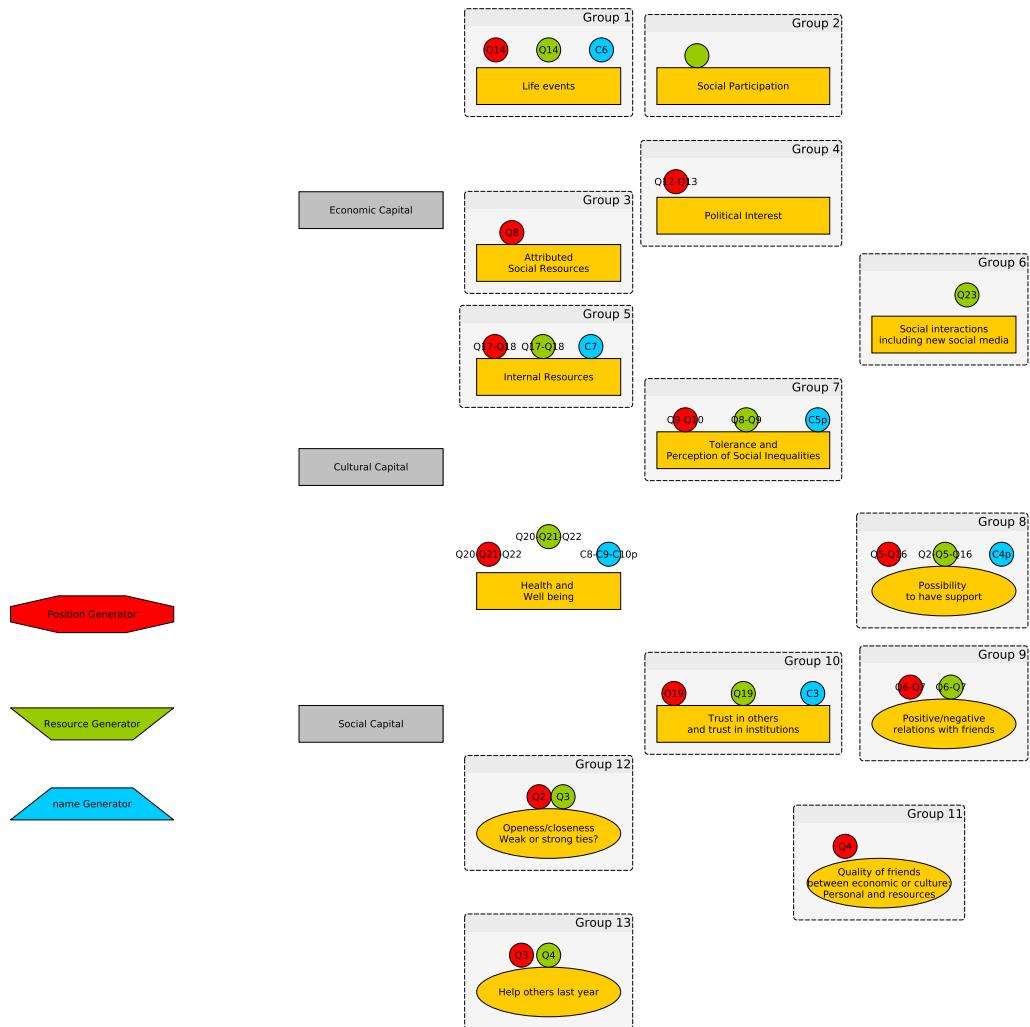[8]A Swiss experiment realised in the context of ESS experiments showed in 2007 a very different response rate according 30, 45 and 60 minutes length.

[9]See a short discussion of this by Edith deFors Leeuw, [LDH07], p. 121.

[10]The effect of burden and respondent's fatigue is for example mentioned in [Lav08]. The literature associated to these entries is nevertheless from the last millennium.

6 Annexe

6.1 Matrix design, repartition of questions between three variants

Group 1 — Q14 Q14 C6 — Life events

Group 2 — Social Participation

Group 3 — Q8 — Attributed Social Resources

Group 4 — Q12-Q13 — Political Interest

Group 5 — Q17-Q18 Q17-Q18 C7 — Internal Resources

Group 6 — Q23 — Social interactions including new social media

Group 7 — Q9-Q10 Q8-Q9 C5p — Tolerance and Perception of Social Inequalities

Economic Capital

Cultural Capital

Social Capital

Q20-Q21-Q22 / Q20-Q21-Q22 / C8-C9-C10p — Health and Well being

Group 8 — Q3-Q16 Q2-Q5-Q16 C4p — Possibility to have support

Group 9 — Q6-Q7 Q6-Q7 — Positive/negative relations with friends

Group 10 — Q19 Q19 C3 — Trust in others and trust in institutions

Group 11 — Q4 — Quality of friends between economic or culture Personal and resources

Group 12 — Q2 Q3 — Openess/closeness Weak or strong ties?

Group 13 — Q3 Q4 — Help others last year

Position Generator

Resource Generator

name Generator

6.2 Example of multiple name generator

	A1. Les personnes avec qui je discute de choses importantes :	A2. Les personnes qui me donneraient du soutien matériel ou me rendraient des petits services lors de difficultés	A3. Les personnes qui me donneraient du soutien émotionnel lors de problèmes légers :	A4. les personnes qui seraient capable de me faire changer d'avis :	A5. Les personnes qui m'énervent souvent, qui me fâchent souvent :	A6. Les personnes qui sont importantes pour moi actuellement :
Personne (1)	Pierre	Pierre	Pierre		Pierre	Pierre
Personne (2)	Marie	Marie	Marie			Marie
Personne (3)	Marc		Marc	Marc		Marc
Personne (4)	Gerald	Gerald				Gerald
Personne (5)	Maude					Maude
Personne (6)	Anna	Anna			Anna	
Personne (7)		François	François		François	François
Personne (8)		Ivan				Ivan
Personne (9)		Julie	Julie			Julie
Personne (10)			Suzanne	Suzanne		Suzanne
Personne (11)			Anna B.			Anna B.
Personne (12)				Michelle		Michelle
Personne (13)					Isabelle	
Personne (14)						Philippe
Personne (15)	—	—	—	—	—	—