

*Aleksandra Fijałkowska**

NON-RANDOM ERRORS IN OPINION POLLS

Abstract. The article deals with the issue of non-random errors in the public opinion research in the context of 2010 Presidential Elections in Poland. Remembering the fact of alarming discrepancy between exit polls' and elections' results in 2005, this problem will be reminded.

The content includes both, the theoretical core of the negligence mentioned above and practical side as well. The problem will be described in the context of different researches carried out by chosen Polish public opinion research centers and the results of the author's experiments and the research, that took place in 2007.

Key words: public opinion research, non-random errors, society manipulation, ethics.

I. INTRODUCTION

Nowadays all modern societies think they know exactly what the public opinion research means and what it is for. However, in most of cases the way of understanding this problem is wrong. People believe that such a tool give them right answers for any questions, but it is not true, unless the mechanism has been taken correctly and occurring of non-random errors has been minimized. Either on micro or macro sphere, the public opinion research is needed to characterize the whole population by using only its part. The theory of sampling gives a good basis to this process, but as very often some manipulation and invigilation have place, it is important to understand and analyze mistakes that come not from the sampling, otherwise the it won't have any substantial power. It is very important as most of results have ability to impact people's believes and opinions. For example, let's consider a survey on the psychological condition of a society. Due to the results, if some percentage of the society suffers from depression, other people might feel more comfortable then as they feel they are sharing it with others. Hence, some of them may admit they have a problem and decide to visit a psychiatrist. Of course, the best example here would be opinion polls during elections' campaigns. Every single one might be deciding as the results can also create political beliefs or change preconceptions. Therefore public opinion research instead of being a tool, which helps to discover people's opinion, can be also a good tool for societies' manipulation.

* MSc, Department of Statistical Methods, University of Lodz.

II. THE PUBLIC OPINION RESEARCH

There are various definitions, functions and uses of public opinion research¹, but generally it is a survey of public opinion from a particular sample. It is a tool, which gives opportunity to recognize opinion, beliefs, views etc. of a population or a group by researching only its representative part². This process requires understanding what “representative” means and the same, that information from such a survey isn’t sure and is burdened with mistakes, as it is dealing with the inferring from a part of the population to its whole.

The main purpose of the public opinion research is to obtain information from a society, it might be either data about consumer’s opinion, voting preferences or social attitudes and the mood of the public³. Basically, most of live aspects and thoughts can be research by this tool. By optimizing time, money and space costs, desirable information can be gained, analyzed and finally- utilized.

Sometimes society can undermine such a approach, but understanding high costs, impossibility of taking measurement from different places at the same time or interviewing the whole population in the agreed time interval, makes it the most optimal one for knowing people’s opinion about the chosen problem. Therefore, it is so important to every society to know how the public opinion research should be understood as it would be misinterpret to take it as the sure information. For the researcher is essential to use all possible tools and techniques to avoid non-random errors, while for the recipients- to comprehend this delicate line between inferences based on representative sample and the reliable knowledge.

III. NON-RANDOM ERRORS VERSUS RESPONSIBILITY AND ETHICS

Let’s define a random error as a sampling error, which means that although sampling mechanism has been used correctly, the obtained sample isn’t well-representative, the same it’s a “price” for obtaining data from a sample and infer it on the whole population⁴. The more elements in the sample, the lower error here. A non-random error is an error, which doesn’t have its source in the

¹ H. Alpert, *Public Opinion Research as Science*, The Public Opinion Quarterly, Vol. 20, No. 3, Oxford University Press, Oxford, 1956

² P. H. Rossi, J. D. Wright, A. B. Anderson, (1983), *Handbook of Survey Research: Quantitative Studies in Social Relations*, Academic Press, Orlando

³ M. Szreder, (2004), *Metody i techniki sondażowych badań opinii*, PWE, Warszawa

⁴ J. Kordos, (1987), *Dokładność danych w badaniach społecznych*, GUS, Warszawa

sampling mechanism⁵. Therefore it can be understood as every single interruption to the casual process of the survey. Higher number of elements in the sample here doesn't mean any errors reductions and often cause even more of these.

The main result of a random error is low precision, while non-random- low accuracy.

Basically in this area, only opinion polls might be verified by the real results and all other not. Hence, there is a wide field to lie and manipulate. Napoleon said "*Power is founded upon opinion*⁶" and that's the reason, why people desire to manipulate the public opinion. Politicians, media, companies etc. can present fictitious public opinion results to create own visions in various life fields. It might be done in many different ways, during the whole survey time, from the stage of its design to the presentation of the results. Table 1 contains the main non-random errors according to the research stages.

Table 1. Chosen sources of the non-random errors by the public opinion research stages

Stage	Source
<i>PLANNING</i>	incompatibility of the sampling population with the target and survey population; wrong metadata and terms; wrong methods; wrong tools; fictitious research problems
<i>SAMPLING PROCESS</i>	wrong sampling frame; wrong, incomplete or not up-to-date population register; incomplete, disrupted or incorrect sampling process; factors that are disruptive for the sampling process (i.e. prejudices and stereotypes); not-sufficient training and knowledge
<i>DATA COLLECTING</i>	problems with finding respondents and communication; omitting some elements due to different difficulties; not-sufficient interviewer's training and knowledge; bad interviewer's personal abilities; lack of professionalism; difficult to understand questions; duplications; wrong questions' interpretations; answers suggestions; questions manipulation; wrong answers due to lack of knowledge; de-coding mistakes; mistakes while populating sheets; inappropriate interface; inappropriate questions order; questions non-neutrality; advanced language level and sophisticated vocabulary; ambiguous questions
<i>PROCESSING</i>	poor measurement accuracy; data entry, coding, editing, compiling, aggregating mistakes; insufficient mathematical and statistical knowledge (i.e. wrong wages); elimination of outliers; incorrect imputation
<i>PUBLICATION/ PRESS RELEASE</i>	incorrect, wrong interpretation or over-interpretation; incorrect way of presenting results; too complicated language used in publication; untidy way of presentation; omitting important facts in presentation; wrong media used for press releases; wrong, suggestive headings

Source: Self-prepared.

⁵ M. Szreder, as above.

⁶ J. A. Manning, G. De Liancourt, (2009), *Political Aphorisms: Moral And Philosophical Thoughts Of The Emperor Napoleon (1848)*, Kessinger Publishing, Whitefish.

The researcher is able to control all of the shown above errors and the same manipulate in those areas. Hence, the ethics is so important, since people wants to know the truth about themselves and there is a potential to manipulate the society by the public opinion results.

If we look at non-random errors by the participants of the research (Table 2), we will see that researchers, coordinators and interviewers have much more impact on it rather than respondents.

Table 2. Chosen sources of the non-random errors by the participants of the research

Participants of the Public Opinion Research	Chosen sources of non-random errors in the Public Opinion Research
<i>CUSTOMERS</i>	non-ethical behaviour and demands, inadequate requirements, lack of appropriate knowledge
<i>CO-ORDINATORS, SURVEY DESIGNERS, RESEARCHERS, POLLSTERS</i>	coverage and other frame issues; measurement errors; respondent burden not minimised; wrong terms and metadata, especially in delicate social matters; lack of coherence; inappropriate question order; suggestive questions or suggestive pollster's behaviour; hard to understand questions, not clear enough sentences; insufficient sociological or/and psychological training; unfriendly behaviour or language; inadequate pollster's look while interviewing; general manipulation - intentional or not; writing down wrong, changed, transformed answers; misunderstanding of respondent's answers; taking wrong assumptions from respondent's answers; processing problems: copying and editing mistakes; lack of accuracy; aggregation mistakes; wrong inferences; inappropriate presentation; manipulation and suggestions with results presentations and inferences publication
<i>RESPONDENTS</i>	non-response aspects; memory problems; incomplete answers or omitting of some questions; psychological aspects; lying; manipulating; lack of trust and knowledge; aspects of education and intelligence

Source: Self-prepared.

Those two classifications above aren't classical, based on completeness and content problems⁷, but give good perspective where and who is the most responsible for the veracity of the public opinion results.

IV. THREE EXAMPLES

This section consist of three different examples that aim to show, that non-random errors can occur everywhere and can be caused by both human sides of the survey.

⁷ J. Kordos , as above.

In 2005 in Poland all opinion polls' results were far away from the elections' ones. Respectively Table 3 and Table 4 show legible differences between the opinion polls and the official results:

Table 3. Presidential election in Poland 2005 – opinion polls and results

Candidate	Researcher					
	CBOS	Gfk	PBS	PGB	TNS OBOP	The official results
Lech Kaczyński	46%	48%	48%	50.20%	49.10%	54.04%
Donald Tusk	54%	52%	52%	49.80%	50.90%	45.96%
Date	13-16 Oct	19-20 Oct	19-Oct	19-20 Oct	18-20 Oct	X
The discrepancy for the winning candidate	8.04%	6.04%	6.04%	3.84%	4.94%	X
The sum of differences for both candidates	16.08%	12.08%	12.08%	7.68%	9.88%	X

Source: Based on Andrzej Sadowski "Ocena zgodności wyników sondaży z rzeczywistymi wynikami wyborów. Wrzesień- październik 2005", Adam Smith Centre, Warsaw, 2005.

Table 4. Polish Parliament election 2005 – opinion polls and results for Platforma Obywatelska

Researcher	PGB	CBOS	OBOP	PBS	GfK	The official results
Date	20 – 21 Sep	14 – 18 Sep	22-Sep	21-Sep	20 – 22 Sep	X
Platforma Obywatelska	29.00%	33.00%	32.00%	32.00%	30.00%	24.14%
The discrepancy	4.86%	8.86%	7.86%	7.86%	5.86%	X

Source: Based on Andrzej Sadowski "Ocena zgodności wyników sondaży z rzeczywistymi wynikami wyborów. Wrzesień- październik 2005", Adam Smith Centre, Warsaw, 2005.

We don't know what schemes different researchers used for sampling, so it is hard to test this problem statistically here, however we can see that in both examples above, the differences were significant. What could cause those discrepancies? Obviously, the sample taken from the population could be not enough representative, or a priori information wasn't used to identify the population in a more complete way. However in this kind of survey, a couple of different reasons also can be found. First of all media, that usually in the election time are the main employers for researching companies, juggled with

words like trust, popularity, support. It can make a big confusion for the voters. Also, media can manipulate by the way how results are presented, i.e. by presenting results only for one candidate and people can't compare it to others or by avoiding presenting metadata like where, when the survey took place and whether the sample was representative or not. As well there is a problem of the number of additional questions to determine if the chosen respondent will actually vote. Moreover, in Poland, where the electorate opinion is vague, an extra attention should be put on it. Obviously, it cannot be just a sentence like "Will you be voting?", but it should be a special chain of numerous questions from which the research will deduct, if the respondent is a potential voter. Also, in the case of presidential opinion polls we can wonder if it wasn't a kind of Bradley's effect. It could be transferred to Polish ground in the meaning that media put too much pressure on the public opinion to vote Mr Tusk. The Media was claiming he would be better candidate for Poland in the context of European Union as more liberate, and people that supported Mr Kaczyński had pointed Mr Tusk in the opinion polls as it was more "appropriate" but finally voted their candidate.

The second example is a panel research with an experiment. Target population were Polish emigrants staying and working legally in Wrexham, Wales, total number of 8309 as of the end of December 2006. Random Sample Generator was used, the sample consisted of 100 elements, 1 was not found and 6 non-responses. There were two surveys taken with the following dates: 1-16 February 2007 and 6-20 August 2007. Let's call the first one as "survey A" and the other one- as "survey B". Both consisted of 17 questions covering the area of social policy of Polish and British government. In survey A all elements were ask to answers those questions straight away. In the second one, before exactly the same questions were presented, a short note made by a lawyer was read to them. The note covered the core of legal regulations in the social policy area in UK and Poland. To present differences in opinions before and after knowing basic facts about the survey topic, two question will be discussed here. Question 7 was "Does in your opinion Polish government care about expectations and needs of pregnant women, single and young mothers?" and Question 8 was "Does in your opinion British government care about expectations and needs of pregnant women, single and young mothers?" Table 5 contains the results of both surveys for the mentioned above questions.

Table 5. Comparison for questions 7 and 8 in Survey A and B

Survey	Question	"Yes"	"No"	"I don't have any opinion"
<i>A</i>	Q7	2%	65%	33%
<i>B</i>	Q7	2%	90%	8%
<i>A</i>	Q8	32%	30%	38%
<i>B</i>	Q8	77%	6%	17%

Source: Self-prepared.

For Question 7, 24 % from those, who didn't have own opinion, decided to say "No", 1% changed opinion from "I don't have any opinion" to "Yes" and 1% from "Yes" to "No". For Question 8: 24% changed answer "No" to "Yes" and 25% "I don't have any opinion" – to "Yes", no one changed opinion from "Yes" or "No" to "I don't have any opinion". This shows that inferences from Survey A don't show realistic opinion, as if so many answers migrated later from "I don't have any opinion" option to a specific one, it means in time of taking Survey A, people didn't have enough knowledge about the target topic. Thus, it is important to make sure that the target population is right and information from the research will actually have methodological power.

The last example is a small experiment of manipulation on non-representative volunteers . From the previous panel survey sample, 30 people were identified with a similar age, education and material status and their level of politics interest. Those 30 were ask to take part in this test, which was: divide group of 30 on two – Group A and Group B, and give them to populate a 10 questions questionnaire, but for Group B one of question smuggles a suggestion for Question 10. The suggestion was unpleasant cite from Polish President about Polish immigrants – UK workers in Question 8 and Question 10 in both cases was: "Do you think there is someone- a private or institution, that actually has a negative influence on the opinion about Polish work immigrants in UK?" Although it is a small, non-representative group and those 30 elements were taken from a different research, it points out on the behaviours' direction while manipulation is done. In Group A 12 on 15 people didn't have any opinion, while in Group B 8 on 15 wrote down Polish President name and only 3 didn't declare anything here. Such a hidden suggestion worked here quite visibly.

V. DISCUSSION

Non-random errors can be easily avoided, therefore more pressure and attention should be put on this aspect. Even correct process of planning, implementation and publication of the results from the Public Opinion Research

doesn't give us 100% reliable information. Hence, even small interruptions or a missed detail can be crucial, not mentioning manipulations during this statistical process or intentional changes in the results. A researcher needs to do everything correctly, using the best practise with full ethics and without fear or favour. Researches, as same as the Media should present all information from the survey accurately and pay a lot of attention on the words they use. Inferences can't be suggestive or based only on partial data from the research, as even here, on the last stage of the Public Opinion Research process it also is a non-random error. All possible extra information as dates or sample number should be provided. Every single receiver should be able to understand the results and know on what the inferences are based. Also, when using results of any Public Opinion Research, there is a need to remember about its imperfection. This applies especially to the situation when the researcher is trying to describe abstract or subjective phenomena and it is not a matter of statistical aspects, but a problem how to measure it.

REFERENCES

- Alpert H. (1956), Public Opinion Research as Science, *The Public Opinion Quarterly*, Vol. 20, No. 3, pages 493 – 500, Oxford University Press, Oxford.
- Babbie E. (2007), *Badania społeczne w praktyce*, PWN, Warszawa.
- Benton E.J., Daly J.L. (1991), A Question Order Effect in a Local Government Survey, *Public Opinion Quarterly*, Vol. 55, No. 4, pages 640 – 642, Oxford University Press, Oxford.
- Berg B.L. (1998), *Qualitative Research Methods for the Social Sciences*, Allyn and Bacon, Boston.
- Bolstein R. (1991), Comparison of the Likelihood to Vote among Preelection Poll Responents and Nonrespondents, *Public Opinion Quarterly*, Vol. 55, No. 4, page 648, Oxford University Press, Oxford.
- Bracha C. (1996), *Teoretyczne podstawy metody reprezentacyjnej*, PWN, Warszawa.
- Czapiński J. (2007), *Badania sondażowe w psychologii 2006/2007*, Uniwersytet Warszawski, Warszawa.
- Domański Cz. (1990), *Testy Statystyczne*, PWE, Warszawa.
- Dyoniziak R. (2003), *Sondaże a manipulowanie społeczeństwem*, Wydawnictwo Adam Marszałek, Toruń.
- Kordos J. (1987), *Dokładność danych w badaniach społecznych*, GUS, Warszawa.
- Manning J.A., De Liancourt G., (2009), *Political Aphorisms: Moral and Philosophical Thoughts of the Emperor Napoleon (1848)*, Kessinger Publishing, Whitefish.
- Popper K.R. (2002), *Logika odkrycia naukowego*, PWN, Warszawa.
- Rossi P.H., Wright J.D., Anderson A.B. (1983), *Handbook of Survey Research: Quantitative Studies in Social Relations*, Academic Press, Orlando.
- Schuman H., Presser S. (1981), *Questions and Answers. Experiments on Question Form, Wording and Context in Attitude Surveys*, Academic Press, New York.
- Sulek A. (1979), *Eksperyment w badaniach społecznych*, PWN, Warszawa.
- Szreder M. (2004), *Metody i techniki sondażowych badań opinii*, PWE, Warszawa.

Szreder M. (2005), Próba głosów. Nasze poglądy czy poglądy badacza: pułapki sondaży opinii publicznej, *Tygodnik Powszechny*, No. 29, 17 July 2005, page 4, Tygodnik Powszechny spółka z o.o., Kraków.

Zasępa R. (1972), *Metoda Reprezentacyjna*, PWE, Warszawa.

Aleksandra Fijałkowska

BŁĘDY NIELOSOWE W BADANIU OPINII PUBLICZNEJ

Artykuł przywołuje zagadnienie błędów nielosowych w badaniach opinii publicznej, w kontekście zbliżających się wyborów prezydenckich w Polsce. Pamiętając o rozbieżnościach pomiędzy wynikami sondaży przedwyborczych w 2005 roku w stosunku do późniejszego wyniku elekcji, temat ten warto ponownie przywołać do dyskusji.

Praca obejmuje między innymi rdzeń teoretyczny problemu błędów nielosowych, na który składają się przede wszystkim ich opis oraz konsekwencje, jakie niosą za sobą takie właśnie błędy. Autorka pragnie zauważyć też istotność kwestii rzetelności naukowej oraz odpowiedzialności etycznej, jakie spoczywają na badaczu podczas realizacji badania opinii. Problem zostanie opisany za pomocą analizy wybranych wyników badań, opublikowanych przez różne jednostki badawcze. Zostaną także zaprezentowane wyniki eksperymentów i badania Autorki, mających za zadanie ustalić do jakiego stopnia badacz może modyfikować i sterować informacjami oraz respondentami podczas realizacji badania opinii.