Analyzing the Interviewers' Evaluative Questions in Phone Polls

Alyaa Roshdy Zahran*, Aya R. Farag†, Hesham M. Aly‡

Keywords: telephone polls, evaluative questions, biplot

DOI: [10.29115/SP-2012-0025](https://doi.org/10.29115/SP-2012-0025)

Survey Practice

Vol. 5, Issue 4, 2012

Analyzing the Interviewers' Evaluative Questions in Phone Polls

INTRODUCTION

Usually national surveys include some questions that are related to the interviewer's characteristics as well as interview evaluative questions to be answered by the interviewer to reflect on the completed interview. Considerable research has been devoted to study interviewer effects (age, race, gender, experience, attitudes) (see Berk and Bernstein 1988; Groves and Magilavy 1986; Groves et al. 2004; Hill 2002; Kish 1962; Singer et al. 1983; Stokes and Yeh 1988). Few papers, however, studied the evaluation questions. In the context of telephone polls, only two studies: The Gallup Organization (1998) and Tarnai and Paxson (2005), studied the interviewer's evaluative questions. These questions could highlight the need of improvement in many directions, like choosing target population, question wording, raising awareness among specific groups in the society, etc.

Beginning in 2009, two questions were added to each poll at the Public Opinion Poll Center (POPC) at the Information and Decision Support Center (IDSC) of the Egyptian Cabinet. The first question identifies a "less than good" interview (defined in terms of some identifiable problem) from the perspective of the interviewer while the second question specifies what kind of problem was encountered. Thirty-four polls of POPC during January 2009-April 2010 are analyzed in this paper. The political and social polls suffer from the existence of high percentages of interviews with problems. In most of our polls, region is significantly associated to interview type, while in all the polls gender is significantly associated with interview type. As respondent education level increases, the interview tends to be good, whereas as the respondent gets older, the interview tends to be a less than good interview. The most reported problem is "not understanding the meaning of some questions." A biplot shows that the reported problems partition in four groups (clusters), where

* **Institution:** STAT Dept, FEPS-Cairo University

† **Institution:** Poll Center, Information and Decision Support Center, Egyptian Cabinet

‡ **Institution:** Poll Center, Information and Decision Support Center, Egyptian Cabinet

group members are positively correlated together. There is no association between poll type and the reported problems.

LESS THAN GOOD INTERVIEW AND THE REPORTED PROBLEMS

Figure 1 presents sample sizes and the proportions of interviews with problems by poll type (political, social, media, and health). The proportions are high and range from 0.075 to 0.259. On average, differences among these proportions by poll type are not significant (Kruskal–Wallis p-value=0.648).

Figure 1 Proportion of interviews with problems and sample size in each poll grouped by type of the poll.

Thirteen reasons were reported for having a less than good interview. These reasons were coded as a multiple response question. Table 1 presents the percentages of each problem within each poll. The maximum percentage is 61.5 percent while the minimum value is 0. On average, one problem is reported most often: “not understanding the meaning of some questions” with

a standard deviation of 0.15. This problem is also reported the most controlling for poll type. Within each poll type, the problems are divided into three groups: low (below 5 percent), medium (5–20 percent) and high (above 20 percent) reported problems.

Table 1 Percentage of problems identified in the polls [weighted data].

Poll Type	Poll Name	Hearing	Knows nothing about Topic	Noise existence	Doubt in answers	Reluctant	Ironical answers	Not understanding meaning of some questions	Not interested in survey subject	Others sharing answers	Responsible	Rushed	Did not want to be called again
Health	Swine Flu_1	0.11	0.27	0.30	0.22	0.10	0.11	0.28	0.15	0.03	0.03	0.03	0.01
	Swine Flu_2 -Oct09	0.06	0.09	0.23	0.30	0.11	0.08	0.54	0.09	0.11	0.01	0.07	0.01
	Swine Flu_3_Nov09	0.05	0.03	0.26	0.20	0.19	0.05	0.60	0.05	0.12	0.01	0.04	0.02
Media	Calculate it correct_1	0.01	0.29	0.21	0.34	0.04	0.02	0.25	0.06	0.06	0.03	0.03	0.01
	Calculate it correct_2	0.02	0.45	0.19	0.22	0.07	0.04	0.22	0.24	0.12	0.03	0.02	0.00
	Performance Pop Media_09	0.18	0.21	0.24	0.22	0.22	0.14	0.54	0.26	0.18	0.00	0.05	0.02
	Calculate it correct_3_Jul09	0.04	0.57	0.10	0.35	0.05	0.02	0.15	0.16	0.02	0.01	0.05	0.00
	Media Performance_Repve Health & Family Planning	0.04	0.28	0.20	0.24	0.09	0.04	0.53	0.09	0.08	0.02	0.05	0.00
Political	Calculate it correct_4_Dec09	0.03	0.48	0.14	0.33	0.11	0.04	0.21	0.17	0.09	0.01	0.06	0.02
	Television	0.05	0.15	0.24	0.35	0.29	0.11	0.28	0.20	0.28	0.01	0.11	0.02
	Eval Gov's Decisions_09	0.07	0.30	0.23	0.20	0.09	0.05	0.19	0.17	0.12	0.04	0.05	0.00
	The Gaza War	0.03	0.25	0.31	0.21	0.07	0.05	0.53	0.18	0.07	0.04	0.06	0.01
	The Credibility of the Gov	0.06	0.12	0.34	0.27	0.03	0.04	0.40	0.12	0.07	0.01	0.06	0.00
	Eval Gov's Performance	0.03	0.17	0.25	0.32	0.08	0.09	0.27	0.13	0.19	0.03	0.06	0.00
	Obama Visit_before	0.05	0.52	0.15	0.23	0.04	0.04	0.08	0.14	0.06	0.01	0.02	0.01
	Obama Visit_after	0.06	0.43	0.23	0.18	0.06	0.07	0.20	0.25	0.06	0.01	0.03	0.01
	Trends States_October	0.05	0.30	0.25	0.25	0.15	0.05	0.10	0.18	0.17	0.01	0.07	0.01

Poll Type	Poll Name	Hearing	Knows nothing about Topic	Noise existence	Doubt in answers	Reluctant	Ironic answers	Not understanding meaning of some questions	Not interested in survey subject	Others sharing answers	Respo ill	Rushed	Did not want to be called again
Social	Eval some Public Services_before match_Nov09	0.01	0.04	0.27	0.39	0.10	0.05	0.46	0.09	0.05	0.01	0.02	0.01
	Eval some Public Services_after the match	0.08	0.14	0.27	0.31	0.19	0.05	0.50	0.10	0.11	0.04	0.05	0.02
	Public Services/ Trends States_Jan 10	0.03	0.19	0.27	0.42	0.32	0.04	0.39	0.09	0.11	0.01	0.01	0.01
	Public Services/ Trends States_Jan10_2	0.05	0.08	0.30	0.34	0.31	0.00	0.41	0.09	0.16	0.00	0.01	0.00
	Evaluation of the Government's Decisions_Jan10	0.05	0.04	0.15	0.33	0.09	0.01	0.26	0.26	0.07	0.00	0.06	0.00
	Mubarak Visit to USA	0.02	0.40	0.21	0.18	0.06	0.03	0.14	0.16	0.02	0.01	0.04	0.01
	Nazif in Beit Beitk_Jul09	0.04	0.52	0.14	0.09	0.04	0.03	0.29	0.27	0.01	0.01	0.03	0.01
	Traffic Pr in Egypt_09	0.03	0.16	0.18	0.17	0.07	0.04	0.45	0.11	0.07	0.01	0.03	0.00
	Renovating Religion Speech	0.04	0.12	0.18	0.27	0.04	0.05	0.32	0.13	0.07	0.03	0.10	0.03
	Women Role in Society	0.03	0.03	0.20	0.26	0.08	0.08	0.53	0.01	0.07	0.00	0.05	0.01
	E-Government Services	0.06	0.54	0.19	0.16	0.03	0.06	0.30	0.11	0.06	0.00	0.02	0.02
	Population Problem	0.04	0.16	0.29	0.13	0.13	0.02	0.49	0.09	0.10	0.03	0.03	0.01
	What do the Egyptians read?	0.02	0.03	0.41	0.17	0.03	0.05	0.14	0.07	0.18	0.01	0.10	0.02
	Quality of Public Transport09	0.09	0.10	0.26	0.32	0.14	0.13	0.29	0.06	0.07	0.01	0.11	0.06
	The Traffic	0.06	0.07	0.29	0.47	0.16	0.04	0.50	0.10	0.02	0.05	0.06	0.00

Poll Type	Poll Name	Hearing	Knows nothing about Topic	Noise existence	Doubt in answers	Reluctant	Ironic answers	Not understanding meaning of some questions	Not interested in survey subject	Others sharing answers	Respo ill	Rushed	Did not want to be called again
	Problems in Egypt_Feb10												
	Role of Public Opinion Polls_Sep09	0.05	0.62	0.12	0.26	0.02	0.03	0.13	0.02	0.01	0.02	0.05	0.01
	Management Corruption	0.04	0.28	0.20	0.21	0.14	0.05	0.46	0.20	0.07	0.02	0.06	0.01
	Grand mean	0.05	0.25	0.23	0.26	0.11	0.05	0.34	0.13	0.09	0.02	0.05	0.01
	Standard deviation	0.03	0.18	0.07	0.09	0.08	0.03	0.15	0.07	0.06	0.01	0.03	0.01

BIPLOT AND ITS USE IN POPC DATA SET

While it is important to report the problem's grand percentage, it is of more interest to look at the underlying association structure (1) among the reported problems, (2) among the polls, or (3) between polls and reported problems. The biplot of Gabriel (1971) helps in visualizing these structures. In a biplot, a multivariate data set with n observations and m variables is represented with n data-points and m axes. The length of the axis approximates the variable variance. How data points are spread in the multi-dimension space (Euclidean distances) reflects the association structure among these points; the closer the points to each other the more association among them. The value of any observation on any variable is measured by the product of axis length and length of the perpendicular projection from the observation onto this axis. Finally, the cosine of the angle between any two axes represents approximately the correlation between the axes-variables (Kohler and Luniak 2005). The biplot is offered in the well known statistical packages (SAS, R, and STATA) and other specialized packages (e.g., GGEPlot and XLS-Biplots). However, we used the Biplot add-in-Excel-macro of Lipkovich and Smith (2002) because it is run on a user-friendly widespread platform.

For our weighted data set, the absolute measure of goodness of fit of the biplot equals 71.6 percent. Although it does not exceed the 90 percent cutoff point defined by Smith and Cornell (1993) for $m > 2$, it slightly exceeds the 70 percent cutoff point of Kohler and Luniak (2005), who emphasized that their cutoff point suffices to approximate key features of the data. Figure 2 shows the biplot of the 34-polls data. Two problems have the highest variability (longest axes): "knowledge lack" and "not understanding the meaning of some questions". This information was also given in the last row of Table 1, however, it is easily visualized in this plot. Four groups are formed:

- Group 1: Rushed, reluctant, not want to be called again, ironically answering, noise existence, doubt in respondent answers, other sharing answering the questions;
- Group 2: Not understanding the meaning of some questions, hearing problems;
- Group 3: No interest in survey topic, knowledge lack;
- Group 4: Other problems, including illness.

RESPONDENT CHARACTERISTICS AND INTERVIEW TYPE

Figure 3 Scatterplot of association measures between interview type and some respondent characteristics vs its p-value.

To quantify the association between interview type and education level (below high school, high school or equivalent, university level or above) or age group (18-less than 30, 30-less than 40, 40-less than 50, 50-less than 60, 60 and above), the gamma measure is used. As the gamma value approaches 1 (zero) in its absolute value, the association between the two variables approaches the perfect association (independence) state. A negative (positive) gamma value indicates that the two variables are negatively (positively) associated. The lower left panel of Figure 3 shows the gamma-vs p-value scatterplot using interview type and education level, while the upper right panel depicts the scatterplot using interview type and age group. Significant negative association exists between interview type and the education level, which ranges between -0.3 and -0.7. Hence, as education level increases, there is a tendency that the interview will be good. On the other hand, weak positive association exists between the interview type and age group, except for five polls. Among those five polls, only one has a relatively high gamma value (-0.138), and it is significant. For all the other polls, as the respondent gets older, the interview tends to be less than good. However, one should notice that not all the polls do have significant association between interview type and age group.

CONCLUSION

The interviewers rated large proportion of interviews in each poll as having problems. On average, this proportion does not significantly differ by poll type. In most of our polls, however, region is significantly associated to poll type, while in all the polls, gender is significantly associated with poll type. As respondent education level increases, the interview tends to be good, whereas as the respondent gets older, the interview mostly tends to be less than good interview. Raising awareness among elder or/and low educated ones would help to decrease the probability of getting a less than good interview.

The reported problems of less than good interviews are divided into three groups with regard to their occurrence percentage (low, medium, high). We should work on the high and medium groups to eliminate/reduce their occurrence. "Not understanding the meaning of some questions" is reported the most on average in both grand average and within poll type average. It is recommended to introduce an extra revision step in the process of writing the questionnaire to ease the language and/or remove any ambiguous questions. Good introduction could help in creating respondent-interest in the topic. Questionnaires of periodic polls that do cluster more around the axis of not understanding the meaning of some questions should be carefully revised.

According to the biplot, poll type does not affect the reported problems. Regarding the correlation structure among the reported reasons, four groups are distinguished from each other, where group members are positively correlated together. A pair of groups are either independent from each other or negatively correlated with each other depending on the angle between the two groups.

REFERENCES

- Berk, M.L., and A.B. Bernstein. 1988. "Interviewer Characteristics and Performance on a Complex Health Survey." *Social Science Research* 17: 239–51.
- Gabriel, K. 1971. "The Biplot Graphic Display of Matrices with Application to Principal Component Analysis." *Biometrika* 58: 453–67.
- Groves, R.M., and L.J. Magilavy. 1986. "Measuring and Explaining Interviewer Effects in Centralized Elephone Surveys." *Public Opinion Quarterly* 50: 251–66.
- Hill, M. 1986. "Race of the Interviewer and Perception of Skin Color: Evidence from the Multicity Study of Urban Inequality." *American Sociological Review* 67: 99–108.
- Kish, L. 1962. "Studies of Interviewer Variance for Attitudinal Variables." *Journal of the American Statistical Association* 57: 92–115.
- Kohler, U., and M. Luniak. 2005. "Data Inspection Using Biplots." *The STATA Journal*, 208–23.
- Lipkovich, I., and E.P. Smith. 2002. "Biplot and Singular Value Decomposition Macros for Excel." *Journal of Statistical Software* 7: 1–15.
- Singer, E., M. Frankel, and M.B. Glassman. 1983. "The Effect of Interviewer Characteristics and Expectations on Response." *The Public Opinion Quarterly* 47: 68–83.
- Smith, W.F., Jr., and J.A. Cornell. 1993. "Biplot Displays for Looking at Multiple Response Data in Mixture Experiments." *Technometrics* 35: 337–50.
- Stokes, L., and M.Y. Yeh. 1988. "Searching for Causes of Interviewer Effects in Telephone Surveys." In *Telephone Survey Methodology*, edited by R.M. Groves and et al. New York: John Wiley and Sons.
- Tarnai, J., and M.C. Paxon. 2005. "Interviewer Judgments about the Quality of Telephone Interviews." In *60th Annual Conference, American Association for Public Opinion Research*. Miami Beach, Florida.
- The Gallup Organization, Inc. 1988. "Ohio Family Health Survey Methodology Report." Part 5. 1988. <http://www2.odh.ohio.gov/Data/OFHSurv/OFHSMeth.htm>.