When students give biased responses to researchers: An exploration of traditional paper vs. computerized self-administration

Raphaëlle Butori, Béatrice Parguel

When students give biased responses to researchers:
An exploration of traditional paper vs. computerized self-administration

Abstract:
This article investigates the impact of two data collection modes (online surveys and paper-and-pencil surveys) and the perceived attractiveness of the experimenter on two types of response biases: social desirability and demand artifacts. Its results highlight the combined effect of the data collection mode and the perceived attractiveness of the experimenter on social desirability and show that signs sensitivity and signs interpretation (two types of demand artifacts) are stronger in the context of online surveys than in the context of paper-and-pencil surveys.

Key words:
Data collection, online surveys, paper-and-pencil surveys, response bias, social desirability, demand artifact

Track 8: *Marketing Research and Research Methodology*

Researchers in social sciences often use student samples for their investigations, because such respondents are easily accessible. Since the 60's, 80% of the empirical studies published in some of the most prestigious journals in social psychology (Sears, 1986) and consumer behavior, especially in *Journal of Consumer Research* or *Journal of Consumer Psychology* (Peterson, 2001), have used self-administered questionnaires on student samples. Yet, the use of students is controversial. On the one hand, students' homogeneity enables a better control of the noise created by exogenous variables, which increases the internal validity of the results (Calder, Phillips, & Tybout, 1981; Greenberg, 1987). But on the other hand, some of their characteristics might alter the quality of their responses. Specifically, their self-concept and attitudes are less precisely defined than those of adults (Sears, 1986), which may increase socially desirable responding. In addition, their higher cognitive abilities (Orne, 1962) may increase demand artifacts, i.e. "*the biases that result from the adoption of a specific role by the subject who believes that he or she has discerned the objective of the study, or provoked by a reaction to the decoding of an experimental manipulation*" (Herbert, 2005). The existence of these biases could explain why the effects observed on student populations differ in intensity and sometimes in direction from the effects observed on nonstudents (Peterson, 2001).

In this article, we study these biases in the light of the opportunities provided by the diffusion of the Internet. Whereas many researchers have highlighted that online surveys enabled to save time and effort in quantitative data collection (Couper, 2000; Cobanoglu, Warde, & Moreo, 2001; McDonald & Adam, 2003), their ability to reduce social desirability bias and demand artifacts has yet to be studied. Our first objective is to compare the impact of two self-administered data collection modes – paper-and-pencil surveys and online surveys – on these biases. Since both administration modes imply the physical (paper-and-pencil) or virtual (online) presence of an experimenter, our second objective is to explore the role of the perceived attractiveness of the experimenter on the two biases.

1. Theoretical Foundations and Hypotheses

Many studies have already compared online surveys to face-to-face interviews (Heerwegh & Loosveldt, 2008) or phone interviews (Dillman et al., 2001; Fricker, Galesic, Tourangeau, & Yan, 2005; Kreuter, Presser, & Tourangeau, 2008). These studies have reached the following conclusions. Interviewer-administered questionnaires (face-to-face and phone interviews) enable to collect higher quality responses than online questionnaires: a lower rate of "*don't know*" responses and non-responses and a higher variance in the responses (Heerwegh & Loosveldt, 2008). As Holbrook and his colleagues (2003) explain, the presence of an interviewer gives rise to non-verbal communication, enables the interviewer to ease the comprehension of the questions, and reduces respondents' distraction. However, interviewer-administered questionnaires generate more social desirability than online questionnaires. Though interesting, these studies do not compare online surveys to self-administered paper-and-pencil surveys. To make up for this gap, we propose and test an explicative framework.

1.1. The impact of data collection mode on social desirability

The concept of social desirability rests on the notions that there are social norms governing some behaviors and attitudes and that people may misrepresent themselves to appear to comply with these norms (Kreuter, Presser, & Tourangeau, 2008). In social science research, Naas, Moon, and Carney (1999) explain that social desirability is triggered by two factors: the

nature of the questions (sensible or related to personal topics) and the presence of an interviewer, who activates the existence of social norms.

As previously mentioned, there is no consensus on the impact of the collection mode (online *vs*. paper-and-pencil) on social desirability. Whereas some authors suggest that the interaction with a computer creates a feeling of intimacy and anonymity that reduces socially desirable responding (e.g., Martin & Nagao, 1989), others find contradicting results (e.g., Lautenschlager & Flaherty, 1990). According to them, respondents to a computerized survey anticipate the fact that their answers will be identified, verified and stored in a database, which increases socially desirable responding. A meta-analysis conducted by Richman, Kiesler, Weisband, and Dragow (1999) on more than 61 studies published between 1967 and 1997 reached the conclusion that it is not so much the medium (computer *vs*. paper) as the existence of moderating variables that triggers social desirability. Among these moderating variables is the presence or absence of an experimenter. Since paper-and-pencil surveys require the presence of an experimenter and online surveys do not, we propose the following hypothesis:

H$_{1a}$: Social desirability bias is stronger in paper-and-pencil surveys than in online surveys.

In the case of an online survey, the researcher is not physically present when respondents answer the questionnaire. However, he or she can signal his or her presence with the use of graphic elements. Several studies in electronic commerce (e.g., Hassanein & Head, 2005) have shown that the use of design elements such as pictures were indeed able to create an online social presence that leads to more favorable consumer attitudes. In this study, we hypothesize that this online social presence activates the existence of social norms, which increases social desirability.

H$_{1b}$: In online surveys, social desirability bias is stronger when the experimenter is visually identified (i.e. in the presence of a picture of the experimenter) than when it is not identified (i.e. in the absence of a picture of the experimenter).

1.2. The impact of data collection mode on demand artifacts

Demand artifacts consist of three biases: a *signs sensitivity* bias (the respondent pays attention to the connections that exist between the questions and tries to guess the objective of the study), a *signs interpretation* bias (the respondent actually figures out the objective of the study) and a *motivation* bias (the respondent modifies his or her responses to go either in the direction or in the opposite direction of the study's perceived objective). These three biases represent the three steps of an ordered sequence (Schwartz & Sudman, 1996; Herbert, 2005).
In this study, we hypothesize that two elements have an impact on these biases: 1) the data collection mode and 2) the interviewer attractiveness. More precisely, the data collection mode furthers (or not) an information processing that increases signs sensitivity and signs interpretation. Then, the interviewer attractiveness leads (or not) to a distortion of the responses according to the perceived objective of the study.

- *The impact of data collection mode on signs sensitivity and signs interpretation*

To understand the impact of the data collection mode on signs sensitivity and signs interpretation, we refer to the satisficing theory (Krosnick, 1991). According to this theory, optional question-answering takes a considerable amount of cognitive work: a respondent

must interpret the meaning and intent of each question, retrieve relevant information from his or her memory, integrate that information into a summary judgment, and then report that judgment taking into account the provided response alternatives. Whereas many respondents may perform these steps, other respondents might take cognitive shortcuts to reduce the required effort (and exhibit satisficing behavior). The motivation and ability of the respondent reduce the probability to use a heuristic whereas the perceived difficulty of the task increases this probability.

When answering an online survey, respondents can be distracted by the simultaneous use of other applications, as well as by the execution elements (colors, buttons to click etc.) that build up online questionnaires. Online surveys also require respondents to understand how to use the data collection software and its peripheral tools, which complicates the answering procedure and decreases respondents' motivation to thoroughly process the questions. In the light of the satisficing theory (Krosnick, 1991), the probability to use a heuristic will therefore be higher in the case of an online survey than in the case of a paper-and pencil survey. Alternatively, the probability of elaboration will be higher in the case of a paper-and-pencil survey than in the case of an online survey. We hypothesize that this higher elaboration translates into a higher signs sensitivity, which in turn translates into a higher signs interpretation.

H$_2$: Signs sensitivity bias is stronger in paper-and-pencil surveys than in online surveys.

H$_3$: Signs interpretation bias is stronger in paper-and-pencil surveys than in online surveys.

- *The impact of the experimenter attractiveness on the motivation bias*

Subjects who are sensitive to the signs of a questionnaire and who have been able to discern the objective of the research may (or not) be motivated to answer in the direction of the survey. According to the source-attractiveness model (McGuire, 1985), a message depends for its effectiveness chiefly on the familiarity, likeability and similarity of the source with the receiver of the message. We hypothesize that the relationship between the signs interpretation bias and the motivation bias too is moderated by the attractiveness of the experimenter, that is, by its perceived familiarity, similarity and likeability.

H$_4$: The motivation bias is stronger in the presence of an attractive experimenter than in the presence of an unattractive experimenter.

Figure 1 displays the conceptual framework of the study and summarizes the hypothesis tested.

Figure 1. Conceptual framework
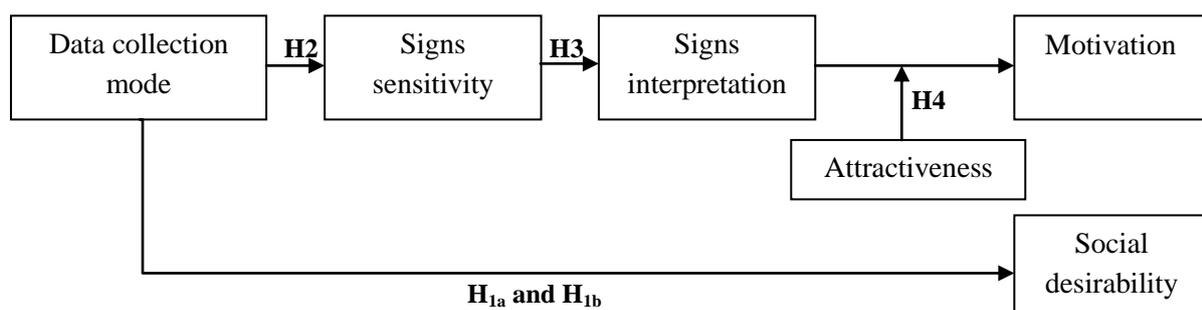
2. Methodology

Three variables were manipulated: the data collection mode (online survey *vs*. paper-and-pencil survey), the presence of an experimenter (presence *vs*. absence) and his attractiveness (attractive *vs*. unattractive). Since paper-and-pencil surveys require the presence of an experimenter, five experimental conditions (i.e. five versions of the same questionnaire) were created: 1) an online questionnaire with the picture of an attractive experimenter, 2) an online questionnaire with the picture of an unattractive experimenter, 3) an online questionnaire without any picture of the experimenter, 4) a paper-and-pencil questionnaire distributed by an attractive experimenter and 5) a paper-and-pencil questionnaire distributed by an unattractive experimenter.

To test our hypotheses, we needed a theme that triggers social desirability. We chose that of consumer's ecological sensitivity. At the beginning of the questionnaire, respondents therefore had to answer general questions about their ecological sensitivity and knowledge in terms of ecology. Then, they were asked to look at an extract of an internet site that promoted the launching of a new ecological car. Two series of questions followed. The first one evaluated the respondents' attitude toward the internet site as well as their attitude toward the car; the second one evaluated their memorization abilities (e.g., respondents were asked specific questions about the informative elements that appeared on the site, and they also had to make the list of all the elements they remembered from the site). These memorization questions were asked to provide an indirect measure of the information processing mode (elaborated or not). Finally, Herbert (2005)'s demand effects scales were inserted, as well as general questions about the survey (perceived boredom and seriousness with which the respondents answered the questionnaire) and the experimenter. The latter enabled to do a manipulation check.

In total, 232 students participated in the study. They were distributed within the five versions of the questionnaire.

3. Results

Social desirability can be measured directly, with the use of a scale (one of the most well-known scales was developed by Crowne and Marlowe (1960)), or indirectly, through an examination of the means and variances obtained on sensitive questions (Holbrook, Green, & Krosnick, 2003; Nass, Moon, & Carney, 1999). Since most of the social desirability scales are long to administer (e.g., 33 items for the Marlowe-Crowne scale), we chose the indirect method. More precisely, we assessed social desirability by running t-tests on two variables: the declared ecological sensitivity of the respondents (3 items) and the difference between their declared expertise in terms of ecology (2 items) and their actual expertise (6 items).

No direct effect of the administration mode (online *vs*. paper-and-pencil, **H$_{1a}$**) and the presence of an experimenter (**H$_{1b}$**) was identified, which supports the meta-analyses conducted on the topic. Interestingly, complementary analyses showed that in the context of an online survey, respondents were significantly more inclined to define themselves as sensitive to ecological matters when the experimenter was attractive than when the experimenter was unattractive. Though interpretable with caution (the comparison of the mean ecological sensitivity scores is only an indirect way to measure social desirability bias),

this result suggests that the experimenter's attractiveness may have an impact on social desirability in the context of online surveys.

To test **H₂**, we performed a t-test with sign sensitivity as the dependent variable. The results indicate that contrary to predictions, respondents are significantly more sensitive to the signs of the questionnaire when they answer an online survey than when they answer a paper-and-pencil survey ($M_{onliune}$ = 5.13 *vs*. $M_{paper-and-pencil}$ = 4.44 ; t = 3.75 ; p<0.001). However, this does not mean that they process information in a more extensive manner. Indeed, respondents to the paper-and-pencil survey were better at memorizing the slogan of the brand and the gas emission rate of the car (34.8% and 87.0% of correct answers for the respondents to the paper-and-pencil version of the questionnaire *vs*. 27.0% and 67.5% for the respondents to its online version) and the number of informative elements they remembered was significantly higher than the respondents to the online version of the questionnaire (M_{online} = 2.25 *vs*. $M_{paper-and-pencil}$ = 3.35 ; t = 4.39 ; p<0.001). This suggests that when students answer a paper-and-pencil survey, they focus more on the *content* of the questions than on what the questionnaire tries to show. In other words, they take care of providing responses that truly correspond to what they think and they more extensively process each piece of information. This could be because the administration context (a class or lecture room) is similar to the context in which students are evaluated. Students are therefore put in a situation that furthers an elaborate information processing.

Since respondents are more sensitive to the signs of the questionnaire in the context of an online survey than in the context of a paper-and-pencil survey, we logically find that they are also more inclined to interpret these signs (M_{online} = 4.76 *vs*. $M_{paper-and-pencil}$ = 4.37 ; t = 1.81 ; p=0.07). Similar to **H₂**, **H₃** is not supported.

Finally, no direct or indirect effect of the perceived attractiveness of the experimenter was identified on the motivation bias. **H₄** is not supported

4. Discussion

This research aimed at studying the impact of two data collection modes (online surveys *vs*. paper-and-pencil surveys) and the perceived attractiveness of the experimenter on two types of response biases: social desirability and demand artifacts. Though not expected, its results hold important implications for researchers. First of all, they highlighted the combined effect of the data collection mode and the perceived attractiveness of the experimenter on social desirability. Second of all, they showed that signs sensitivity and signs interpretation (two types of demand artifacts) were stronger in the context of online surveys than in the context of paper-and-pencil surveys.

Despite the increasing use of online surveys, and despite all their advantages (when conducted on students, online surveys enable to save time and effort in data collection, and they give access to a wider sample than paper-and-pencil surveys), researchers therefore have to be aware of the specific biases they trigger.

References

Braunsberger, K., Wybenga, H., & Gates, R. (2007). A comparison of reliability between telephone and web-based surveys. Journal of Business Research, 60 (7), 758-64.

Calder, B.J., Phillips, L.W., & Tybout, A.M. (1981). Designing research for application. Journal of Consumer Research, 8 (2), 197-207.

Cobanoglu, C., Warde, B., & Moreo, P. (2001). A comparison of mail, fax, and Web-based survey methods. International Journal of Market Research, 43 (4), 405-410.

Couper, M.P. (2000). Web surveys: A review of issues and approaches. The Public Opinion Quarterly, 64 (4), 464-494.

Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24 (4), 349-354.

Dillman, D.D., Phelps, G., Tortora, R., Swift, K., Kohrell, J., & Berck, J. (2001). Response rate and measurement differences in mixed mode survey using mail, telephone, interactive voice response and the Internet, draft paper.

Fricker, S., Galesic, M., Tourangeau, R., & Yan, T. (2005). An experimental comparison of web and telephone surveys. The Public Opinion Quarterly, 69 (3), 370-392.

Greenberg, J. (1987). The college sophomore as guinea pig: Setting the record straight. Academy of Management Review, 12 (1), 157-159.

Hassanein, K., & Head, M. (2005). The impact of infusing social presence in the web interface: An investigation across product types. International Journal of Electronic Commerce, 10 (2), 31-55.

Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus Web surveying in a high-Internet-coverage population: differences in Response Quality. *The Public Opinion Quarterly*, 72 (5), 836-846.

Herbert, M. (2005), Comportement de réponse de l'individu en situation de questionnement : le biais du répondant en comportement du consommateur, Thèse de doctorat en sciences de gestion, Université Paris-Dauphine.

Holbrook, A.L., Green, M.C., & Krosnick, J.A. (2003). Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *The Public Opinion Quarterly*, 67 (1), 79-125.

Kahle, L.R., & Homer, P.M. (1985). Physical attractiveness of the celebrity endorser: A social adaptation perspective. Journal of Consumer Research, 11 (March), 954-961.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and Web surveys: the effects of mode and question sensitivity. The Public Opinion Quarterly, 72 (5), 847-65.

Krosnick, J.A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. Applied Cognitive Psychology, 5(3), 213-236.

Martin, C., & Nagao, D. (1989). Some effects of computerized interviewing on job applicant responses. Journal of Applied Psychology, 74, 72-80.

McDonald, H., & Adam, S. (2003). A comparison of online and postal data collection methods in marketing research. Marketing Intelligence and Planning, 21(2), 85-95.

McGuire, W.J (1985). Attitude and Attitude Change. In Gardner, L., & Aronson E. (eds) *Handbook of Social Psychology*. NY: Random House.

Nass C., Moon, Y., & Carney, P. (1999). Are respondents polite to computers? Social desirability and direct responses to computers. Journal of Applied Social Psychology, 29(5), 1093-1110.

Lautenschlager, G., & Flaherty, V. (1990). Computer administration of questions: More desirable or more social desirability. Journal of Applied Psychology, 75, 310-314.

Orne, M.T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 17 (11), 776-783.

Peterson, R.A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. Journal of Consumer Research, 28 (3), 450-461.

Richman, W., Kiesler, S., Weisband, S., & Dragow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. Journal of Applied Psychology, 84, 754-775.

Schwartz, N., & Sudman, S. (1996), *Answering questions: methodology of determining cognition and communication processes in survey research*. New York: Springer-Verlag.

Sears, D.O. (1986). College sophomores in the laboratory: influences of a narrow data base on social psychology's view of human nature. Journal of Personality and Social Psychology, 51 (3), 515-530.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133 (5), 859-883.