The Theory-Testing Questionnaire Business Survey Has Passed Its Expiry Date. What Is The Alternative?

Tony Hak
RSM Erasmus University

Abstract

Response rates in business surveys are low, which implies a huge risk of selection bias. Usually no attempt is made to assess the extent of selection bias and published survey results might, therefore, not be a correct reflection of actual population characteristics. In this paper, it is argued that response rates cannot be improved to a sufficient degree and that assessment of selection bias is difficult in practice. It is concluded that academic questionnaire surveys of businesses should be abandoned and that an alternative way of testing theory in populations must be found.

Keywords: Business surveys, Nonresponse, Response rates, Selection bias, Theory-testing

Introduction

Quality standards for academic business surveys have steadily declined. A milestone in this decline was Baruch's (1999) article on response rates in academic studies. Baruch examined 175 studies which were published in five of the top leading journals in the management and behavioral sciences in three years (1975, 1985, and 1995). Most notable was a decline through the years. The average response rate in 1995 was 48.4. Top managers or representatives of organizations had a far lower response rate (36.1) than other categories of respondents. Baruch did not make a distinction between surveys of populations of persons (such as employees) in which respondents are asked to answer questions about themselves and surveys of organizations or other entities in which informants are asked to answer questions about these entities. The low response rate for "top managers or representatives" indicates a relatively lower response rate for business surveys in comparison to employee surveys.

Baruch did not discuss – not even as an aside – reasons why a high response rate is important, nor why a low response rate could be a problem. Instead of discussing the methodological implications of his findings, he turned the issue into one by which an author can demonstrate academic competence:

> "The first guideline resulting from the study is an indication of a norm for response rates in the behavioral sciences. As suggested, it seems that the norm should be within 1 SD from the average. But what average? It seems that there should be a distinction between studies directed toward top management (CEO/MD etc.) or representatives of organizations, and others such as rank and file, mid-level managers, or conventional populations. For the former the norm may then be 36 +/- 13 whereas for most other populations it may be about 60 +/- 20. Any deviation from this norm, especially downward, must be explained." (Baruch, 1999: 434).

The implication of this "guideline" is that a 30% response in a survey of CEOs is good because others do not do it better on average, and a response rate of 50 is excellent because it is 1 SD above the average. The implied criterion for the quality of a survey is not the absence of selection bias but whether the obtained response rates are not lower than the "average".

Baruch's "guideline" has been embraced by survey researchers, reviewers and editors. His article has, at the time of the writing of this paper (mid-September 2007), been cited by 91 publications in the ISI Web of Science database, most often as part of a reasoning in which obtained low response rates are represented as quite good in comparison to the average.

Published surveys, even in top journals, report steadily decreasing levels of response and do not report serious attempts to assess the extent of selection bias that might have occurred in the study. The main aim of this paper is to reverse the trend of declining standards of survey quality by applying the only criterion that matters, namely whether there is potential selection bias and that, thus, results might not be generalizable to a population. This paper consists of three parts. In the first part, it is discussed whether the situation regarding response rates is as bad as is it seems to be. Business surveys published in the most recent volume (2006) of the *Academy of Management Journal*, the top journal in management, will be analyzed. These studies have unacceptably low response rates indeed. In the second part, current approaches to addressing this problem are discussed and the conclusion is drawn that academic questionnaire business surveys cannot be sufficiently improved and, therefore, must be abandoned. Next, in the third part, an alternative to the questionnaire survey is discussed.

1. The quality of academic business surveys

This analysis is limited to one volume (volume 49, 2006) of the top journal in business research, *The Academy of Management Journal* (AMJ). The reasoning behind this is that the papers in this journal can be considered to represent the absolute top quality research in business research. Problems regarding nonresponse and potential selection bias in papers recently published in this journal might, therefore, be considered to reflect the current best practice in academic business surveys.

A total of 59 research papers were published in volume 49 of AMJ. A minority (17 out of 59) of these are questionnaire surveys. Six of the surveyed populations are populations of persons (mainly of employees), five are populations of teams or dyads of persons, and one is an analysis of a sample of data collected by Statistics Canada as part of its Workplace and Employee Survey (WES). The remaining five surveys are "business surveys" conducted by the authors of these articles. We will now first describe (in 1.1) and discuss (in 1.2) each of these five articles and, particularly, look at how the problem of potential selection bias is treated in each of them. Then we will discuss (in 1.3) whether it matters that we do not know the actual extent of selection bias in these five surveys and draw overall conclusions regarding their quality.

1.1. A description of five studies from a top journal

Study 1 (Agle et al. 2006) reports a test of a theory of the effect of strategic charismatic leadership on organizational performance in a population of (all) companies listed in the Monitor Publishing Company's *Financial 1000* and *Corporate 1000 Yellow Books*. A random sample of 500 CEOs is drawn from the population. The response in this sample is 128 of 500 (26%). The 128 respondents are compared with the nonresponding CEOs and firms on dimensions such as age, size, abnormal stock returns, capital intensity, and CEO tenure. No significant differences are found. Questionnaires are sent to the population of all (other) members of the management teams of the 128 CEOs that participate in this study. The response in this (second) population is 80%. Differences between respondents and nonrespondents in this population are not discussed. Regarding the limitations of the study which result from nonresponse, the article states: "We were unable to determine if the sample (*sic*) was skewed for CEO charisma and, thus, it is possible that our sample was overly populated with CEOs who agreed to participate in the research because they considered themselves to be charismatic".

Study 2 (Collins & Smith 2006) is a test of a theory of the effect of human resource practices on a firm's performance. Two populations of companies are selected for this test (one from the Washington DC region and another from Austin, Texas), but they are treated as one population with 397 members. The number of responses in this population is 136 (34%). Organizations that agreed to participate did not differ from nonparticipating firms in reported sales or number of employees, i.e., on known characteristics. All "core knowledge workers" in the 136 participating firms are approached. The response in this population is 61%. The paper does not contain any information or discussion of potential selection bias on the variables of interest.

Study 3 (Krishnan et al. 2006) is a test of a theory of the effect of uncertainty on the trust-performance relationship in alliances. The article states that a sample was identified of 700 dyadic international strategic alliances operating in India, but it is not explained from which population this sample was drawn (or "identified"). The response is 126 (18%). The article justifies this response rate by stating that it "is comparable to those of recent surveys of alliance managers in other emerging economies, such as 14.4 percent for China and 19 percent for Mexico". The potential for nonresponse bias was checked by comparing the characteristics of the respondents to those of the targeted "population sample" (*sic*). There were no significant differences between respondent and nonrespondent groups. A comparison was made between early and late respondents but, interestingly, *not* on the variables of interest (such as uncertainty, trust and performance). The paper does not contain any information or discussion of potential selection bias on the variables of interest.

Study 4 (Wasserman 2006) is a survey in which a theory of determinants of executive compensation is tested in a population of all companies on a list compiled by three American professional services firms. The response in this population is 20%, which is presented as "relatively high, considering the sensitivity of the questions and the level of the executives targeted". No statistically significant differences were found between the group of respondents and the group of nonrespondents on known characteristics. The paper does not contain any information or discussion of potential selection bias on the variables of interest.

Study 5 (Carson et al. 2006) reports a test of a theory of the effectiveness of types of (outsourcing) contracts in constraining opportunism in a population of all Indian companies within categories defined by three-

digit SIC codes that were considered to have a high degree of R&D outsourcing. A random sample of 2600 managers was drawn from this population. All managers in this sample were contacted in order to assess their eligibility. 1305 managers (50%) could not be contacted. 722 of the 1295 managers who were contacted appeared not to be eligible because their company was not involved in R&D outsourcing. 168 (29%) of the 573 managers that were eligible refused to participate. These eliminations left 405 eligible informants who received a questionnaire. The response in this group was 125 (31%). The last number (31%) is treated by the authors as the response rate in their study, but this number refers to only their third round of response. The response rate in their eligibility study was 50%. The participation rate that they achieved when they asked eligible managers to participate in the survey was 71%, and only 31% of the latter group returned the questionnaire. If we assume that there was no difference in eligibility between the 1305 managers that could not be contacted and the 1295 that could, then the original number of eligible managers could be estimated to have been 1150. The eventual response rate, therefore, is 11% (125 of 1150). Early and late responders within the group of 125 respondents were compared on all variables. The paper does not contain any other information or discussion of potential selection bias on the variables of interest that could have occurred in each of the three stages of this study (accessing potential respondents; the request to participate if eligible; and the questionnaire). In their discussion of the limitations of this study, the authors state that "it is a study of 125 firms, so replication is needed".

1.2. Evaluation of these five studies

All five studies are aimed at the testing of a theory. A population is chosen for this test, and either the entire population (census) or a sample from that population is selected for the test. In some cases, this process of selecting the sample (or population) for the study consists of two phases. In Study 1, for instance, a sample of 500 CEOs is approached in a first stage and, next, the other members of the management teams of the responding CEOs are approached. The response rates in these studies range from 11% (in Study 5, but not reported there) to 34% (in the first round of Study 2). These response rates are presented as "normal" or "relatively high". In all studies nonrespondents and respondents are compared on known characteristics, but *not* on the variables of interest in these studies. Early and late respondents are compared in two studies, again on known characteristics but not on the variables of interest (except in Study 5, in the third round only).

Only two articles (Study 1 and Study 5) show some evidence that the concept of selection bias (as well as the implication that no conclusion regarding the population can be drawn) is understood. When referring to the potential skewedness of the "sample", the authors of Study 1 apparently refer to the possibility that the group of respondents is different from the group of nonrespondents (which are also members of the sample) on the variable of interest.

It is remarkable that Study 5 is published in AMJ at all. First, it does not correctly compute the actually obtained response rate, which is considerably lower than the reported one (11% rather than 31%). Early and late responders are compared in the third round of selection only, and no attempt is made to compare respondents and nonrespondents in any of the three rounds. Finally, this survey is presented as "a study of 125 firms", whereas it actually is a study of a large population from which a sample of about 1150 eligible firms was drawn of which only 125 participated. This failed survey of a large sample of a large population, thus, is misleadingly presented as if it was a survey of a small population of 125 firms.

A rather depressing picture emerges, particularly when we take into account that these are five studies from the *Academy of Management Journal*, the top journal in management studies. All five business surveys published in one volume (2006) of this journal fail to achieve their aim to generate results that are generalizable to the population that is surveyed. Some of them cover up this failure by referring to other studies with equally low response rates (and, for that matter, equally low generalizability to a population). All five claim to have produced evidence of the absence of selection bias by presenting the results of a comparison between respondents and nonrespondents on known characteristics. In none of these five studies an attempt is made to assess selection bias on the variables of interest. Only Study 5, in its third round of selection, compares early and late respondents on the variables of interest.

1.3. Does it matter that the extent of selection bias is not known?

This negative assessment of the quality of the five business surveys published in volume 2006 of the AMJ is based on the fact that selection bias was not measured and, therefore, not excluded in these studies. It is assumed that selection bias has occurred because it was not shown that it did not occur. Is this reasonable? Or, in other words, does it matter that it is not known whether and, if so, to what extent selection bias has occurred in these studies?

We will answer this question by demonstrating how easy it is to obtain erroneous results in surveys with the response rates as reported in these studies. Assume we have a variable with only two values (Yes and No) and the ("true") frequencies of these values in the population are both 50%. Also assume that no data error occurs, so the true values are measured correctly in all cases. Assume also that the size of the population is 1000 and that we conduct a census, so sampling error is zero. If the response rate is 100%, we will (correctly) measure the value Yes in 500 respondents (50%) and the value No in the other 500 (50%). Now assume that there is nonresponse and that the response is "biased" or "skewed" such that population members with the value No have 20% higher inclination to nonresponse.

Table 1. Selection bias for different response rates if persons with value No have a 20% higher nonresponse rate than persons with value Yes, if true frequencies are 50% for both Yes and No.

Response rate	Found Yes	Found No	Found distribution of Yes/No
100	500	500	50/50
90	454	446	50/50
80	409	391	51/49
70	364	336	52/48
60	318	282	53/47
50	273	227	55/45
40	227	173	57/43
30	181	119	60/40
20	136	64	68/32
10	81	19	81/19

Table 1 lists the distribution of Yes and No answers within the group of respondents for different response rates. The figures in this table show that, in this hypothetical scenario, the observed Yes/No distribution begins to deviate considerably from the true distribution if the response rate decreases to values lower than 50. The five surveys published in 2006 in AMJ have response rates between 11% and 34%. The Table shows that selection bias might be considerable, even under the assumption of a quite moderate selectivity in response between persons with a No and a Yes value.

The table also demonstrates that the three approaches to nonresponse that are applied in the articles published in the 2006 volume of AMJ do not address the problem of selection bias:

(1) The fact that other researchers also have response rates of 30% does not prevent the selection bias as illustrated in the table to occur. It only suggests that these other surveys were equally liable to potential selection bias.
(2) A comparison between early and late respondents within the group of respondents (e.g., within the group of 300 respondents at the 30% response level in the table) cannot reveal how both early and late respondents are preselected (as it were) by different response inclination.
(3) Comparison of nonrespondents and respondents on known characteristics will in no way reveal a difference between these two groups on the variables of interest.

It is often argued that the potential selection bias as illustrated in Table 1 does not affect the results of theory-testing studies (e.g., Blair & Zinkhan 2006). This is correct. Even if response is seriously biased for all relevant variables in a study, this will not change their *relationship* in the group of respondents compared to the group of nonrespondents. However, it is a realistic assumption that response will be biased in terms of relationships rather than in terms of values of variables only. The authors of Study 1 above raise the possibility that their group of respondents "was overly populated with CEOs who agreed to participate in the research because they considered themselves to be charismatic". This is possible indeed. But it is not relevant because such an overrepresentation of charismatic leaders would not result in an overestimation of the effect of strategic charismatic leadership on organizational performance. These authors should, rather, have been concerned about the possibility that their group of respondents was overly populated with charismatic CEOs who have reason to believe that their style of leadership has had a positive effect on organizational performance.

Let us assume, for example, that effective charismatic CEOs have a 20% lower nonresponse rate than all other CEOs (i.e., charismatic leaders who do not consider themselves successful as well as all CEOs who are less charismatic); that there is not an actual effect of leadership on performance; and that both charismatic leadership (with values Yes and No) and successful performance (values Yes and No) have a 50/50 distribution in the population. Table 1 shows that under these assumptions a response rate of 26% (i.e., the actually obtained response rate in Study 1) would result in a distribution of about 60% successful performance and 40% non-successful performance in companies with charismatic leaders, whereas this distribution would be 50/50 in responding companies with non-charismatic leaders. A statistical test would reject the null hypothesis and the hypothesized positive relation between charismatic leadership and

organizational performance would be confirmed erroneously.
The assumption (in this example) of a 20% difference in response inclination between persons or companies with different characteristics in terms of their pattern of values on the variables of interest is realistic. Nonresponse research has shown that nonrespondents and respondents tend to differ in terms of the relations between the variables of interest. For instance, respondents who are more sensitive to specific types of advertising are usually also more inclined to take part in research on the effects of advertising, resulting in research findings that suggest that the researched types of advertising are effective, whereas this is a characteristic of the respondents rather than of the entire population that is surveyed.

The hypothetical example of a 20% higher response inclination of effective charismatic CEOs demonstrates that the combination of (a) a slightly higher response of instances (persons, companies) that possess the combination of values of the independent and dependent variables as predicted by the hypothesis and (b) response rates that currently are considered "normal" or "the best possible", can easily result in artefactual results. Thus, it *does* matter that it is not known whether and, if so, to what extent selection bias has occurred in the AMJ research papers that report response rates of 11% to 34%. It is very likely that their findings do not reflect actual relations in the populations that are surveyed. The bottom line is that the findings of these studies are suspect and that we have no way of knowing how correct they really are.

This begs the question how solutions could be found for the current situation. Two main current approaches to addressing the problem of low response rates and potential selection bias are discussed in section 2. A more radical alternative is discussed in section 3.

2. Two current approaches

A recent topic issue of the journal *Organizational Research Methods* (Volume 10, issue 2, April 2007) was devoted to "understanding and dealing with organizational survey nonresponse". Although the papers in this issue deal with nonresponse in employee surveys rather than in business surveys, they are useful for the discussion in this paper because they represent the state of the art in approaches to potential error resulting from nonresponse. The editors of this topic issue (Steven G. Rogelberg and Jeffrey M. Stanton) explicitly criticize the practice in which researchers "justify an obtained response rate on the basis of how it is consistent with industry standards or what is typically found in a given area of research. Although such descriptions do put a response rate into context, the fact that everyone else also achieves 30%, 40%, or 50% response does *not* help to demonstrate that the reported research is free from nonresponse bias" (Rogelberg & Stanton 2007: 197-198). Two types of approaches to the issue of potential selection bias are discussed in this topic issue: (a) trying to improve response rates and (b) understanding the magnitude and direction of bias caused by nonresponse.

2.1. Attempts to improve response rates

Research on determinants of nonresponse reported in the April 2007 topic issue of *Organizational Research Methods* is only the latest in a long tradition of nonresponse research in which it is shown that some incentives and some approaches (in a tailored design) sometimes have and sometimes do not have a positive effect on response rates. These effects might be quite large in size. Rose et al. (2007), for instance, obtained a response rate of 43.5% in a sample with a monetary incentive compared to a response of 31.7% without the incentive. However, our increasing knowledge of the potential positive effect of specific incentives and approaches on response rates has failed as yet to stop the pace of decline in response rates.

Trying to improve response rates, though important in itself, is not a promising approach to the problem of selection bias, for two main reasons:

(1) Even if response rates increase dramatically in relative terms, such as 34% (from 31.7% to 43.5%) in the Rose et al. (2007) study, the absolute level of response is still quite low.
(2) It is not known how such methods with positive effects on the response rate affect selection bias. They might even introduce new selection bias if, for instance, sensitivity to a monetary incentive is (positively or negatively) correlated to conceiving yourself as a successful charismatic leader.

Against this background, Rogelberg & Stanton (2007) advocate another strategy, namely that researchers "conduct a nonresponse bias impact assessment, *regardless* of how high a response rate is achieved" (2007: 199). This is sound methodological advice, which should be heeded whenever survey results are published.

2.2. Nonresponse bias impact assessment

Rogelberg & Stanton (2007) provide a comprehensive overview of methods that can be used to assess the extent of the selection bias that actually has occurred in a survey. These techniques are not remedies for actual selection bias, but they could be seen as remedies for methodological uncertainty if the result of such an

assessment is that no relevant selection bias has occurred in a survey. Rogelberg & Stanton (2007) discuss nine different techniques. None of these techniques is conclusive, but results of several of them together might provide a basis for an estimation of the selection bias that might have occurred in a survey.

Two of these techniques are the ones that were also used in the five business surveys published in the 2006 volume of AMJ; (a) "archival analysis" (a comparison between nonrespondents and respondents in terms of known characteristics) and (b) "wave analysis" (a comparison between early and late respondents). Archival analysis might result in information about differences in response rate between different types of respondents but it does not give any information at all about selection bias, i.e., whether the pattern of values of the variables of interest in the group of respondents differs from the one in the group of nonrespondents. Wave analysis is a more relevant technique because it compares groups in terms of such patterns, but in each specific survey it is not known to what extent late respondents differ from nonrespondents.

A third technique, labeled "familiar" by Rogelberg & Stanton (2007) is "follow-up", a comparison between a sample from the group of nonrespondents with the respondent group. This is the only approach which is aimed at a direct assessment of selection bias (i.e., of differences between respondents and nonrespondents). Although this technique is "familiar" indeed, it is hardly ever practiced in actual surveys and it is, e.g., not reported in the five studies published in 2006 in AMJ. It is normally applied only in methodological studies in which the survey is the object of study rather than a research strategy used to obtain scores from members of a population.

Five of the other six techniques for nonresponse bias impact assessment discussed by Rogelberg & Stanton (2007) can be characterized as estimation techniques in which different types of evidence (about respondents or about nonrespondents) are used to derive a judgment about the likelihood that selection bias of a relevant size might have occurred. The ninth technique is replicating the survey. In a replication study a sample is selected from the same population in a different way or the study is conducted under different conditions.

An important feature that all these techniques (except archival analysis and wave analysis) share is that they are rarely put into practice. This is understandable if we think about the implications for the researcher of the results of a serious nonresponse bias impact assessment. The chance that evidence for the existence of selection bias of a relevant size will be found is far from zero, and what should one do if selection bias appears to have occurred? Honest reporting of such evidence will considerably decrease the likelihood that the results of the survey will be published in a journal. Because of the immense risks for the researcher, it is clearly wise to *not* conduct any serious nonresponse bias impact assessment. Much less risk is involved in justifying "an obtained response rate on the basis of how it is consistent with industry standards or what is typically found in a given area of research".

The conclusion of this overview of current approaches aimed at (a) increasing response rates and at (b) estimating the extent to which selection bias actually has occurred is that they offer no solution. Response rates can be improved but not to sufficiently high levels. Moreover, currently available techniques for nonresponse bias impact assessment are either feasible but inconclusive (archival analysis and wave analysis) or less feasible as well as risky for the researcher (and therefore avoided in practice). The business survey as we know it, thus, is in a deep crisis, because its results cannot be trusted and there are currently no methods available for improving their trustworthiness. The best methodological advice regarding business surveys is, therefore, to *not* conduct this type of survey any more. Fortunately there is an alternative, which will be discussed now.

3. An alternative to the questionnaire survey

A team of researchers wanted to test a model that links managerial and employee turnover in a business unit with its performance. The model states that high turnover results in lower efficiency and that lower efficiency results in lower performance. A population must be selected for this test. The model was assumed to be applicable to any type of business unit with managerial and employee turnover. The research team decided that a population of fast food restaurants would be an excellent choice for the test. They got support from the general management of Burger King and sent a questionnaire to the managers of all corporately owned Burger King restaurants in the continental United States. The response rate of 50% was quite good compared to the response rates typically found in this area of research, presumably due to the support of Burger King general management for the research. However, there was some anecdotal evidence that managers that had successfully dealt with high levels of employee turnover had responded in larger (relative) numbers than managers who had been less successful in dealing with equally high levels of employee turnover. The lack of support for the model found in this study might, therefore, be due to an

overrepresentation of "good" managers who had been able to avoid the effects of turnover that the model correctly predicts. There was also anecdotal evidence that managers had not taken the effort required for providing correct answers to a number of questions. It seemed they had not looked into their records to find correct turnover and efficiency data but rather had estimated these values. In sum, this was a survey plagued by nonresponse error and measurement error.

This survey of corporately owned Burger King restaurants was actually conducted, but in a very different manner. Kacmar et al. (2006) collected data from all 262 corporately owned Burger King restaurants in the continental United States for which complete data over a two-year period were available in the POS system installed in each restaurant. The POS system feeds all local data to a central database that is routed to the BKC restaurant support center. Kacmar et al. (2006) did not send questionnaires to local managers but retrieved data directly from the POS system. They had zero nonresponse, and the data were as good as they were when recorded in the POS system by local managers. This study, published in the 2006 volume of AMJ, was not listed at the beginning of this paper as one of the business surveys in that volume because it does not present itself as a "survey".

Kacmar et al. do not provide any label for the research strategy in this study. They refer to the study in generic terms, such as "this study" or "the present study", only. If we would ask coders to categorize the research strategy of papers published in the 2006 volume of AMJ in categories such case study, experiment, and survey, it would be quite hard for them to find the appropriate category for this study. It would certainly not count as an experiment or case study, but coders would also hesitate to categorize it as a survey because it misses the features that make a survey recognizably a survey (such as sections discussing questionnaire construction or evaluating the obtained response rate).

Interestingly, most of the established textbooks on survey research do not provide a definition of the survey. Dillman (2000), e.g., apparently assumes that it is evident what a survey is and that we know it when we see it. Fowler (2002:1-2) does not give a definition either but states that his book focuses "on those surveys that have the following characteristics". These characteristics are:

- The purpose of the survey is to produce statistics.
- The main way of collecting information is by asking people questions.
- Generally, information is collected about only a fraction of the population.

Fowler does not discuss which types of surveys do not have these characteristics and are, therefore, excluded by his focus on only surveys that have them.

Of the standard texts on the survey, only Groves et al. (2004: 2) define a survey, namely as "a systematic method for gathering information from (a sample of) entities for the purposes of constructing quantitative descriptors of the attributes of the larger population of which the entities are members". This formulation is (linguistically) constructed as a definition of a "method of gathering information" (namely as a "systematic" one), but the "system" of this method of data collection is not specified and the only specification that this definition offers concerns the *aims* of data collection ("for the purposes of constructing quantitative descriptors of the attributes of the larger population"). Kacmar et al.'s (2006) study of the population of corporately owned Burger King restaurants in the continental United States fits this definition.

Although the definition of Groves et al. does not specify the method for gathering information, their book discusses only data collection by means of questionnaires (in various modes). This raises the question how we could call a research strategy of which the purpose is to produce statistics for a population but does not use questionnaires as a means of data collection, e.g. if tax filings are used as a source of information in a survey of companies, or if "interviewers" in a survey of housing conditions visit households and count the number of rooms for each household without asking questions to the occupants.

These examples point to the necessity to treat the purpose of a study (such as the aim of constructing quantitative descriptors of the attributes of the larger population) and its method of data collection (such as a standardized questionnaire) as independent aspects of a study. The question is then whether a survey should be defined by its aims or by its data collection method. The proposal of this paper is to define the survey by its purpose (as is actually done by Grover et al.) and to adopt the following definition suggested by Dul & Hak (2008: 289): "A survey is a study in which (a) a single population in the real life context is selected, and (b) scores obtained from this population are analysed in a quantitative manner". Using this definition, a study such as Kacmar et al.'s (2006) *is* a survey, but without its methodological problems! If this study (in which data were retrieved from the company's information system) is compared to a corresponding questionnaire survey (in which questionnaires are sent to local Burger King managers, resulting in potential selection bias and data error), it is immediately clear which is the preferred one.

Groves et al (2004: 67) state that surveys "describe or make inferences to well-defined populations" but do not discuss how such a well-defined population should be selected for a study. The aim of describing characteristics of a specific, well-defined population (such as the population of corporately owned Burger King restaurants in the continental United States) is considered to be a given. The aim of academic survey research, however, is rarely to describe characteristics of a specific population but rather to test a proposition regarding a theoretical domain (such as a proposition regarding business units with personnel turnover) for which one population from the domain (such as the population of corporately owned Burger King restaurants in the continental United States) is chosen for the test. If the proposition is shown to be true for the population of corporately owned Burger King restaurants in the continental United States, this is not treated as information about Burger King restaurants but rather about the predictive value (for performance) of managerial and employee turnover.

Because academic research is not primarily aimed at discovering facts about Burger King restaurants but rather at enhancing the robustness of propositions, the outcome of a single survey can never be decisive. The test of the proposition needs to be replicated in other populations and, therefore, next populations need to be selected for such tests. Because it is not the case that all populations in the theoretical domain have the same theoretical importance, in particular if a proposition has already been tested in some populations, the selection of a population for a replication survey is not an entirely arbitrary matter. From this perspective, it is quite astonishing that neither textbooks on survey methods nor, for that matter, general texts on methodology discuss criteria and methods of selecting a population for a survey.

However, replication is very rare in business research and currently most published studies are first tests of a theory. The selection of a population for such a first test of a theory is arbitrary in a sense. The only important restriction is that the selected population should be a member of the theoretical domain. If we again take the example of testing a model of the effects of employee turnover on unit performance, this model could be tested in any population of business units that show variation in employee turnover as well as in unit performance. This means that a population can be selected of which one knows beforehand that good data will be available in an information system. The preferred procedure of selecting a population for the test, then, would not consist of (a) selecting a population, such as Burger King, and then (b) making decisions on how data will be collected, but rather a reverse order, consisting of (a) discovering for which candidate populations complete and good data are available in, e.g., information systems and then (b) selecting one of them for the test.

Conclusions

The theory-testing questionnaire business survey has passed its expiry date. Response rates have reached such low levels that significant selection bias is to be expected. Effective techniques of assessing the extent of selection bias in a survey are not available and, if they were, it is not clear how researchers should deal with the results of such assessments. Selection bias, however, can be avoided by collecting data from other sources than respondents.

References

Agle, B.R., et al. (2006), "Does CEO charisma matter?**",** *Academy of Management Journal*, 49, 161-174.

Baruch, Y. (1999), "Response rates in academic studies", *Human Relations*, 52, 421-438.

Blair, E., & G.M. Zinkhan (2006), "Nonresponse and generalizability in academic research", *Journal of the Aacademy of Marketing Science*, 34, 4-7.

Collins, C.J., & K.G. Smith (2006), "Knowledge exchange and combination", *Academy of Management Journal*, 49, 544-560.

Carson, S.J., et al. (2006), "Uncertainty, opportunism, and governance", *Academy of Management Journal*, 49, 1058-1077.

Dillman, D.A., (2000), *Mail and internet surveys. The tailored design method* (2nd Edition), Wiley.

Dul, J., & T. Hak (2008), *Case study methodology in business research*, Butterworth-Heinemann.

Fowler, F.J., (2002), *Survey research methods* (3rd Edition), Sage.

Groves, R.M., et al. (2004), *Survey methodology*, Wiley.

Kacmar, K.M., et al. (2006), "Sure everyone can be replaced … but at what cost?", *Academy of Management Journal*, 49, 133-144.

Krishnan, R., et al. (2006), "When does trust matter to alliance performance?", *Academy of Management Journal*, 49, 894-917.

Rogelberg, S.G., & J.M. Stanton (2007), "Introduction: understanding and dealing with organizational survey nonresponse", *Organizational Research Methods*, 10, 195-209.

Rose, D.S., et al. (2007), "A penny for your thoughts", *Organizational Research Methods*, 10, 225-240.

Wasserman, N., (2006), "Stewards, agents, and the founder discount", *Academy of Management Journal*, 49, 960-976.