Seizing the Opportunity: Using Availability Samples in Policy Programs for Creating Relevance in Broader Contexts

Lien Beck[1], Jill Coene[1], Friso den Hertog[2], and Tinne Lommelen[1,2]
[1]Hasselt University, KIZOK, Hasselt, Belgium
[2]Maastricht University, UNU-MERIT, Maastricht, The Netherlands
lien.beck@uhasselt.be

Abstract: There is a growing interest within research in management studies to regard 'organizational practice' as a highly relevant field of empirical observation. Planned change projects turn out to be promising platforms for theorizing about practice, theory testing, and the innovation and improvement of organizational practices. However, natural contexts can also seriously limit the possibility to transfer the outcomes to broader contexts. Both external and construct validity are evidently a matter of serious concern in this kind of research. Field researchers are placed for difficult assessment of the strengths and weaknesses of their data set. What is the richness and added value of the data, what are the flaws that limit their value for other contexts, and what can be done to reduce the potential threats to external and construct validity? The present paper offers a practical roadmap that leads field researchers through a number of basic design considerations. The steps in the roadmap are illustrated by examples from a large-scale strategic development program for SMEs in the Euregion Meuse-and-Rhine.

Keywords: field research, external validity, induction, statistical generalization, theoretical generalization

1. Introduction

There is a growing concern about organizational practice as a field of research. 'Outdoor research' offers the promise to get close to practice and practitioners. In this way it offers the opportunity to contribute to the improvement and innovation of practices (Romme and Damen 2007, Van Aken 2004), and theorizing about practice (Jarzabkowski and Spee 2009). Furthermore, this research might be regarded as one of the key answers to the urgent call to make research in the fields of management and organization more relevant for both practice and practitioners (Hambrick 1994, Lawler et al. 1985, Van Aken 2004). The encouragement is a two-way effort: not only researchers are stimulated to regard practice as a field of empirical study, but practitioners are stimulated also to reflect on their work, and carry out inquiries in their own fields of practice (Jarzabkowski and Spee 2009). The key motive in both perspectives is to make local knowledge useful for other contexts (Van Aken 2004).

In this study we discuss a specific strand of 'outdoor research'. We refer to the strategy in which researchers nestle their research in 'naturally occurring' change and policy programs, or organizational experiments (Lawler 1977). This strategy offers a broad window of opportunities for doing research that is both relevant for theory as for practice (Lawler et al. 1985). The full range of social science methods can be applied in this 'outdoor research': rigorous quantitative and qualitative observation, and a mixture of both. However, we can not conceal that these kinds of settings also carry serious limitations for academic research since the stakeholders in these change processes determine to a large extent the degrees of freedom for the researchers. This can seriously limit the possibilities to create the necessary conditions for solid research. The limitations of samples originating from change and policy programs concern both external validity and construct validity. External validity is threatened because actors that participate in these programs have mostly passed a process of selection (Denrell and Kovács 2008) or self-selection (Harris 2001), which can result in a selection bias (Heckman 1979). Problems regarding construct validity can arise because researchers might be strongly restricted in their choice of measurement instruments, and consequently, can give less certain guarantees that the instruments used are indeed informing us about the attribute they are supposed to measure.

The samples originating from policy programs are often opportunity (or convenience) samples. As such, researchers that want to use these samples are in a dilemma. Do they give priority to the added value and richness of their material or do they give priority to the methodological expectations hold by

Reference this paper as: Beck, L, Coene, J. den Hertog, F and Tinne Lommelen, T. "Seizing the Opportunity: Using Availability Samples in Policy Programs for Creating Relevance in Broader Contexts "*The Electronic Journal of Business Research Methods Volume 8 Issue 2 2010 (pp.146-155), available online at www.ejbrm.com*

their peers? A key challenge for researchers working in this area is to make it plausible to their peers, that their conclusions are valid or informative for wider contexts (Thompson and Perry 2004).

There is a huge literature on sampling, external validity, and construct validity (e.g., Campbell and Fiske 1959, Cronbach and Meehl 1955, Denrell and Kovács 2008, Heckman 1979). However, there still remains a wide gap between the abstract and fragmented views from these studies, and the complexity of a dynamic field work setting. Researchers that want to make use of data collected by means of a policy program are in need of an overview that can be used to make an iterative assessment about what is desirable, and what is possible with their data. As such, the aim of this study is to offer a practical roadmap that guides field researchers through a sequence of design considerations to evaluate their data and to make a justification of the use of this data in academic research. This roadmap consists of three sequential steps: (1) a first assessment of the data, (2) a choice about the contribution, and (3) a plan of action to upgrade the data. In this vein, this paper aims to contribute to the methodological literature about '*doing* research outdoors' (Bryman 1988). The paper is based on experiences in a large-scale strategic development program ('Strategic Innovation') in which more than 700 Small and Medium-sized Enterprises (SMEs) participated.

2. Policy program 'Strategic Innovation': An overview

To illustrate the roadmap, we will use examples of the policy program 'Strategic Innovation'. 'Strategic Innovation' was launched in the Euregion Meuse-and-Rhine by regional governments in cooperation with the Interreg Fund of the European Commission, and aimed to support SMEs in that region to strengthen their innovation capacity. The program was targeted towards 650 firms. The design of the intervention program was based on a pilot project in which 14 SMEs participated. The developed method is an integration of well-known intervention methods, like search conference, survey feedback, cognitive mapping, SWOT analysis, and action planning. The main active ingredient of the method was a serie of two small group sessions with staff, and in some cases external stakeholders (suppliers or customers). The positive reactions from the entrepreneurs on the pilot program made the regional policymakers decide for a large-scale rollout of the program. The large-scale implied the outsourcing of the field work to business/management consultants of private firms. The outsourcing of the field work had the consequence that the program team had to find ways to instruct, monitor, and coach the consultants. The decision was made to follow two additional routes to comply with these different goals. First, protocols were developed to meticulously instruct the consultants how to carry out the interventions. Second, an effective information system was set up, serving three functions: as learning tool in the workshops, as quality management indicator of the work of the consultants, and the building of a database for scientific research. The program has been completed according to plan in July 2008. By now the regional governments in The Netherlands and Belgium have decided for two new follow-up projects involving respectively 300 and 250 SMEs. This decision opens the possibility to use these samples for replication. The research strategy was to build an extensive database fed with qualitative and quantitative information, both by the consultants, and the participants: reports from interviews and meetings, surveys, SWOT-analyses, action plans and so on. All texts have been embedded in software for handling qualitative data (NVivo) and quantitative data (SPSS).

3. The roadmap

Typical mainstream research follows a sequence of steps (Kumar, 2005): First, a problem statement is proposed and formulated. Next, the research design is conceptualized. These two steps are based on a literature review. After that, instruments for data collection are constructed and a sample is selected and collected. Then, the research method is set out and data is processed and tested. Finally, results are presented. Many textbooks are available on how to progress through these steps (e.g., Creswell 2003, Kumar 2005). However, in data that is the result of a policy program, this rarely is the case. The stakeholders of the program determine to a large extend the data collection. Nonetheless, these data can provide rich information, since they are often collected in a meaningful context. Therefore, it would be regretful to sweep them away on the grounds that they do not follow the mainstream research logic.

As such, this study introduces a roadmap to evaluate data collected through a policy program. This roadmap will guide researchers and contains three sequential steps. The first step is making an assessment about the potential strengths and weaknesses of the data, both regarding robustness and meaning. The researcher need to take advantage of the data's strengths and must try to eliminate the data's weaknesses. The second step comprehends making a choice regarding the contribution of the study to theory and practice. What is the researcher's objective? This choice has implications for the

third step of the roadmap, which involves making a plan of actions to upgrade the data, and reducing the threats to external and construct validity. When the researcher has followed these three steps, he/she has made a thorough evaluation of the data, and as such, he/she can make a justification of the use of the data in academic research. Below, each step will be explained and illustrated when necessary with examples of the policy program 'Strategic Innovation'.

3.1 STEP 1: Assessment of the data set

In mainstream research the research design can predominantly be considered as a derivative from the interests and aims of the researchers. That means that the *aims* of the research, either regarding theory building, theory testing or practical application define the point of departure. The subsequent design of the study is meant to determine what kind of data would fit those aims. However, in the context of policy programs, discussed in this paper, the choices for researchers tend to be far more limited. Either data are already available, or the possibilities for data gathering are strongly restricted by influential stakeholders in the program. The consequence is that the *opportunities* offered by the data constitute the point of departure in this kind of research. The researcher has to find out what the strong and weak points in the data set are, how the possibilities of the data set can be further enlarged, and how the external validity can be strengthened. The design of the study cannot be formulated in a linear way, but requires iteration between the possibilities of the data set and wishes of the researcher. This assessment must be carried out with a critical eye and vivid imagination, because in 'natural' projects, data sets do not only have hidden flaws, but also hidden opportunities. The following checklist can be used as a guide through this first assessment process of the data. As we already mentioned, the answers on these questions are influenced by the researcher's wishes and objectives about the data set:

- Do the data offer the opportunity to provide new insights?
- What are the qualities of the data in terms of validity and reliability?
- What are the qualities (advantages and shortcomings) of the sample(s)?
- What are the options for repairing, completing or upgrading the data?

Taking 'Strategic Innovation' as an example, we made the following assessment. The data set of 'Strategic Innovation' indicated that the data offered rich opportunities for cross-sectional research into the interrelationships of a broad range of variables of high academic interests, such as: innovation climate, strategic conversation, organizational characteristics and customer interaction (e.g., Slater and Mhor 2006; Deshpandé and Farley 2004, Liedtka and Rosenblum 1996). Second, methodological strengths and weaknesses were analysed in this process. The strengths of the data set regarded the *broad variety* of qualitative and quantitative sources. Data is collected through surveys, and reports from interviews and group sessions. Data is also collected from *more than one source*, making them more reliable (e.g., surveys are completed by different individuals within the company). A possible shortcoming of the sample was that the acquisition procedures in the program could create a *(self-) selection bias*. Also *construct validity* was considered as a serious problem in the data set, because no validated procedures were available, or could made to fit into the intervention program. Furthermore, the 'Strategic Innovation' data set has some options to repair, complete or upgrade the data. Because the identity of the firms is known, there are possibilities to add relevant data of *secondary sources*. For example, there is the possibility to add objective and financial data. At last, the decision of policy-makers to continue the program for new groups of firms, opened the chance to *replicate* the analyses.

3.2 STEP 2: Basic choices

While the assessment of the data set advances and the researchers become well informed about the potentials and the flaws of their data, they become ready to make a basic choice on the perspective of their inquiry. The perspectives are based on different strategies the researcher can choose to pursue generalizability. The first perspective is deductive and aims to inform us about social groups and categories by making inferences from samples from those groups and categories. It is the perspective of statistical generalization. The second perspective also follows deductive logic but uses theory as a path to generalization (Blair and Zinkhan 2006). Theories serve to produce propositions that can be tested by making empirical observations. Once the theory is tested, it can be applied to new situations. One refers in this respect to theoretical generalization (Yin 1989). The third perspective is mainly based upon induction. Observations from the field are interpreted, coded and integrated in order to construct new conceptual frames that serve to understand, or change that field. This

perspective refers to theory-building. The choice of the perspective also influences the criteria by which the sample needs to be evaluated (step 1). When the researcher for example wishes to make statements about the population, it is required that the sample is random, whether when the researcher wishes to build theory, sample criteria are different (e.g., broad variety of cases).

3.2.1 Statistical generalization

Statistical generalization is a first perspective researchers can choose. In statistical generalization inferences are made about a universe (or 'population') based on empirical data about a sample from that universe. Researchers working along this line have statistical procedures and formulas available that enable them to make judgements about the level of confidence regarding the generalizations. These confidence levels depend mostly on the size and internal variations, both in the universe and in the sample (Yin 1989). This perspective fits in the missions of disciplines that are focussed to inform about characteristics and trends in populations, like labour economics, election research, and epidemiology. The 'sacred' principle of this generalization strategy is the use of randomized samples, meaning that chance determines which members of the universe are assigned to a sample (Cook and Campbell 1979). Randomization purports to eliminate all possible rival explanations of final effects ('rival hypotheses') "without specifying whatever they are" (Yin 1989 p. 8; Yin 2003 in Sørensen, Mattsson, and Sundbo 2010).

However, in many areas of social research this principle of using randomized samples is difficult to realize. In the case of opportunity (or availability, convenience) samples the data-selection is driven by circumstances or by decisions made in the past. The result can be that certain types of respondents can be either over- or underrepresented (Winship and Mare 1992). In that case the risk exists that the sample can give a seriously distorted picture of the universe. At that stage the sample has to be regarded as a non-randomized sample. The implication is that the certainty that the control of the infinitive number of unspecified biases (or: 'rival hypotheses') is lost. As such, researchers that want to make inferences about a population based on a sample created through a policy program are forced to take measures to reduce the threats of external validity. This can be done by a systematic search for potential biases, and by taking corrective action. Economists have built statistical selection bias models to detect potential bias, test their impact and to adjust the data to the requirements of their statistical analyses (Heckman 1979, Kalton 1983, Winship and Mare 1992). In order to detect the bias, researchers first need to analyse the process by which respondents have been selected, or are 'self-selected'. In the case of 'Strategic Innovation', participating firms are selected in two ways: directly on a one-by-one basis by external consultants, who approached their local networks, and collectively, by representative organizations (such as professional associations). A thing to bear in mind is whether these different approaching strategies cause any differences. One should be aware that the statistical modelling strategies serve to 'restore' the image of the universe that is based on the sample. When researchers pursue other aims in their analysis, for example analysing co-variance, other strategies might be more effective to reduce the threat to external validity.

3.2.2 Theoretical generalization

The second perspective researchers can choose for their inquiry is theoretical generalization. Researchers taking the perspective of theoretical generalization are not predominantly interested in the characteristics of the universe, rather than in the interrelationship between variables. They are searching for *co-variance,* and not for means, and other characteristics of the universe. The basic strategy is the *falsification of rival hypotheses* (Cook and Campbell 1979). The point of departure for this perspective is a previous developed theory. Empirical observations are used to test the theory, by the acceptance or rejection of propositions that are derived from that theory. Before propositions can be accepted, researchers have to inspect whether other variables might have been responsible for the established effect of co-variance. These 'rival hypotheses' might have an origin in the way theoretical constructs are measured, or might be the result of the way the sample is built up. The researchers have the responsibility to trace possible error, and to disarm these threats. The critical question deals with the propensity that the established relationship (co-variance) can be rightfully attributed to the dependent variable, and not to rival explanations.

With regard to sample requirements, literature suggests that academic research tends to study phenomena that are largely resistant to sampling bias (Blair and Zinkhan 2006). Relationships between variables remain relatively accurate even if the sample is disproportionate (Blair and Zinkhan 2006). The basic condition is that the sample is diverse enough. As such, opportunity samples or

samples coming from policy program can be used when researchers choose for theoretical generalization. However, it is neither wise to use too narrow opportunity samples, because they tend to eliminate extraneous variation. Thereby, it is still necessary to search for potential biases in opportunity samples, and this starts in the same way as in the case of statistical generalization. Once these biases are identified, the analysis takes a different turn. The essential test regards the way these biases moderate the relationships of the variables under study, and not the differences between the characteristics of the sample and the universe.

3.2.3 Theory-building and -testing by inductive strategies

A third perspective researchers might choose for their inquiry is theory building. Empirical research data can be used for theory-building by following an inductive route. Researchers can use amply structured and systematic procedures for that purpose. One might place the *grounded theory* procedures in that category (Strauss and Corbin 1990). This approach leads the researcher through a sequence of coding steps, in which one iterates continuously between observations (data) and the emerging theory. In a similar vein, Eisenhardt (1989) offers a roadmap for theory-building from case studies. Both approaches are confined to theory development. In theory-building, generalization to larger populations is not part of the game. For the followers of grounded theory, inductive analysis is indeed the end of the game, because their aim is to specify the conditions concerning a specific phenomenon. Statements about new situations require new observations (Strauss and Corbin 1990). Eisenhardt (1989) is more positivistic and regards the empiric testing as the logical continuation of the scientific process. Induction helps to produce theory, deductive logic serves testing of that theory.

Theory-builders use mostly *purposeful* or *theoretical sampling* procedures to gather their data. These samples are primarily selected because of their conceptual informativeness, that is a high potential to yield rich information (Miles and Huberman 1994). Policy programs can offer rich opportunities to learn about specific phenomena which are closely related to either the target group, the content or the context of the program. In this vein the data from 'Strategic Innovation' have been used to theorize about large-scale organizational change (Sluismans et al. 2008). In that sense, samples originating from policy programs may be also considered as purposeful samples, and hence, can be useful for theory-building purposes. However, one should be aware that researchers can make also mistakes in selecting samples for their theorizing. Selection bias can occur when the researcher, consciously or unconsciously, selects cases that will confirm already existing theories, conjunctures or prejudices. At last, it should be underlined that the rich data gathered by explorative, and qualitative strategies can also play an important role in the *testing* of theoretical inferences. Yin (1989) suggests that previous developed theory can be used as a template with which to compare the empirical results of a case study in order to establish *analytical generalization.* In the inductive mode, qualitative data can be a powerful tool in the hunt for potential 'rival hypotheses', both regarding external and construct validity in later quantitative analyses. In a deductive mode, they can strengthen the researcher's claims for the validity of the conclusions when mutual confirmation of results can be established. In this *triangulation strategy* hypothesized relationships are tested by a variety of measures (Neuendorf 2002). The basic idea is that the strengths and weaknesses of the various methods tend to balance out, and when these methods lead to similar outcomes they can strongly support specific hypotheses (Neuendorf 2002).

3.3 STEP 3: Disarming validity threats: practical approaches

The two previous steps of the roadmap determine the actions the researchers need to undertake to upgrade the data set which originates from policy programs. The assessment of the data set is meant to work out as a SWOT-analysis, helping to define the opportunities the researcher can catch, and threats (s)he has to avoid by using the data's strengths and compensating their weaknesses. The actions the researchers need to take also depend on the perspective they choose for their inquiry. There are four general strategies that can be followed to upgrade the data set: (1) additional data gathering, (2) data conversion, (3) external control, and (4) internal control. In this paragraph we illustrate how these approaches can contribute to the validity of studies, using 'Strategic Innovation' as an example. Figure 1 provides examples of each strategy to upgrade the data set.

3.3.1 Additional data gathering

Several strategies exist to collect additional data for a given data set. A first strategy encompasses that extra information is accumulated for the same respondents (or other units of analysis). This is

possible by turning to secondary data (e.g., objective financial measures) or by doing a follow-up research (e.g. asking additional questions). This strategy however only works when the researcher knows the identity of the respondents (such that information can be linked together). A second strategy is to replicate the analysis on a different set of respondents. Do certain findings hold true in this new sample? A disadvantage here is that this is very time and resource consuming. A final strategy to gather additional data, is to compare the data with data of a comparable group of respondents. This group does not receive a 'treatment' and as such can provide insight of the existence of certain phenomena without a policy intervention. In this sense, it is a control group.

Regarding the used example 'Strategic Innovation', additional data has been gathered in a number of ways. First, a multiple case study among the 14 participating firms of the first project has been carried out. Second, two new program releases made it possible to replicate the analyses with different SMEs, and to add a small number of additional questions. Yin (1989) regards replication as a strategy to establish analytical generalization. Third, a survey among a local federation of entrepreneurs, meant to interest them in participation in the program was used to create a control group of 109 non-participants. Furthermore, a series of follow-up interviews with entrepreneurs has been started. At last, the intention is to link the data from the Belgian SMEs with objective financial data that are available in the database of the National Bank. These additional data make it possible to test 'rival hypotheses' concerning the sample (see examples a and b in Figure 1), and to strengthen construct validity. The follow-up interviews, for example will hopefully indicate whether the action plans made by the firms have actually been enacted. Furthermore, the linking with objective financial data will evidently strengthen the validity of the performance measures used in the questionnaires, which are consequently self-reported data.

3.3.2 Data conversion

Data conversion has been borrowed from the information sciences and refers to the conversion of the format of the data. In social sciences most usual is the conversion from qualitative (or 'text') data into quantitative data, in other words: 'content analysis'. In the program 'Strategic Innovation' there was a huge amount of text material available: interview reports, group session reports, SWOT-analyses, and action plans. This material has been coded according to standard procedures (Neuendorf 2002, Strauss and Corbin 1990). After the conversion, this data could be linked to the survey data, and ensures construct validity of part of the measures (see examples c and d in Figure 1). However, data can also be converted in the opposite direction. This can be useful when researchers are confronted with unexpected statistical relationships (or: lack of these relationships) between variables in their quantitative analysis. Going back to qualitative material in the files of individual cases can help to theorize about the explanation of such findings. The use of software for qualitative analysis (e.g., NVivo) can be instrumental in this respect. Regarding the used example 'Strategic Innovation', the reports of the company (in which the firm's outline, past strategic changes, SWOT analysis and action plan to advance the company's operations is reported) were used to analyse the social responsible behaviour of these firms (Arnoldi 2009).

3.3.3 External control

External control is the approach in which a comparison with external data is used to hunt for plausible 'rival hypotheses'. Questions that need to be answered in this respect are: do the findings hold up in other samples, are the findings congruent with previous research, do the findings fit with what literature states? One of the most followed approaches in this respect is to compare the composition of the sample with other representations of the universe, in order to see whether certain subgroups might be either over- or underrepresented. In the program 'Strategic Innovation' a control group of non-participants (see 3.3.1) was created, which could be used to replicate survey outcomes, and to check potential self-selection bias (see example a in Figure 1). One can also search for other similar data sets and samples and compare for example demographic characteristics (see examples e and f in Figure 1).

3.3.4 Internal control

Internal control also serves to hunt down potential threats to validity ('rival hypotheses'), and to disarm them by secondary analysis. Internal control regards the data which is available *within* the given data. Concerning the policy program 'Strategic Innovation', the survey data was used to check for self-selection bias (see examples g and h in Figure 1). In a similar way the moderating effect of firm size

on established associations between variables could be disaffirmed (see examples i and j in Figure 1). Furthermore, relating data from different sources can help to strengthen construct validity, as is shown in 3.3.2 (see example d in Figure 1).

'STRATEGIC INNOVATION' SAMPLES USED IN THE FOLLOWING EXAMPLES:

Sample 1: Sample of the policy program in the Euregion Meuse-Rhine (2004-2008): N=650 SMEs; without missing values N=486 SMEs.
Sample 2: A replication sample of the policy program in the province of Limburg (The Netherlands) and the Flemish part of Belgium (2008-2011): N=198 SMEs
Sample 3: A control sample of non-participants: N=109 SMEs. These SMEs filled in the same questionnaire as the SMEs in sample 1 en 2, but did not receive any 'treatment').

APPROACH 1: ADDITIONAL DATA GATHERING

The hypothesis is that SMEs that participate in the project are not different than the ones that decide not to participate. A rival hypothesis could be that the innovation climate of the participating SMEs is biased, thus that they are more innovation-minded. The survey of 'Strategic Innovation' contains a number of items that inquire about the firm's attitude regarding innovation, referred to as 'innovation climate'.
Example a. We compared the mean innovation climate score in Sample 1 with Sample 3 (control group). We found that the mean scores in both groups did not significantly differ.

The confirmation of a relationship in several samples is another test to check for 'rival hypotheses'.
Example b*:* The hypothesis is that there is a relationship between the extent of cross-functional cooperation within the firm (referred to as 'horizontal connectivity') and the degree strategic issues are alive in the discourses of management and staff (referred to as 'strategic conversation'). A regression analysis on Sample 1 (without missing values) reveals that 'horizontal connectivity' shows a linear positive relationship to 'strategic conversation'. R^2 is 0.33 and unstandardized beta is 0.69, with a significance level of 0.001. This association holds in Sample 3 (control group) (R^2 = 0.33, beta = 0.56, significance of 0.000) as well as in Sample 2 (the replication sample), which can be said to be a sample of firms that participated in a post-crisis era (R^2 = 0.37, beta = 0.68, significance of 0.000). This 'robustness check' makes rival hypotheses unlikely.

APPROACH 2: DATA CONVERSION

Qualitative data can be coupled to quantitative data resulting in rich stories.
Example c*:* An explorative study has been carried out regarding the Social Responsibility Policies of SMEs (Arnoldi, 2009). Qualitative analysis software (NVivo) was used to identify firms that carried out initiatives in this field. Next, a statistical analysis of the quantitative data served to draw the organizational profiles of these firms.

The route can also be the other way around: from quantitative to qualitative data.
Example d*:* Company X in Sample 2 scores high on proactive market orientated behaviour measured by the survey (3.6 on a score of 5). We can draw the same conclusions when we take a look at the report of the SWOT-analysis. One of the strengths of company X is: "we work closely with our clients and try to think along with them to recognize solutions or needs the clients did not yet think of". Company Y provides another example: it scores very low on strategic conversation measured by the survey (2.1 on a score of 5). Zooming in on their reported weaknesses during the SWOT-analysis corroborates this: one of the weaknesses of company Y is: "we have no clear strategy for the future".

APPROACH 3: EXTERNAL CONTROL

Concerning over- or underrepresentation of certain subgroups, a comparison can be made with the universe.
Example e*:* In our program we used data of 1317 SMEs from the Amadeus business register from the region Meuse-and-Rhine to compare the sample profile. This comparison indicated for example that micro-firms, and wholesale and retail firms were more strongly represented in our samples (1 and 2), while medium-sized firms were underrepresented.

A comparison can also be made with samples of research papers focusing on the same particular subgroups.

Example f: Sample 2 was used to investigate market orientation within family firms (Beck et al. 2010). For this reason the sample was compared with the samples of two other family firm studies in the same region. This revealed that the 'Strategic Innovation' sample was similar in age and size. Only regarding sector, there were small differences. The family firms in the 'Strategic Innovation' sample tended to be less active in services, retail, and wholesale. To limit possible biases, sector, age and size are used as control variables in the regression analysis.

APPROACH 4: INTERNAL CONTROL

A rival explanation that could be formulated for the data from 'Strategic Innovation', regards the problem of sample (self-) selection. The participating SMEs might be more strategy- and innovation-oriented than their non-participating peers. The survey of 'Strategic Innovation' contains a number of items that enquire about the degree strategic issues are alive in the discourses of management and staff (referred to as 'strategic conversation').

Example g*:* The first control regards the rival hypothesis that it is likely that SMEs with a high score on strategic conversation, are more inclined to participate because of their preference for strategic debate. However, in Sample 1 (without missing values) this variable follows a normal distribution (both low and high scores occur), with a mean of 2.96 (standard deviation = 0.60). Hence, mean strategic conversation is relatively low.

Example h*:* Another solution to check this bias is to control the motivation to participate, which is asked only in Sample 2, the replication sample. The conclusion is that motives to participate are diverse.

Furthermore, it could be that a third variable has an influence on the observed relationship.

Example i: Assume that size has an effect on the established association (for example larger organizations cooperate more in order to know what is going on in other units, or larger organizations are 'growing too big' such that cooperation decreases). In Sample 1 (without missing values) Pearson correlation between size and horizontal connectivity is however only -0.19, and between size and 'strategic conversation' r = 0.004.

Example j*:* Finally, controlling for an interaction effect of size and 'horizontal connectivity' in regressions in Sample 1 (without missing values) and Sample 2 respectively, results in a non-significant interaction term, denoting that when controlled for size, 'horizontal connectivity' is positively related to "strategic conversation".

Figure 1: Falsification of rival hypotheses

4. Conclusion

The awakening interest among researchers for organizational practice and practitioners must be welcomed as an answer to the serious call for making research more relevant both to theory and practice (Bell et al. 2006, Lawler et al. 1985, Vermeulen 2007). Mainstream researchers should no longer ignore the real world of practitioners in organizations as a field of study. This paper strongly supports the urgent call to researchers in the field of business studies to get off our veranda's to get a good deal closer to the actual work (Johnson et al. 2003, Van Aken 2004). However, once researchers decide to leave their veranda's (Johnson et al. 2003) and to nestle their research in a natural environment of planned change, they are confronted with a double challenge: (1) to make optimal use of the rich opportunities, and (2) taking the right measures to attack the many methodological problems that can arise from 'outdoor research'. We argue that these problem are not to be played down, but have to be tackled openly. The threat to validity is one of these problems, that has to be addressed openly. In this effort research should not be discouraged by the excess of limitations and warnings from the methodological mainstream literature. In this paper it is suggested to start from the opportunities offered by the available data and continue to systematically build a case for the relevance of the results for wider contexts. However, the basic condition for making trustful

interferences from these data is a ceaseless effort to falsify rival hypotheses. That is our game. The contribution of this paper is to draw a roadmap that guides field researchers through this difficult process of matching what is desirable and possible. Based on the choice they make concerning the research perspective (statistical generalization, theoretical generalization or theory-building and -testing), this roadmap can be a tool to evaluate their data. It starts with an assessment of potential weaknesses and strengths of the data, regarding robustness and meaning. This can be done by taking into consideration 4 basic questions: (1) do the data offer the opportunity to provide new insights? (2) what are the qualities of the data in terms of validity and reliability? (3) what are the qualities (advantages and shortcomings) of the sample(s)? And (4) what are the options for repairing, completing or upgrading the data? The next step in the roadmap involves making a basic choice regarding the contribution of the study to theory and practice. What is the aim the researcher strives for? Statistical generalization, theoretical generalization or theory building/theory testing? Each of these aims poses different requirements regarding validity. When the goal is to build theory for example, one could use a purposeful sample, while statistical generalization requires the data to be generalizable towards a wider context. Finally, one needs to take actions to upgrade the data in accordance with the chosen research strategy, in order to reduce threats to external and construct validity. One should bear in mind that actions such as additional data gathering, external and internal control serve mostly the external validity, while data conversion helps in improving construct validity.

However, this paper only touched an instrumental aspect of 'outdoor research'. The real challenge is the 'practice-turn' itself: the awareness that in many fields of management and organization, practice is ahead of theory. Phenomena, methods, and strategies are invented or discovered in the field (Daft and Levin 1990, Lawler et al. 1985). Researchers should be ready to capture the relevance of these inventions and discoveries and make them available for broader contexts. This means in a technical sense, that researchers have to advance their skills and tools to deal with the threats they can expect to encounter. Methodologists might support those working out there in the field by improving the skills and tools that can be used to deal with the threats of 'outdoor research'.

References

Arnoldi, E. (2009) "Corporate social responsibility (CSR) in small-and medium sized enterprises (SMEs): A content analysis", Maastricht University.

Beck, L., Debruyne, M., Janssens, W. and Lommelen, T. (2010) "A study of the relationships between generation, market orientation and innovation performance in family firms", Hasselt University.

Bell, J., Den Ouden, B., and Ziggers, G.W. (2006) "Dynamics of cooperation: At the brink of irrelevance", Journal of Management Studies, Vol 43, No.7, pp1607-1619.

Blair, E., and Zinkhan, G.M. (2006) "Non-response and generalizability in academic research", Journal of the Academy of Marketing Science, Vol 34, No. 1, pp4-7.

Bryman, A. (1988) Doing research in organizations, Routledge, London.

Cook, T.D., and Campbell, D.T. (1979) Quasi-experimentation: Design and analysis issues for field settings, Rand McNally, Chicago.

Campbell, D., and Fiske, D. (1959) "Convergent and discriminant validation by the multitrait-multi matrix method", Psychological Bulletin, Vol 56, pp81-105.

Creswell, J.W. (2003) Research design: Qualitative, quantitative, and mixed methods approaches, Sage, Thousand Oaks.

Cronbach, L.J., and Meehl, P.E. (1955) "construct validity in psychological tests", Psychological Bulletin, Vol 52, pp.281-302.

Daft, R.L., and Lewin, A. (1990) "where are the theories for the "new" organizational forms? An editorial essay", Organization Science, Vol 4, No.4, pp.i-vi.

Denrell, J., and Kovács, B. (2008) "Selective sampling of empirical settings in organization studies", Administrative Science Quarterly, Vol 53, No.1, pp.109-144.

Deshpandé, R., and Farley J.U. (2004) "Organizational culture, market orientation, innovativeness, and firm performance: An international research odyssey", Research in Marketing, Vol 21, No.1, pp.3-22.

Eisenhardt, K.M. (1989) "Building theory from case study research", Academy of Management Review, Vol 14, No.4, pp.532-550.

Hambrick, D.C. (1994) "1993 presidential address: What if the academy actually mattered?", Academy of Management Review, Vol 19, pp.11-16.

Harris, H. (2001) "Content analysis of secondary data: A study of courage in managerial decision-making", Journal of Business Ethics, Vol 34, No.3/4, pp.191-208.

Heckman, J.J. (1979) "Sampling selection bias as a specification error", Economica, Vol 47, No.1, pp.153-161.

Jarzabkowski, P., and Spee, A.P. (2009) "Strategy-as-Practice: A review and future directions of the field", International Journal of Management Reviews, Vol 11, No.1, pp.69-95.

Johnson, G., Melin, L., and Whittington, R. (2003) "Guest editors introduction: Micro strategy and strategizing: Towards an activity-based view", Journal of Management Studies, Vol 40, pp.3-22.

Kalton, G. (1983) Compensating for missing survey data. Michigan: Survey Research Centre, Institute for Social Research, University of Michigan.
Kumar, R. (2005) Research methodology: A step-by-step guide for beginners, Sage, Thousand Oaks.
Lawler, E.E. III, Mohrman, A.M. jr., Ledford, G.E., Cummings, Th., and Associates (1985) Doing research that is useful for theory and practice, Jossey-Bass, San Francisco.
Lawler, E.E. III (1977) "Adaptive experiments: An approach to organizational behaviour research", Academy of Management Review, Vol2, No.2, pp.576-585.
Liedtka, J.M. and Rosenblum, J.W. (1996) "Shaping conversations: Making strategy, managing change", California Management Review, Vol 39, No.1, pp.141-157.
Miles, M.B., and Huberman, M.A. (1994) Qualitative analysis, Sage, Thousand Oaks.
Neuendorf , K.A. (2002) Content analysis guidebook, Sage, Thousand Oaks.
Romme, A.G.L., and Damen, I.C.M. (2007) "Toward science-based design in organization", Development Journal of Applied Behavioral Science, Vol 43, No.1, pp.108-121.
Slater, S.F., and Mhor, J.J. (2006) "Successful development and commercialization of technological innovation: Insights based on strategy type", Journal of Product Innovation Management, Vol 23, No1, pp.26-33.
Sluismans, R., Den Hertog, J.F., Kunst, P., and Lommelen, T. (2008) "Readying the unready, a small business initiative a ground for theory", Maastricht University.
Sørensen F., Mattsson J., Sundbo J. (2010) "Experimental research in innovation research", Research Policy, Vol 39, pp.313-322.
Strauss, A., and Corbin, J. (1990) Basics of qualitative Research, grounded theory procedures and techniques, Sage, Newbury Park.
Thompson, F., and Perry, C. (2004) "Generalising results of an action research project in one work place to other situations", European Journal of Marketing, Vol 38, No.3/4, pp.401-417.
Van Aken, J.E. (2004) "Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules", Journal of Management Studies, Vol 42, No.2, pp.219-246.
Vermeulen, F. (2007) "'I shall not be insignificant': Adding a second loop to matter more", Academy of Management Journal, Vol 48, No.6, pp.978-982.
Winship, C., and Mare, R.D. (1992) "Models for sample selection bias", Annual Review of Sociology, Vol 18, pp.327-350.
Yin, R.K. (1989) Case study research, design and methods, Sage, Newbury Park.