Big Conclusions with Small N's?

Case Study as a Method in Evaluation Research

Petri Virtanen

ABSTRACT

If the research focus might appear to be small in terms of quantitative numbers, it does not necessarily mean that the research questions are automatically "small scale". The case study approach has gained a highly popular status in the social sciences, and in evaluation research in particular, during recent years. The case study method has been applied extensively also in the programme and project evaluations of EU structural funding. The debate on the case study method has concerned the issue of the transferability and generalizability of the result findings. However, the case study method is a systematic process that calls for specific competence in those researchers applying the method. There are a number of possibilities to enhance the trustworthiness of the case study approach. These possibilities - such as the use of the meta-evaluation approach - are discussed in the latter part of the article.

Key words: Case study, method, generalization, evaluation research

1. INTRODUCTION

In 1947 Knut Pipping published his doctoral thesis entitled Kompaniet som samhälle - laktagelser i ett finskt frontförband 1941-1944 (Pipping 1947/1978). In his study, Pipping described a machine-gun company during the Continuation War in various concrete situations: on the march, in battle, during episodes of subdued trench warfare, at rest behind the front, and when being transported by rail. Pipping's work has been considered to be, among other things, a "clear transition phase work in the modernization of [Finnish] sociology" and its value was not recognized outside a small academic circle until years later, particularly after the publication of Väinö Linna's Tuntematon sotilas (Unknown Soldier) (Eskola 1992, 247-249). Pipping's work, however, offers an eye-opening perspective into the battles and life of the 12th Infantry Regiment's Second Machine-Gun Company during the Continuation War. Pipping analyzes in detail the composition of the company, its casualties, operating practices and social-group formation in a way that probably could not have been done by any other research method. Pipping is a contemplator of contexts, a collector and a recorder of data - and, above all, a soldier among other soldiers.

Pipping (ibid., 251-252) concludes that the social structure of the company is a complex texture, the sum of different roles. Applying the concept later introduced by Pierre Bourdieu, it could be said that the habitus of the men of the company was constructed from a total of six different group roles: value group, military group, local group, age group, home district group and "eating buddy group". At the end of his doctoral thesis Pipping (1978, 255) states: "I hope that my study has shown that the above-mentioned holds true for the 2nd Machine-Gun Company/12th Infantry Regiment and probably for the whole army." Later Pipping wrote in the introduction of the Finnish translation of his doctoral thesis, published in 1978, (ibid., 18-19): "I could not say anything absolute about the representativeness of my study when Kompaniet was published in 1947. Since then, however, a number of books have been published about life on the front and many of these support the view I had that general conclusions about the army as a whole can be drawn from my study." (italics - PV).

Pipping's work is made more convincing by positions expressed later that emphasize the advantages of the case study method. For

example, Robert Donmoyer (2000) highlights three advantages of applying the case study method. Firstly, with the aid a case study, one can access fields that cannot be probed using other research strategies (the accessibility criterion). This idea of accessing the research subject is also well illustrated via another Finnish example. Risto Alapuro's study Suomen synty paikallisena ilmiönä 1890-1933 (The birth of Finland as a local phenomenon) (Alapuro 1995) analyzes politically the birth of modern Finland, the entry of major forces on the political scene. Alapuro's research task was to understand the origin and organization of the social class and stratum division within the reference frame of the parish of Huittinen. Alapuro writes (ibid., 315): "By examining a single community it is possible to obtain much information about these matters that extends beyond its boundaries. (...) Only in a compact environment can one study concretely how people are linked as well as set against each other. (...) One can uncover the mechanisms that regulate the links and oppositions of different groups and as well as the fundamental nature of organization and collective action in Finland. (...) The levels of support for different parties, ideological-political statements or other forms of political expression do not tell us much about these."

The second factor supporting the case study approach is, according to Donmoyer, that in a research project implemented using case study principles the readers of the research report get to see the research subject "through the researcher's eyes" (the "seeing through the researcher's eyes" criterion). This, of course, also sets tough demands for the writing of the research report. The writing must take into account the requirement to record the research subject authentically. The third advantage of the case study approach is, according to Donmoyer, that the report of the case study process rarely encounters harsh opposition, for the very reason that the research project has been implemented close to the research subject, meaning that research conducted in this way is difficult to condemn out of hand as incomprehensible.

What then is the nature of the case study, particularly from the perspective of evaluation research? It is above all a matter of a holistic approach (see e.g. Weiss 1998, 261). But what is the nature of this holistic approach? And what is the nature of the case study? And what value can the case study approach have at best in evaluation research? How is a case study conducted and how are its results reported? These questions will be the focus of discussion in this article. Means of strengthening the credibility of case study analysis will be presented at the end of the article.[1]

2. CASE STUDY APPROACH AS A DATA ACQUISITION METHOD IN EVALUATION RESEARCH

In the "magnum opus" works of evaluation research - such as, for example, Guba and Lincoln (1989) and Patton (1987 and 1997) - qualitative methods of evaluation research are not presented in detail. Very often, as far as qualitative evaluation research is concerned, it appears enough to state in terms of the data acquisition method that the material is collected using a case study approach or applying a case study method. Moreover, outside the evaluation research method literature - thus relating more generally to social and administration science research method literature - the case study is considered as one "basic alternative" when seeking solutions to research problems (see e.g. de Vaus 1996, 6-7). In the area of evaluation research, the case study approach is not accepted unreservedly: some respond to it disparagingly (e.g. Fischer 1995), while some others speak highly of it (e.g. Yin 1993 and 1994; Patton 1987).

Michael Q. Patton (ibid., 19) says that the case study approach is more valuable the more diverse the objectives one is trying to achieve with programmes and projects. According to Patton, utilizing the case study method is useful specifically in cases where one is trying to identify essential and detailed features from individual programmes and projects. In slightly the same vein, the handbook on evaluation methods published by the European Commission in 1999 (MEANS Handbook 1999; see MEANS 1999) states that in the field of evaluation research the case study has been utilized as a method in recent years particularly in such programmes and projects where the objectives and goals are highly diverse and complex (ibid., 59, 73-77).

In the field of evaluation research, Robert K.

Yin has been perhaps the most influential figure in terms of methodology since the early 1990s (see e.g. Yin 1993 and 1994). According to Yin (1994, 13-15), the case study is empirical research which can be used to analyze social phenomena in real-life contexts, particularly in situations where the boundaries between the phenomenon under examination and the context are unclear. Yin adds that the case study in terms of its approach is a research method in which several different data sources will be used and in which one relies on theoretical propositions prescribed in advance (and which, furthermore, direct the collection and analysis of the data). The case study is, moreover, bi-directional in nature. It can be the analysis of one case (single-case design) or many cases (multiple-case design).

All in all the case study appears to have high value in empirical social and socio-scientific research. For example, Robert E. Stake (2000) predicts that the case study will in future continue to be a favoured research genre, chiefly because it can be used to build a bridge between survey research and qualitative research (see Männikkö 2002, 132-134, for a Finnish assessment). According to one view (Eckstein 2000), the case study is extremely useful particularly in that stage of research when theoretical hypotheses are being formulated. It should be said that Stake's conceptual innovation relating to the case study method debate is naturalistic generalization, with which he refers to the analysis of the differences and similarities between different cases.

The method classics who write about the case study approach distinguish several case study types. In this respect, Yin's method, outlined above, of understanding the nature of the case study remains superficial to some extent. On the other hand, Harry Eckstein (2000, 132-152) identifies five different forms of case study: firstly, there is the configurative-idiographic case study, in which the research data is descriptive in nature and in which theoretical generalizations are not the objective; secondly, there is the disciplined-configurative case study, in which the basis of the research, including question setting, is reminiscent to some extent of positivist survey research. This is a matter of testing theoretical hypotheses; thirdly, one can distinguish heuristic case studies, which are laboratories for theory development, in which close analytical induction is used to promote Truth by utilizing qualitative methodology; fourthly, there exist so-called plausibility probes, which are performed as if for further assurance. Here one wishes further confirmation about the crystallization of already developed theoretical postulates and empirical observations into generalizations ("let's do just one more test!"); fifthly, one can identify critical case studies, in which the aim is to compare two or three cases in order to find differences and similarities between them.

3. IS GENERALIZATION FROM INDIVIDUAL CASES POSSIBLE?

Case study positions are always based on theoretical and conceptual propositions. This distinguishes the case study from, for example, the principle of qualitative evaluation research (particularly in the form presented by Guba and Lincoln), the ethnographic approach to research and the grounded theory method. The above-mentioned Yin (ibid., 30-31) speaks in positivist spirit when emphasizing that the aim of the case study is generalizing from case to theory. But what are the methodological questions relating to the case study approach? According to Hammersley and Gomm (2000, 5-6) there are four dimensions in this respect.

The first of these relates to the general application of observations and conclusions. To what extent can conclusions relating to individual events be generally applied to wider populations? The second methodological dimension relating to the case study approach is the question of causality of analysis and interpretation. In simple terms, this involves the extent to which one can assume a link of phenomena between cases, especially in cases where the same case is analyzed temporally as a panel arrangement at multi-dimensional points (on a timeline t ' t-1 ' t-2 etc.). The third methodological dimension is connected with the idea of theoretical perspectives relating to the case study method. This means that one can think about which parts or which entity of a case can be considered to represent a wider entity in the surrounding social reality ("what part/entity can be generalized?"). The fourth and final methodological dimension is connected with a deliberation about the authenticity of cases, namely the extent to which each case is unique. As one can see, there

are aspects connected with the case study method that have to be approached almost in a "methodological" spirit when one is planning the data collection stage.

The case study method literature strongly emphasizes the question of empirical observations and generalization (see e.g. Gomm et al. 2000b - in this anthology the main tension is precisely in the kind of interpretation the writers formulate in generalizations made from empirical observations). In philosophy of science terms, inductive thinking produces generalizations. The products of inductive thinking - observations, conclusions - also relate to individual events, which have not been studied yet. In this lies the undisputed power of inductive thinking, but at the same time also its weakness. Thus there are two sides to a coin, as David Hume (1711-1776) in his time remarked. Hume justifiably asked what grounds we have for believing the validity of the inductive principle. When inductive thinking produces generalization beyond individual cases, undeniable uncertainties are contained within the induction. Is generalization merely an uncertain jump into the unknown?

Why then has the question of generalizability raised its head? And why expressly in connection with qualitative research methodology? It is considered a basic principle of qualitative research methodology that generalization is not as a rule connected with the positivist idea of finding universal laws by which one can explain the appearance of a studied phenomenon in a population wider than the research sample. Such a basis is not a standard that applies to the generalization of qualitative methodology. On the other hand, in studies conducted from the perspective of qualitative methodology an acceptable basis is that conclusions about the studied phenomenon can be related to wider entities. Thirdly, qualitative methodology has been considered to offer a new opportunity to relate research results to a wider phenomenal field, applying various phrases such as "searching for fittingness", "transferability" and "dense recording".

The titles of articles in the above-mentioned anthology edited by Gomm et al. illustrates well interpretations made about the possibilities of generalizability: one extreme is represented by Stanley Lieberson in his article "Small N's and Big Conclusions" (Lieberson 2000) and the other by Yvonna S. Lincoln's and Egon C. Guba's (2000) view that, in scientific research, "The Only Generalization is: There is No Generalization". According to Lincoln and Guba, "the trouble with generalizations is that they don't apply to particulars" (ibid., 27).

It should be said that Yin's (1993 and 1994) way of interpreting the generalizability of the case study is undeniably very simple and limited. For example, Lincoln and Guba construct a substantial argument in the mentioned article when shooting down, to their mind heretical, hopes of positivist-spirited generalization. Lincoln and Guba state that it is pointless to imagine that generalizations are phenomena that appear in nature. The writers in question flatly deny this. To them, generalizations are mere - to intensify slightly, artificial - creations of the human mind. To Lincoln and Guba this represents, in a way, a new type of mechanomorphism (the belief that the universe is a machine and that everything works within the framework of universal laws). To their mind, generalization is an illusion, because the most that can be attained by this approach is probabilistic uncertainties. They consider it a mistaken belief that the ideal of nomothetic science can be achieved by the case study method, employing the expression of the German philosopher Wilhelm Windelband: with a case study no natural scientifically explicable position can be achieved on the one hand in terms of causes and effects and on the other in terms of dependent and independent variables.

What value then does the case study have, if Lincoln and Guba are correct? In their opinion, the case study can be used to produce working hypotheses. It is a matter of the ideas and conclusions that can be derived from the cases studied and whose value can be determined on the basis of their degree of transferability and fittingness. We have therefore taken a step from the nomothetic ideal of the philosophy of science: the effectiveness of the working hypotheses will be decided based on how transferable they are between cases. In the opinion of Lincoln and Guba generalizability is thus the wrong expression for something more generally valid which one is trying to describe with observations achieved by qualitative methods.

To sum up the matter of generalizability, I am on the opinion that there exists a link between the conduct of a case study, the

conclusion made with the aid of data collected by a case study, and generalization. The case study method can be utilized only and if the cases are connected by appropriate theoretical and conceptual fundamentals. As far as the case study method is concerned, it is therefore not meaningful to speak about generalization emphasizing the positive connotation of the word, but rather about the referencing and transferability of conclusions relating to one individual case or a limited number of cases. Or then it is a matter at most of generalization to the level of theory.

How then can interpretations based on the numerically limited observation data of the case study be strengthened? How can the referencing of results and conclusions be performed? Jane Ward Schofield (2000, 88-92) outlines a host of ways whereby the credibility and representativeness of the case study can be increased. Two ways are worthy of mention in this context. Schofield speaks first about the case survey method, which means in practice the systematic referencing of individual cases to the existing research literature.

Systematic referencing in this context really refers to the detailed analysis of observations and conclusions in relation to conclusions in the relevant research literature. According to Schofield, the procedure is carried out by locating the appropriate research literature and by making for it a systematic content analysis in terms of the conclusions raised in the case study. Although a way of thinking such as this is interesting in principle, it is evident that there is a host of problems associated with it, not the least of which is whether the said appropriate research literature can actually be found. Another problem relates to the fact that ways of producing the relevant research literature (method selection) might vary greatly, which will weaken the possibility of systematic comparison.

Another term raised by Schofield is meta-ethnography, which is reminiscent of the case survey method described above, but which fundamentally differs from the case survey method in the setting of the objective. While the objective in the case survey method is to reference the conclusions of the conducted case analysis to conclusions in the research literature, the aim with meta-ethnography is to reference, to each other, case studies that may be implemented in different ways in terms of data collection methods. This is therefore a question of converting case analyses implemented in different ways into others. Individual case analyses are thus translated into the language of the others so that the results of different case analyses can be examined in reference to each other. It should be stated that Yin (1994, 121-122) uses in respect of this way of approach the terms case survey: secondary analysis across cases, which therefore differs from Ward Schofield's case survey definition. In the field of evaluation research, people have been in the habit of speaking about meta analysis when referring to ways of approach that analyze a number of individual evaluation research studies (e.g. Weiss 1998, 236-238). The 'meta' term used by Schofield relates in this context to the fact that individual case studies are examined in the light of some single criterion or several criteria, in the same way as meta-evaluation in the form presented by Ray Pawson and Nick Tilley (1997) in their work Realistic Evaluation (specifically in the form that individual evaluation studies and their quality are analyzed in the light of pre-selected criteria, for example so-called MEANS criteria).

4 CASE STUDY IMPLEMENTATION IN PRINCIPLE...

In the dispute relating to case study generalization outlined above, could we just as well be content with a statement like that of Robert Gomm et al. (2000, 98-99), namely: Is generalization necessary on the whole? And do we really need the squabbling associated with generalization? Would it be more useful - instead of arguing about the overall possibilities and impossibilities of generalization - to focus on pondering how analysis can really be done in the case study method. Gomm et al. (2000, 104-111) propose two ways of implementing a case study: firstly, empirical generalizations between researched and unresearched cases and, secondly, aiming for generalization on the basis of a single case. When aiming for empirical generalizations between researched and unresearched cases, one must start by pondering the number and nature of the cases. What, for example, is the degree of heterogeneity and homogeneity of the cases? In other words, read,

deepen your knowledge and understand the subject of the case study better! When embarking on the actual conduct of the case study, the first stage is specifically the extent to which one can infer, given the existing information, whether the cases in question are typical or atypical based on some criterion to be examined. Connected with this is the selection of the cases themselves, for which different strategies exist.

For the case study one can select on the one hand typical cases, the kind that are very similar; on the other hand cases that fall within the sphere of the research "with a little trouble"; a third option is to select cases that enable as large an examination variation as possible (in relation to some examination theme or a number of examination themes); or fourthly by supplementing data and seeking cases that complement one another on the basis of what seems interesting from the standpoint of developing the themes of the study. The essential point in this strategy is to realize above all how, for example, the research literature will be utilized in building research hypotheses and in referencing interpretations made from the cases.

Robert Gomm et al. (2000, 108-111) call generalization "within" one case internal generalization. This can be done in many different ways, for example by utilizing different kinds of empirical data as well as interpretations made by different researchers, in the spirit of triangulation. Besides approaching the subject to be studied from various standpoints, the idea of triangulation is also to generate anomalies or even possible conflicting views in the phenomenon to be studied (reductio ad absurdum). The phenomena to be studied are always relative, and a certain randomness is associated with the interpretation of the phenomena. I confess that here I am adapting the basic idea of epistemological relativism: it is difficult to imagine finding unequivocal or absolute truths. In this context one can, moreover, emphasize the significance of time in case study data acquisition. Time is one of the most decisive credibility criteria of research results. A single case can be analyzed in its context as if in a parallel study - at many momentary cross-sections in time.

Finally, it is worth recalling that the case study approach has a special status in evaluation research. Why? With the case study one can, according to Yin, explain the causal links between phenomena that remain out of reach when using the survey method, for example. It is, in a way, a question of the "case study" conversion of the basic thesis of the realistic evaluation school, the Context-Mechanism-Result position (see e.g. Pawson & Tilley 1997). On the other hand, the case study can help to describe a certain phenomenon to be studied much more precisely than with a traditional survey or the qualitative method. Thirdly, the case study is suitable for such evaluation positions in which the subject to be evaluated (a project, programme or suchlike) has no clear assumed results or impacts and, fourthly, the case study is an excellent meta-evaluation tool when evaluating, for example, the method of evaluation already undertaken. The link between the case study and evaluation activity has, in recent years, been clarified also by familiarization with the TORs of different evaluation projects funded by the European Union - the case study is offered as a research tool in one or more projects.

5. ...AND THE CASE STUDY RESEARCH MODEL IN PRACTICE

Case study textbooks rarely tell how a case study is conducted in practice. It is worth starting a case study analysis by determining the volume of cases to be analyzed (number of cases investigated) and thereafter the level of intensity with which each case will be analyzed (amount of detailed information). And then one requires a strategy for selecting cases, a process about which the evaluation research manuals, almost without exception, say nothing. One exception, however, is How to Use Qualitative Methods in Evaluation (Patton 1987), published by Michael Q. Patton 15 years ago, in which the writer presents a host of various ways of proceeding in the selection of cases for case analysis (see especially ibid., 51-58). In Patton's ten-point typology, the tenth and final way of selecting cases for analysis is interesting and such formulations surely tend to increase the suspicions directed at the case study method (particularly in terms of transferability). Patton's guidelines is that appropriateness (convenience sampling) can also be selection criteria. Therefore select cases that are appropriate and, above all, cases whose study can be carried out quickly.

Patton is certainly aware of the limitations and dangers of such a strategy. In his view, however, this is the most common approach and, according to Patton, the least wished for, particularly in view of factors connected with enhancing the transferability of the case analysis.

A position based on a case study has, according to Yin (1994., 20-27), five critical factors, which the researcher must always tell to the reader: 1) Research questions (Yin urges the asking of how and why questions in particular); 2) Theoretical assumptions/hypotheses and conceptual models (thus the research does not emerge from a tabula rasa, but from a specific advance understanding of the research subject); 3) Unit[s] of analysis, namely that which is the subject of the case study (the case can be an individual, group, event etc.; the unit of analysis is derived from the research questions); 4) The chain of reasoning between the empirical data and the theoretical assumptions/hypotheses (Yin refers to Campbell and highlights the "finding of common patters" [pattern-matching] as one of the most promising case study logics - see more about this in this article, below); 5) The interpretation criteria of the empirical data, i.e. in practice the connecting thread between research questions, the theoretical frame of reference, and observations derived from the data (in other words, this aspect is connected to the credibility of the case study).

When examining Yin's rationale it is worth noting the factor that distinguishes him from Guba and Lincoln's very absolute application of qualitative analysis. Yin emphasizes that the case study method is based on detailed research, whose cornerstones are specific evaluation questions, examination of control groups, hypotheses about the subject to be evaluated and the scaling of data collection practices specifically to these methodological fundamentals. This approach is totally absent from Guba and Lincoln, who refuse to accept such a method of conducting an evaluation. In their opinion, the strict delineation of the evaluation approach is a waste of time and even harmful. The difference between Yin and the above-mentioned Guba and Lincoln is very clear. (See e.g. Weiss 1998, 181-182, about these problems.) There are four refined case study research model variations (Yin 1994, 38-52). The distinguishing factors between these "basic models" are the number of units of analysis and the number of cases being analyzed (single-case designs vs. multiple-case designs). The following outline describes this in more detail.

Justifications when selecting the number of cases for one study. Single-case analysis is justified when the case in question is critical, extreme or unique. In other circumstances, multiple cases are chosen. The selection of cases has been written about by Patton (1987), among others, and the matter has been discussed above in this paper. The number of units of analysis is decided by the nature of the research questions: sometimes it is sufficient to study a very limited and narrow aspect of some phenomenon. In such circumstances, a single unit of analysis (i.e. one basic research question) is sufficient. Ordinarily, however, units of analysis are linked to one another.

The three principles of case study data collection are (Yin 1994, 90-99) utilization of several data sources (data triangulation); a data documentation system; and keeping the chain of reasoning clear. Of these, the latter has

Table 1. Case study models.

Unit of analysis	Number of cases	
	Single-case design	Multiple-case design
Single unit of analysis (research theme), i.e. holistic model	Type 1	Type 3
Multiple units of analysis connected with one other	Type 2	Type 4

Table 2. The process of the case study method, according to Robert K. Yin.

1. Designing the research model	Theory development and theoretical perspectives
	Selection of cases (number of cases)
	Plan (unit of analysis) relating to the collection of data and a documentation system (data collection protocol). It is useful if the analysis can utilize several different kinds of empirical data (triangulation).
2. Carrying out the collection of research data	Analysis of the first case, analysis of the second case, analysis of the third etc.
	Writing of an individual case report for each case analyzed
3. Analysis stage and formulation of conclusions	Writing of a synthesis report covering all cases
	Modification of theory (i.e. a return to theoretical propositions)
	Formulation of conclusions
	Preparation of final report

(Yin 1994, 49, 90-95; see also MEANS 1999, 76-77).

wider validity in the conduct of evaluation research. The key characteristic of an evaluation researcher's expertise is the capacity to distinguish observations, conclusions relating to observations, and action recommendations constructed from conclusions.

Yin (1994, 106-107) uses the descriptive term pattern matching, by which he means the systematic analysis of observations collected from different cases, such that the researcher compares features common to certain phenomena being studied (which have been defined in connection with the pondering of the research questions) using data collected from different cases. The term pattern matching in connection with the case study therefore differs substantially from the definition used in connection with policy evaluation and programme-theory thinking - that the actual results and impacts of a programme are compared "Campbell-like" with the assumed results and impacts of the programme (see e.g. Weiss 1998, 66-67 for more about this). Another useful way of analysis in Yin's opinion is connected with time-series analysis, namely the "monitoring" of a case in time. Without an analysis conducted at different points in time, understandings of what is happening in the phenomenon that is the subject of the case analysis and with what consequences can easily remain only deductions that seek consistency (see also Alapuro 1995, 317-318). In the ideal situation the case study model would be one in which the performer of case study is able to construct panel arrangements at certain points in time when analyzing the case in question.

6. FINALLY: WHAT ABOUT THE CREDIBILITY OF THE CASE STUDY?

The credibility of the case study has been the subject of intense debate in recent years. It is a question of the reliability and validity of studies conducted applying the case study idea, in other words evaluating the reliability of research results. Prejudices associated with the case study method - to sum up earlier thoughts - relate to shortcomings in the systematic nature of the analysis method (lack of rigour), to restrictions

on generalizability and to practical problems introduced by its application (appropriate case analysis takes time, documentation and reporting might seem like wearisome lists of detailed crumbs of information etc.). The credibility of the case study is linked to how case study projects are reported. Here, Yin (1994, 147-152) is very precise: according to him, case study reports must 1) be directed at the essential questions, 2) they must be systematically constructed (covering all questions defined in the research model), 3) they must take into account the effectiveness of alternative working models (depending, of course, on the subject of the analysis), 4) they must be based on sufficiently extensive empirical evidence, and finally, 5) they must be written in an accessible and engaging manner.

Timo Toivonen (1999, 140-141) is to a certain extent even more demanding than Yin when he relates the ways of reporting case studies to the objectives of the method itself. According to Toivonen "...the key idea of the case study lies in the fact that with its help one thinks one can disentangle simply some process complete with its cause and effects". It might be possible to achieve a symmetrical relationship of two variables (a situation in which neither is the cause nor the effect of the other) or reciprocal causality (the entities are the cause and effect of each other, even though it is difficult to put one ahead of the other, i.e. the "chicken and egg" phenomenon) with the aid of a case study, but how reliable and transferable the conclusions are is always open to question (see Virtanen & Uusikylä 2002 for this debate on the level of programme evaluation). Here we come close to where this article began, namely to the question of the generalizability of the case study and of generalizability in relation to what.

Case study credibility can be considered to be governed by four criteria, which are: 1) construct validity: the success of the operationalization of the concepts used; 2) internal validity: the demonstration of causal relationships (not always the objective of the study, but particularly in explanatory case study models); 3) external validity: to what extent are the research results generally applicable?; and 4) reliability: the repeatability of the study procedures does not depend on the researcher. This list is, of course, long and wearisome. It could be condensed to state that the researcher applying the case study ensures above all that he describes the course of the research process in such detail that the critical reader can follow his chain of reasoning via the research questions to the data and on to the observations, conclusions and recommendations. Without his connecting thread, the study may at worst be suspected of randomness.

Finally, we can still ask about the terms of transferability of the conclusions, namely how conclusions achieved with the aid of case analysis are suitable in other contexts (fittingness); to what extent can we be convinced of the principle of reconcilability or are we blindly following the delusion of fittingness? To what extent are the principles of triangulation genuinely applied and how much critical thought has gone into how different types of conclusions, produced in different case studies, relate to each other and where they even conflict (cf. Denzin 1978)? The key requirement in an evaluation research project implemented by the case study method is - and this, of course, is also valid more widely for research work - that observations and conclusions can be interpreted. Without interpretation the observations and conclusions have no value. This is, of course, a stiff challenge for practitioners of evaluation research, because this requirement means that besides technical research skills they are also expected to have the wisdom brought by life experience and a high level of competence.

The general quality requirement connected with evaluation research reporting also applies to case study reporting. The reports must be able to explain the differences between the observations, conclusions and recommendations. Indeed, case study reports should contain a prescriptive assumption (i.e. the giving of instructions) - the idea that the report includes very concrete and, in terms of implementation, realistic proposals for action. In this respect, we can also go along with the comment, which at first glance perhaps seems rather laconic: Writing is indeed a disciplined activity. All that is said above is naturally connected with the fact that we are aware of the asymmetry in size between the research problem and the research subject. Although the research subject, too, may sometimes be small and restricted in size, this does not automatically mean that the problem to be investigated is of little significance.

NOTE

[1] I am grateful to Pertti Ahonen and Petri Uusikylä for their critical and constructive comments to earlier drafts of this article.

REFERENCES

Alapuro, Risto (1995). Suomen synty paikallisena ilmiönä 1890-1933. Porvoo: Hanki ja Jää.

Eckstein, Harry (2000). Case Study in Political Science. Pp. 119-164 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Eskola, Antti (1992). Sosiologian uudistuminen 1950-luvulla. p. 241-285 teoksessa Alapuro, Risto, Alestalo, Matti, Haavio-Mannila, Elina (toim.) Suomalaisen sosiologian historia. Porvoo: WSOY.

Denzin, Norman (1978). The Research Act. Chicago: Aldine Publishing Company.

Donmoyer, Robert (2000). Generalizability and the Single-case Study. Pp. 45-68 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Fischer, Frank (1995). Evaluating Public Policy. Nelson-Hall Publishers: Chicago.

Gomm, Robert, Hammersley, Martyn (2000). Introduction. Pp. 1-16 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Gomm, Robert, Hammersley, Martyn, Foster, Peter (2000a). Case Study and Generalization. Pp. 98-115 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Gomm, Robert, Hammersley, Martyn, Foster, Peter (eds.) (2000b). Case Study Method. Key Issues, Key Texts. London: Sage.

Guba, Egon, Lincoln, Yvonna (1989). Fourth Generation Evaluation. Newbury Park: Sage.

Lieberson, Stanley (2000). Small N's and Big Conclusions. Pp. 208-222 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Lincoln, Yvonna S., Guba, Egon C. (2000). The Only Generalization is: There is No Generalization. Pp. 27-44 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Männikkö, Jarmo (2002). Voluntaristinen alueellinen innovaatiojärjestelmä. Tapaustutkimus Oulun alueen ict-klusterista. Acta Universitatis Lapponiensis 46. Rovaniemi.

Patton, Michael Q. (1987). How to use Qualitative Methods in Evaluation. Thousand Oaks ym.: Sage.

Patton, Michael Q. (1997). Utilization-focused Evaluation. Newbury Park: Sage.

Pawson, Ray, Tilley, Nick (1997). Realistic Evaluation. London ym.: Sage.

Pipping, Knut (1947/1978). Kompaniet som samhälle - laktagelser i ett finskt frontförband 1941-1944. Acta Academiae Aboensis. Humaniora XVI.I. Suom. Komppania pienoisyhteiskuntana. Sosiologisia havaintoja suomalaisesta rintamayksiköstä 1941-1944. Keuruu: Otava.

Stake, Robert E. (2000). The Case Study Method in Social Inquiry. Pp. 19-26 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Toivonen, Timo (1999). Empiirinen sosiaalitutkimus. Filosofia ja metodologia. Porvoo. WSOY.

de Vaus, D.A. (1996). Surveys in Social Research. London ym.: UCL Press.

Virtanen, Petri, Uusikylä, Petri (2002). Exploring the Missing Links Between Causes and Effects. New Conceptual Framework for Understanding Micro-Macro Conversions in Programme Evaluation. A paper presented at the 5th EES (European Evaluation Society) Conference, 8.10.2002, Seville, Spain. 2002.

Ward Schofield, Jane (2000). Increasing the Generalizability of Qualitative Research. Pp. 69-97 in Gomm, Roger, Hammersley, Martyn, Foster, Peter (eds.) Case Study Method. Key Issues, Key Texts. London: Sage.

Weiss, Carol H. (1998). Evaluation. Methods for Studying Programs and Policies. New Jersey ym.: Prentice Hall.

Yin, Robert K. (1993). Applications of Case Study Research. Newbury Park ym.: Sage.

Yin, Robert K. (1994). Case Study Research. Design and Methods. Thousand Oaks ym.: Sage.