A Systematic Process for Selecting a Typical Case in School Setting

Said Morad Babaheidari [1], Lena Pareto [1], Maria Spante [1], Lars Svensson [1]

[1] University West, Department of Economy and IT,
461 86 Tollhättan, Sweden
{said-morad.babaheidari, lena.pareto, maria.spante, lars.svensson, LNCS}@hv.se

Abstract. This paper describes the rationale behind the concrete steps how we systematically have gone through the case selection process as the first phase of a research project named Consequences of the Digitalization of Schools. The project is a single-case longitudinal (5-year long) study and adopts a so called Whole-School Approach. Previous research has criticized case studies for lack of both systematic approach and transparency in the presentation of the case selection process. The purpose of the presentation of this case selection process is to contribute to enhancing understanding of and broadening knowledge on systematic and transparent case selection techniques. The SIRIS database has been used to select a typical Swedish primary school in the chosen municipality where the project will be conducted. The SIRIS database contains annually statistics from all primary schools in Sweden, and is available online for public use. We demonstrate how and why a particular school representing a so called typical school was selected as our case.

Keywords: Case Study Research, Case Selection, Purposive Selection, Whole-School Approach, IS Research Framework

1 Introduction

The Swedish school is currently experiencing a strong push for implementing new Information and Communication Technologies (ICT). Schools are increasingly investing in classroom technology such as interactive whiteboards and so-called 1-to-1 ventures. There is an upward competitive trend to equip each student and teacher with laptops. Schools increasingly use information systems (IS) to e.g., report student daily attendance, to support and monitor student progression and achievement. Furthermore, IS are also increasingly used by schools to communicate with parents, with the hope of bridging the gap and increasing interactivity between home and school. However, little is known regarding how this digitalizing of schools affects the diverse activities that signify schools as organizations and the involved actors such as students, teachers, headmasters, administrative personnel and parents.

In this paper, after briefly introducing the initial phase of a longitudinal (5-year long) case study research project entitled "*The Consequences of Digitalizing Schools"*

at University West, Sweden, we describe in detail how the case selection process is done. Previous research on case selection processes has called for a more transparent and systematic approach while selecting a case. Gerring (2007, p. 6) reports that there is a lack of documentation of: "…why a specific case or set of cases has been selected." In addition Seawright and Gerring (2008, p. 294) postulate: "Despite the importance of the subject, and its evident complexities, the question of case selection has received relatively little attention from scholars since the pioneering work of Eckstein (1975), Lijphart (1971, 1975), and Przeworski and Teune (1970)." The main purpose of this paper is to demonstrate how we've gone through the case selection process and argue for the benefits of a systematic and transparent approach.

2 Research Method

Braa and Vidgen (2000) provides a framework in which case studies can be positioned based on what rationale the case study design was driven by as well as intended outcome of the conducted case study. The framework is based on the idea that research method can be mainly grouped into two distinctive categories of (1) positivist, and (2) interpretivist.

The framework is represented by a triangle shown in Figure 1 below. The triangle is comprised of dotted arrows, sides, and a constrained space. According to the authors (Braa and Vidgen, 2000), different positions on the dotted arrows represent intended research outcomes. If the research intention is to predict, then the case study would be positioned closer to the left in the triangle. Whereas, when intention is to gain understanding, then the case study would be positioned to the right. However, we are interested in the understanding notion which according to the authors (ibid) is aligned with the interpretive approach to research following Braa and Vidgen (2000).

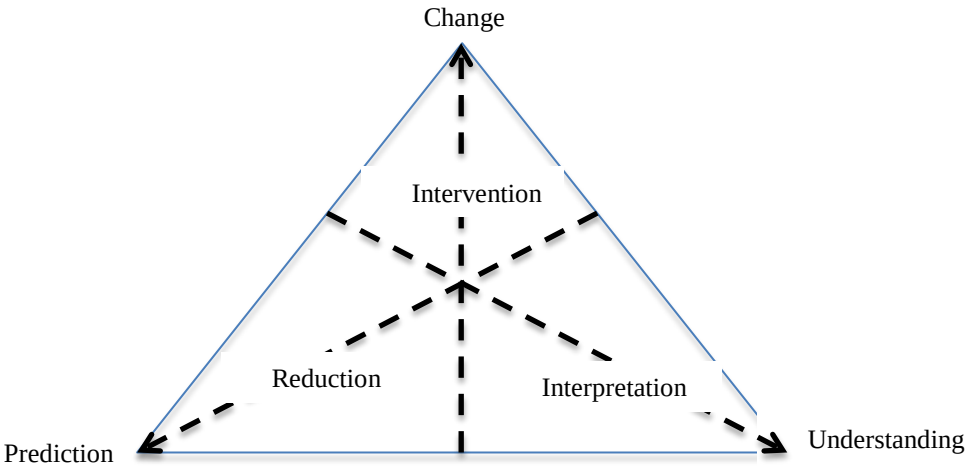

Figure. 1. An IS research framework for the organizational laboratory (Braa and Vidgen, 2000; p. 255)

Moreover, we can link the fundamental philosophical underpinnings of interpretation, following Braa and Vidgen (2000), to our single case Whole-school approach since the overall project aim is to getting a better understanding of the consequences of the digitalizing of Swedish schools, investigating a 'typical' school.

3 Literature on Case Study Research

Authors on case study research have defined case study to be a research strategy (Yin, 2012; 2003; Merriam, 1998), a research choice (Stake, 1995), and a research methodology (Bryman, 2008; Creswell, 2007; Gerring, 2010). Yin (2003, p. 13) defines the case study research as: "…an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used."

Conducting single case has many advantages such as investigating a phenomenon both in-depth and during a long period of time to see if such a case is e.g., deviant, representative, or unique (Yin, 2003; Stake, 2005; Merriam, 2008). Gerring (2010, p. 187) states, when a case study researcher seeks to explain the outcomes of a investigated single case, he/she will base his/her reasoning on why something has occurred, or why; "something might have happened but, in the event, does not. This is, the outcome may be "positive" or "negative."

According to George and Bennet (2005, p. 21), within a single case, the researcher looks at; "...a large number of intervening variables and inductively observe any unexpected aspects of the operation of a particular casual mechanism". Similarly, Eisenhardt (1989, p. 534) defines case study as; "…a research strategy which focuses on understanding the dynamics present within single settings." Literature on information systems (IS) reveals the use of single case study by many IS researchers (e.g, Walsham, 1995; Benbasat et al., 1987; Cavaye 1996; Darke, Shanks and Broadbent 1998).

Table 1: Purposeful sampling strategies and their definitions (Patton, 1990; pp. 182-183)

Types of purposeful sampling strategies	Definition of the sampling strategy
Extreme case	The case demonstrates unusual manifestation of the phenomenon, such as outstanding success and notable failures.
Intensity case	The case is information rich but not an extreme case.
Maximum variation	Cases, despite having diverse variations, exhibit important common patterns that cut across variations.
Homogeneous	Variation between cases is minimized, analysis is simplified and study is focused.
Typical case	Case illustrates what is typical, normal or

	average.
Stratified purposeful case	Case illustrates characteristics of a particular subgroup to facilitate comparison and not for generalization or representation.
Critical case	Case that permits logical generalization to other cases because if it is true to this one case, it's likely to be true to all other cases.
Snowball	Cases of interest from people who know people who know people who know cases, rich information rich, good examples for study, etc.
Criterion	Cases picked because they meet some predetermined criterion.
Theoretical	The cases are manifestation of a theoretical construct and are used to examine and elaborate on it.
Confirming and disconfirming	Cases that elaborate on initial analysis to seek exceptions or test variations.
Opportunistic	Cases that emerge from following leads during field work.
Random purposeful	Cases are randomly selected from a large sample for the purpose of increasing credibility and not for generalization or representation.
Politically important case	Cases are selected or eliminated because they are politically sensitive cases.
Convenience	Cases are selected on the basis of minimum effort, time and money. They are candidate examples of low credibility, information rich cases.
Combination	Cases are flexible and meet different interests and needs cases.

Besides various definitions of case study research which indicates that there is a lack of consensus among researchers, Gerring (2007, p. 6) reports that there is a lack of documentation of: "…why a specific case or set of cases has been selected." In addition Seawright and Gerring (2008, p. 294) postulate: "Despite the importance of the subject, and its evident complexities, the question of case selection has received relatively little attention from scholars since the pioneering work of Eckstein (1975), Lijphart (1971, 1975), and Przeworski and Teune (1970)."

However, regardless of the aforementioned views on case study research concerning its discussed purposes, and particularly the literature on case selection techniques, we argue whether choosing a single-case or a multiple-case to study, it's necessary to adopt a systematic and transparent approach to case selection process which is the main topic of this paper. In our study, we are searching for a Typical case (see table 1, based on Patton 1990) looking for a 'typical Swedish school'.

4 Procedures in Previous Research Regarding Case Selection

In this section of the paper, we explore common ways of case selection. Yin postulates; as a first step, the researcher must initially define the *"case"* which should be derived from the primary research question which, in form of worded phrase, states exactly the focus of the research. Yin (2003, pp. 21-22) further argues that five components of research designs are particularly important (see Table 2 below).

Table 2: Five components of case study research designs

Components of Research Designs	Description
1. Question(s)	Those fundamental research questions to be addressed and thus answered by the researcher should be "what", "where", "who", "why" and "how". Considering the case strategy, choosing "why" and "how" are the two most appropriate questions.
2. Propositions, if any	Propositions would deal with those possible links that a researcher will find. Example, the possible grounded outcomes, or expected finding results.
3. Unit(s) of analysis	Fundamental problem of "What" the case to be studied really is. E.G., a group of exemplary students is a possible unit of analysis.
4. Logical link between data and the propositions	E.G., the choice of some deployed techniques used to analyze the collected information, and to be later compared with the proposition(s).
5. Criteria for interpreting the findings	How to evaluate if findings support the propositions made or not.

He continues to say the researcher's tentative definition of the third element of the unit of analysis, and thus of the "case", is related to the way the researcher has defined her initial research questions. Yin (2003; p. 23) labels this as a "general guide." It means only after the researcher has accurately specified the research questions, she can thereafter select some appropriate unit of analyses. In our longitudinal single-case study research project (i.e., The Consequences of Digitalizing Schools), the unit of analyses will be all those previous mentioned diverse activities that signify schools as organizations and the involved actors such as students, teachers, headmasters, administrative personnel and parents.

Other case study researchers such as Stake (1995; p. 4) states that the first criterion before deciding to select a case "should be to maximize what we can learn. Given our purposes, which cases are likely to lead us to understandings, to assertions, perhaps even to modifying of generalizations?" Learning is not the only criterion; Stake additionally proposes a list of criteria that are: (1) a typical unique case, (2) representative of other cases, (3) easy accessibility, (4) hospitality, (5) identification of informant/(s), and (6) the intrinsic interest in the case.

As an example here, the hospitality criterion pointed out above means that the research environment, like a case study research in a school, must be indeed hospitable to our investigation. Similarly, the various positive contributions of informants/actors e.g., when doing case-study research in a school, are also invaluable. Informants, in our case, would be those identified students, teachers, and administrative staffs to simply cooperate with our research team e.g., in form of be willing to comment on a certain draft materials etc. Therefore, this is an important criterion as well. However, such an aforementioned list of criteria is what we can consider to be a part of the purposeful approach to case selection, or sampling strategy.

Moreover, Merriam (1994, pp. 3-4) discusses the issue of case selection criteria via her definition of case studies of interest. Merriam recommends the importance of one general- and three particular additional questions to be asked by the researcher before deciding to choose the best case of interest to study. These four questions are:

1. What types of questions the researcher should ask?
2. What degree of control the researcher has?
3. How the researcher would think what the final findings might be?
4. Could the researcher possibly identify a bounded system as focus of her research?

She argues question number four is probably the ultimate factor before deciding to select the best case of interest to be studied. She raises the issue of a bounded system which according to her is a system where the boundaries are clear as it's shown by question number 4 above. As an example and in school setting, a particular teacher, school, or a research method can all serve as clear boundaries. The bounded system could be counted as a good case selection criterion candidate as well.

Rowley (2002) has noticed three case selection factors that are: (1) Time, (2) Accessibility, and (3) Resources. Not having just access, but easy access to the case is also one crucial criterion. These factors are supported by other case study research authors such as Yin (2003), Stake (1995), Seawright (2008) and Gerring (2010). Furthermore, Seawright and Gerring (2008, p. 295) note that due to the lack of a well-formulated "formal treatments" researchers continued to focus on a pragmatic approach meaning they continue to lean primarily on following factors: (1) Time, (2) Money, (3) Expertise, (4) Access, and (5) Theoretical prominence of a given case.

However, the process of case selection is dealing with search for finding all necessary criteria in order to justify the final selected case(s). A case study researcher must be prepared to defend her final selected case, it'll say; must be able to answer questions such as why not selecting other potential available alternative case(s). In other words, one main crucial question which to be asked is: what makes e.g., case A to be preferred to case B? This is what we've attempted to demonstrate in this paper by describing our systematically purposive case selection process.

5 Selecting the Typical Case

While Stake and some other well-referred case study research researchers, such as Yin and Merriam, have covered the earlier mentioned aspects of case selection

criteria, Seawright and Gerring (2008) discuss the advantages of using purposive sampling case selection technique but rather in a statistical way. Purposive sampling can be shortly described as a stratified sampling technique by which some of those earlier mentioned criteria can be linked to whether a single case or a subset of the whole population of interest in order to see which case/cases can fulfill the pre-determined required criteria. However, in the subsequent sections we'll discuss both the rationale behind and as well describe how we use mixed methods sampling techniques.

The '*typical case'* has been defined by Seawright and Gerring (2008, p. 299) as a case which has its main focus on exemplifying: "…a stable, cross-case relationship. By construction, the typical case may also be considered a *representative* case…" These authors (p. 295) state: "In the absence of detailed, formal treatments, scholars continue to lean primarily on pragmatic considerations such as time, money, expertise, and access. They may also be influenced by the theoretical prominence of a given case. Of course, these are perfectly legitimate factors in case selection. Yet they do not provide a *methodological* justification for why case A might be preferred over case B." Thus, this is one reason why we need a more transparent and robust procedure when selecting a case, relevant for the research approach adopting a whole school approach combined with our adoption of an IS research framework discussed in research method section of this paper.

Gerring and Seawright (2007) advocate the use of various regression diagnostics to identify cases that would be of (1) typical, (2) deviant, and (3) influential amongst other types of cases. They describe how to use statistical matching procedures as a useful way to formalize a notion of most-similar cases. As Gerring (2010, p. 92) wrote: "…the more common employment of the typical-case method involves a *causal* model of some phenomenon of theoretical interest."

In comparison with the more common selection procedure of finding a good case based on pre-established selection criteria, the statistical matching method is argued to have some advantages The main argument in the case selection via statistical matching method is its more rigorous character and thus preferable particularly when there are many relevant variables (Nielsen, 2012; Gerring and Seawright, 2007; Gerring, 2010). Gerring (2010, pp. 88-90) provides techniques of case selection for nine case study types, one of which is the "typical case" defined to be "representative" by definition. According to Gerring (p. 91), in order for a focused case study to provide insight into a broader phenomenon, it must be representative of a broader set of cases. It is in this context that one may speak of a *typical-case* approach to case selection.

The typical case exemplifies what is considered to be a typical set of values, given some general understanding of a phenomenon. By construction, the typical case is also a representative case; I employ these two terms synonymously.

Teddlie and Yu (2007, p. 78) offer a taxonomy of sampling techniques shown in Table 4 below. These authors view purposive sampling such as typical case to be criteria-based sampling. Furthermore, and in line with what Gerring's definition of typical case, Teddie and Yu (2007, p. 81) underline the purpose of selecting a typical case is to achieve "representativeness or comparability."

Table 3: Taxonomy of sampling techniques for the social and behavioral sciences (Teddie and Yu, 2007, p. 78)

I. Probability Sampling
A. Random Sampling
B. Stratified Sampling
C. Cluster Sampling
D. Sampling Using Multiple Probability Techniques
II. Purposive Sampling
A. Sampling to Achieve Representativeness of Comparability
B. Sampling Special or Unique Cases
C. Sequential Sampling
D. Sampling Using Multiple Purposive Techniques
III. Convenience Sampling
A. Captive Sample
B. Volunteer Sample
IV. Mixed Methods Sampling
A. Basic Mixed Methods Sampling
B. Sequential Mixed Methods Sampling
C. Concurrent Mixed Methods Sampling
D. Combination of Mixed Methods Sampling Strategies

5.1 Data Analysis and Results: Framing the Typical School

For the purpose of framing the typical school, we used the SIRIS[1] database available online at The Swedish National Agency for Education (Swedish: Statens skolverk, commonly known as Skolverket) website. Each year, Skolverket provides school performance statistics for all public and private primary schools (Grades 1-9) that are spread across the whole country. Our mixed methods (MM) sampling techniques for case selection comprises of two steps: Firstly, based on quantitative data retrieved from SIRIS database we compared all primary schools (N=1665) in Sweden with respect to 128 variables (e.g. average grades, number of staff, percentage of students that qualifies for upper secondary school etc.). By calculating the means and standard deviations of all variables it was possible to explore the characteristics of a "typical Swedish school". Secondly, we created a set of criteria for a purposive selection of a single case.

In SIRIS database, all data about primary schools (Grades 1-9) were retrieved at (a) the school level, and (b) at the municipality level. The data were consisted of more than 200 variables measuring school performances such as the final average degree in each subject for each school in Sweden. But we've decided to select only 128 out of 200 variables that we identified to be relevant variables. We identified those

[1] http://siris.skolverket.se

remaining 72 (i.e., 200-128=72) variables to be irrelevant because, for instance, variables such as "The percentage of students whose parents are migrants" or "The school annual budget" lack useful information in our case. For example, we don't know how much of a school's total annual budget actually goes to ICT budget or IT department if there is any such, therefore annual budget was excluded. Similarly, a variable such as the percentage of students whose parents have migrated to Sweden is very low in the most schools that are particularly located in neighborhoods with very low-to-zero percentage of migrants. Therefore, these types of variables have been omitted due to their irrelevancy to our case selection reasoning.

After reducing the original 200 variables to 128, we computed the means and standard deviations of these128 variables. In order to define what constitutes typical values for each variable, we calculated the width of an interval surrounding the mean value. Chebyshev's inequality theorem was used for this purpose. Chebyshev's theorem states that regardless of the shape of the distribution (e.g., bell, skewed, bimodal, etc.) for any given K>1, at least $(1-1/K^2)$ of the population lie within $\mu \pm K\sigma$.

We decided to use K=1,414 which captures at least 50% of the population for each variable. In order to provide a practical solution for the longitudinal study, we selected the municipality of Trollhättan as the best choice based on proximity to the university as well as falling within the area of operation of the research sponsor. Trollhättan has 9 schools that met the requirements (public school with grades 1-9).

For each of the nine schools we explored how many of the 128 variables that were included or excluded in relation to the selected K-value. In Table 4 below the percentage of included and excluded variables are summarized.

Table 4 The outcome of statistical purposive sampling technique: 9 selected cases from the whole population

School/case names	**%Variables included of total 128**	**%Variables excluded of total 128**	**%Variables missing of total 128**
Kro1	29%	28%	44%
Syl1	58%	10%	33%
Sta1	48%	4%	48%
Skg1	44%	8%	48%
Sjt1	48%	5%	47%
Str1	52%	4%	44%
Hjm1	51%	4%	45%
Par1	60%	3%	37%
Lyr1	57%	3%	40%

We decided on the ≤5% limit regarding excluded variables and ≥50% limit regarding included variables. Based on that decision, four schools were included in the range of the decided definition of a 'typical school' (i.e., Str1, Hjm1, Par1, and Lyr1). Based on firstly the percentage of included variables and secondly on percentage of excluded variables, thus these below four selected schools were ranked in this order: Par1, Lyr1, Str1 and Hjm1.

As our final step of the case selection process, and derived from our conclusion based on the data showed in Table 4 above, we've thus contacted those above four

purposive selected schools since we needed to address other types of criteria. These criteria are linked to the requirements for long-term commitment particularly necessary for the longitudinal Whole-school research approach focusing on the consequences of the digitalization of schools. The suggested criteria in the next phase are presented in Table 5 showed below.

Table 5: A list of case selection criteria (inspired by Stake, 1995 & Rowley, 2002)

Management commitment; Time Budget	Easy accessibility	IT infrastructure & Resources: Hardware Software Internet access	Availability and interest of informants and actors

First we did an interview with the IT manager of all public schools (Grades 1-12) in Trollhättan municipality. Our main aim of the interview was to get some useful information on those criteria shown in Table 5 above. Derived from the mentioned interview with the IT manager, we've now know about parameters such as IT investments in those schools, and particularly we asked him about teacher attitudes toward the inclusion of ICT in their daily classroom teaching practices.

During the interview with the IT manger, it became clear that some schools have a clear articulated IT budget, and particularly a clear vision as well positive tendency toward the implementation of ICT in their organizations.

As the first and preferred candidate for the single case study based on the work with the SIRIS data base, the school Par1 was contacted. First, the project was presented for the headmasters of the school. Then, a week later, the same presentation was made for the whole school staff. It was of uttermost importance to get their consent and willingness to participate in the longitudinal project which will be conducted in their school. Thus, they were willing to participate in the 5-year long single-case study research project at their school where the project will be conducted.

6 Conclusion and Further Work

As a contribution to the case study research, we have discussed the case selection issue in order to both enhance the knowledge and also to broaden the awareness of case selection process as the first part of case study research, and regardless of choosing to conduct a single-case or a multiple-case study research. In educational settings and with emphasis on adapting what is known as whole-school approach to education for considering all involved stakeholders, a thorough case selection process should be the first phase before starting with the investigation process of the case study research.

We showed how four out of the nine finally selected schools were statistically qualified as potential candidates. These schools were generated by our proposed case selection process in which we searched for 'typical' schools. All four schools fell outside of the defined interval for less than 5% of all variables. Based on the rank order of the four schools, Par1 has emerged as the primary candidate. Consequently, these schools are defined as 'typical' in relation to the schools in Sweden.

Furthermore, compiling the quantitative data from the SIRIS database will make it possible for us to follow change over time.

As the next step, interview with headmasters and other representatives for those schools have finally led to an agreement and long-term access to the particular school which is now Par1. This school will be investigated during our longitudinal 5-year long single-case study research, and we have been able to demonstrate why Par1 was selected as the preferred candidate with a systematic and transparent approach.

As our final conclusion, we recommend case study researchers to go through the case selection process systematically and present the approach as transparence as possible. It is not only of value for the justification of the selected case. It is also of high value as a learning process about the specific area of investigation and getting an understanding of the case as such beneficial for the next phases in the research project.

Acknowledgements

The research project is sponsored by Länsförsäkringar Älvsborg, Sweden.

References

1. Benbasat, I, Goldstein, DK & Mead, M (1987). The Case Research Strategy in Studies of information Systems. *MIS Quarterly*, vol. 11, pp. 369-386.
2. Braa Kristin and Vidgen Rickard (2000). Research: from observation to intervention. Chapter 12 in Planet Internet. K. Braa, C. Sörensen, B. Dahlbom (eds.). Studentlitteratur, Lund, Sweden.
3. Bryman, A. (2008). Social Research Methods. 3th edition. Oxford University Press. ISBN 978-0-19-920295-9.
4. Cavaye, ALM (1996). Case Study Research: A Multi-faceted Research Approach for IS, Inform*ation Systems Journal*, vol. 6, pp. 227-242.
5. Creswell, J., W. (2007). Qualitative inquiry & Research Design: Choosing among Five Approaches. 2th edition. Sage Publications. ISBN 978-1-4129-1607-3.
6. Darke, P, Shanks, G & Broadbent, M (1998). Successfully Completing Case Study
7. Research: Combining Rigour, Relevance and Pragmatism", *Information Systems Journal*, vol. 8, pp. 273-289.
8. Eisenhardt, K., M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, Vol. 14, No. 4, pp. 532-550.
9. Gerring, J. (2007). *Case study research: principles and practices*, Cambridge University Press Cambridge.
10. Gerring, J. (2010) Case Study Research: Principles and Practices, Cambridge University Press.
11. George, A., L., Bennet, A. (2005). Case Studies and Theory Development. Cambridge MA: MIT Press. ISBN: 0-262-07257-2.
12. Merriam, S. B. (1998). Qualitative Research and Case Study Applications in Education.
13. Nielsen, R. (2012). Case Selection via Matching. Working Paper. Harvard University. Available at http://people.fas.harvard.edu/~rnielsen/Case Selection via Matching.pdf, accessed 28 September, 2012.

14. Patton, M.Q. (1990). Qualitative Evaluation and Research Methods, Sage Publications, Newbury Park, California.
15. Rowley, J. (2002). Using case studies in research, Management Research News, Vol. 25 Iss: 1 pp. 16- 27.
16. Seawright, Jason and Gerring, John (2008). Case Selection Techniques in Case Study Research: a Menu of Qualitative and Quantitative Options. Political Research Quarterly, Volume 61 Number 2, June 2008, pp. 294-308, © University of Utah, 10.1177/1065912907313077.
17. Stake, a. Robert (1995). The Art of Doing Case Study Research. ISBN 978-0-8039-5767-1.
18. Stake, R., E. (2006). Multiple Case Study Analysis. Guilford Press, New York: New York.
19. Teddlie C., Yu F. (2007). Mixed Methods: A Typology With Examples. Journal of Mixed Methods Research, Volume 1 Number 1, January 2007. 77-100_ 2007 Sage Publications 10.1177/2345678906292430.
20. Walsham, G. (1995). Interpretive Case Studies in IS Research: nature and method. *European Journal of information Systems, Vol. 4, No.2, pp.74-81.*
21. Yin, R., K. (2003). Case Study Research, Design and methods. 3th edition, ISBN 07619-2553-8.
22. Yin, R. K. (2012). Applications of Case Study Research, 3th edition, ISBN 978-1-4129-8916-9.