The handle http://hdl.handle.net/1887/25307 holds various files of this Leiden University dissertation.

Author: Elayah, Moosa Abdo Ali
Title: Donors-promoted public sector reforms in developing countries and the Local Knowledge Syndrome: the Dutch-Yemen NPT Program for developing the Higher Education Sector in Yemen
Issue Date: 2014-04-17
CHAPTER 5

STUDY METHODOLOGY ISSUES
5.1 Introducing the Research Argument

A methodology is a strategy that helps a researcher to systematically study a research phenomenon. In social research, examples of methodologies are positivism using quantitative methodology and qualitative methodology, which is often concerned with inducing hypotheses from field research. In addition, there is a mixed methodology approach, including both qualitative and quantitative methods. The methods are research techniques including quantitative techniques like statistical correlations, as well as techniques like observation, interviewing and audio-recording. In themselves, techniques are not true or false. They are more or less useful, depending on their compatibility with the theories and methodologies used in the tested hypothesis and the selected research topic (Silverman, 1993).

The main purpose of this chapter is to provide an outline of the adopted mix-research method in this study and to explain the procedures that were followed to collect data. This is needed to investigate the straightforward and primary question “why are foreign aid programs ineffective in prompting public sector reform in Yemen despite all the involved resources?” The answer to this question is not limited to the Yemeni context only, but applies to many poor nations where the effectiveness of foreign aid activities is equally unsettling.

To address the main question, I have used literature on the justification and effectiveness of donor aid programs in developing countries. As we outlined in Chapter 1 this study is will be undertaken in three stages. The first stage includes the literature review of foreign aid effectiveness in Chapter 2. In that stage, we saw two contradictory approaches to aid. The PIP approach argues that foreign aid does work and that it should be continued to solve the “poverty trap” in developing countries. The PCP approach on the other hand argues that foreign aid does not work and moreover that it harms developing countries instead of helping them; and so it should be stopped to avoid future harm. The factors of aid ineffectiveness provided by many leading scholars of the PCP approach was reviewed and the argument of the PIP approach was isolated.

PCP approach scholars have divided the factors of why foreign aid is ineffective into two main strands. Firstly, much of the focus of macro studies argue that the poor policy environment, such as weak and corrupt policies and institutions in developing countries, is the main cause of the shortcomings of foreign aid programs (see for example: Burnside and Dollar, 1998; Ear, 2006; Alesina and Dollar, 2000). Secondly, other PCP scholars found, based on country level analysis, that most donors are still
tying their aid to egoistic interests that can be characterized as political, strategic, commercial, cultural, and religious. They are highly motivated to achieve those egoistic interests and less interested in holding recipient countries accountable for achieving anything productive with the aid. The egoistic interests of donors often works against the altruistic objectives of aid assistance and might lead to corruption of the aid projects’ outcomes, reducing the effectiveness of aid interventions (e.g., Lancaster, 1999, 2006; Easterly, 2006; Lindsay Whitfield, 2009).

This study attempts to argue against the PCP approach and its presupposed theory of aid ineffectiveness, by presenting an alternative explanation of aid programs (in) effectiveness. As discussed in previous chapters, one purpose of this dissertation is to examine the prevalence of the “Local Knowledge Syndrome” as a potential factor in the low effectiveness level of foreign aid programs for promoting the public sector reforms in recipient countries. The proposition is formulated thus:

(I) Aid programs and projects could certainly work and achieve satisfactory results in host countries with poor governance and a complex set of multiple development goals. Nevertheless, aid practitioners in the field and those in charge of development policies at the highest institutional levels should be in touch with reality and incorporate “local knowledge” in the project design and implementation process, as well as adopting delivery mechanisms to insure effectiveness.

Chapter 3 defined the concept of LKS and its implications for foreign aid effectiveness. LKS is conceptualized as the lack of adequate knowledge of social, cultural and political aspects of a society in which aid programs act. This syndrome may affect the processes and the methods by which development aid programs and projects are formulated and implemented. The ability of development planners to predict and control the outcomes of their programs under conditions of uncertainty is quite limited. International assistance agencies use rationalist planning and management procedures, which often require knowledge and data that are simply not available in most developing countries. To reduce “uncertainty”, the demands of rationalistic planning have forced administrators to use whatever data is at hand, regardless of its appropriateness or accuracy1.

To deal with factors influencing the process of incorporating local knowledge during policy making and implementation, practitioners and policy makers have to increase the success of donor-promoted public sector reform by taking into account some of the obvious dimensions by which recipient countries, or host organizations

---

1 - Before the end of this stage, we intend to review some aspects of the local knowledge in Yemen (Chapter 4) to be used as a background for the analysis in the coming chapters.
within these countries, can be distinguished from each other. It is also crucial to avoid any reform attempts or aid interventions based on components copied from elsewhere, or to change the ways in which specific components of host organizations function along lines which seem successful elsewhere. Accordingly, this study is mainly concerned with testing two main hypotheses that can elaborate how LKS could substantially contribute to the failure of development programs:

(I) Incorporation of local knowledge in the policy design stage, that integrates formal and informal institutions and organizations in host governments or organizations, increases the effectiveness of aid programs. Reduced incorporation of local knowledge based on this integration increases ineffectiveness of aid programs.

(II) Incorporation of local knowledge during the implementation stage leads to greater openness to local realities not foreseen in the design stage and reduces the probability that aid programs or projects need redesigning, and increases their effectiveness.

The independent variable is drawn from the two main hypotheses of the study, based on the theoretical paradigm presented in Chapters 2 and 3. Thus, the independent variable is the process of incorporating local knowledge in policy design and implementation of donor-prompted public sector reform programs. The dependent variable is linked to the outcomes of aid programs. However, we could achieve different results or outcomes, ranging from projects results, partial success (outcomes falling short) to programs, that failed in their entirety to achieve any significant outcomes.

Before explaining the empirical methodology of how we deal with the main argument of this dissertation, it is important to keep in mind the following issues. First, this study defines the concept of the aid project by making use of two approaches, the ‘development associability approach’ and the ‘validity development approach’. The latter attributes the success of development projects to the inherent value of the project. In other words, it draws attention to the connection between success and aims (Scott-Smith, 2013). The former focuses on the process more than the outcomes (see section 2.2.2 of this study). It is critical to note that for one or more of the aid

---

2 - It is the main idea of the actor network approach (see Chapters 1 and 3) as it was applied in this study, which describes human interaction, providing a better understanding for aid programs involving actors on how to incorporate local knowledge. In this regard, it offers a practical way for dealing with the LKS during aid policy design and/or implementation.
projects under investigation in this study, the generative mechanisms of LKS may change as a result of some degree of changes in the implementation process. Secondly, place and time knowledge can differ among aid recipients, aid organizations and individuals, which influences the process of designing and implementing aid programs and projects. Thirdly, perhaps the policy design of aid programs, which was previously inadequate, was promptly corrected or incorporated local knowledge during the implementation stage of aid programs.

This chapter deals with methodological issues. It shows how I undertake the analysis and investigation in order to test the above-mentioned main hypotheses. The chapter provides a description of the research design, population and sample, instruments, data collection procedures, data analysis methods, validity issues and the problems that emerged conducting the research. In the final section, a summary is presented.

5.2 The Research Design

The research design started by setting up the first empirical response to the PCP approach using the case of the Dutch-Yemen development co-operation, in order to support or reject the corresponding hypothesis. By providing a brief historical background of the Netherlands development aid history, determinants and its role in Yemen; the empirical results support the hypothesis that the existence of multiple egoistic goals of aid assistance from a donor country to a recipient country, may not be the main reasons for ineffective aid interventions in poor nations- which was mainly argued by the Public Choice Perceptive (PCP) approach. The strategic interest of the Netherlands in aiding Yemen is to achieve security and stability in the country. The Netherlands has a long history of funding development programs in Yemen, especially in the focal sectors such as water, health, and basic and higher education, with a focus on cross-cutting themes like gender, governance and capacity building. We assess if development efforts can stimulate improvements in terms of development outcomes that have a meaningful impact on security and stability in Yemen in Chapter 6.

Some projects in poor policy environments will have consequences for the resilience of the analysis based on the second argument of the PCP approach, which states that foreign aid does not work in a poor policy environment. By using the example of the Dutch-Yemen NPT programme, our empirical work explains the “aid programme effectiveness” dependent variable as fully as possible, by controlling at
both policy design and implementation levels all independent variables that may positively affect the process of incorporating local knowledge and aid program effectiveness. At the same time, the investigation includes other explanatory variables that may reduce aid program effectiveness and increase the LKS in a poor policy environment.

It should be made clear, that this study seeks to gain insights on the influence of a single causal variable (incorporating local knowledge in policy design and policy implementation) on the outcomes of interest (the successes or failures of aid programmes in Yemen). To investigate this causal relationship, it is logical to adopt a research design that includes some cases which clearly reflect the dependent variable of this study: successful and failed aid projects. This means that the study employs a mixed study method, including a qualitative and comparative case study method to show which aid projects succeeded and which projects failed. This leaves room for local knowledge explanations and arguments against the explanation of the PCP approach, as we mentioned before.

Yin (1994) identified three types of research projects that may use the case study methodology. First, there is the exploratory case study, which intends to develop theory. Then there is the explanatory case study, which intends to test a hypothesis by demonstrating its applicability either in specific circumstances or in general. The third type, the descriptive case study, refers to unstudied situations under a well-theorised area. This study has been designed on the basis of explanatory purposes as long as there is a hypothesis based research. Furthermore, an extensive theoretical background (Patton 1990) does not support the subject under investigation. It is expected that testing the study hypotheses will, in some scope, bring hope for theoretical improvement (in an inductive manner). Extensive preliminary work is needed to get sufficient understanding of the research problem before we can proceed to the hypothesis testing stage. The hypothesis testing stage is where the researcher examines whether or not the hypothesized relationships have been substantiated (Sekaran, 1992, as cited in Elbana, 2008).

Four aid projects or case studies (n=4) are researched and compared in this thesis. All four projects are funded by the bilateral3 donor agency Nuffic, which is the

---

3 In the context of foreign aid, we have two main channels of aiding poor nations. The first is the bilateral channel: the individual donor country. The major bilateral donors in volume terms are the United States, Japan, France, Germany, Britain and the Netherlands, followed by Italy, Canada, Sweden, Norway, and Spain. The second is the multilateral aid organizations, such as the World Bank, the United Nations, The International Monetary Fund (IMF), the European Union (EU) and the Organization for Economic Cooperation and Development (OECD), as well as many non-governmental organizations (NGOs).
Netherlands organization for international cooperation to promote the capacity of post-secondary organisations in Yemen. Nuffic works on behalf of the Netherlands bilateral aid Programme for Institutional Strengthening of Post-secondary Education and Training Capacity (NPT) in developing countries. The overall objective of the NPT program in Yemen is to strengthen the teaching learning conditions, the higher education system, and the institutions which are part of the system. It aims to raise the capacity and quality of higher education (HE) to provide services for promoting border capacity development within the Yemeni public, private and non-governmental sectors (Visser and Almoassib, 2008:1). As we mentioned before, the main reason to select The Netherlands from the donor countries is that it is the second biggest bilateral donor in education sector reform in Yemen. Furthermore, I selected the education sector reform in particular, because more than 40% of foreign aid is going directly to the education sector at all levels (MOPIC of Yemen, 2009). The projects under investigation are:

(1) Establishment of an Executive Masters in Public Administration (MPA) degree at the Faculty of Commerce, Sana’a University. The project was implemented by the Dutch Institute for Public Administration (Dutch acronym ROI), the School of Public Administration of Leiden University, the Department of Political Sciences and Public Administration at Sana’a University, and the National Institute of Administrative Sciences, Yemen. The overall aim of the project was to assist the Government of Yemen in the process of restructuring its public administration through increasing the capacity and capability of the higher education sector in Yemen. The specific objectives of the project were: (1) to have a viable Executive Master Degree programme in Public Administration, for senior and mid-level civil servants at management positions. The MPA was designed to offer the opportunity for civil service managers in the various public administration institutions, to obtain the necessary skills to help them in further strengthening the organization and management of their respective departments. (2) To have a renewed Institute of Administrative Sciences (NIAS) in charge of the development and implementation of the human resource training within the government structures at central, regional and local level.

(2) Establishment of an executive Masters in Business Administration at the Faculty of Commerce, Sana’a University. The Maastricht School of Management, the Institute of Social Science (ISS) and Sana’a University implemented the project. The project targeted the development of an MBA delivery capacity at the Department of Business Administration of Sana’a University. It met the need for management education in
Yemen, as it was evident that there is no comprehensive in-country training available to cater for the emerging private sector needs. The project envisaged the establishment of an MBA programme, through the strengthening of the staff base, supporting the development of the curriculum and educational infrastructure. In parallel, institutional development was foreseen, gearing up the Department to operate in a market-driven environment.

(3) Developing a diploma programme in Gender Studies at the Women’s Research and Training Centre (WRTC) of Aden University. The project was funded by Nuffic and implemented by Wageningen University and Aden University from 1 July 2004 to 31 August 2008. The overall objective of the WRTC project was to contribute to the improvement of the status of women and their participation in the development of Yemen, by increasing the employment of women, their occupation of management positions, women’s participation in politics, fostering growing readiness to address gender issues in policy making at various levels of the Yemeni government, and a greater knowledge of how to do that by way of consultancies and dissemination. The specific objectives of the WRTC project were to strengthen the WRTC to become a sustainable organisation that had the capacity and capability to (i) teach and train present and future gender practitioners; (ii) conduct consultancies on gender issues to organizations and institutions engaged in gender and development; and (iii) manage an information and communication system for distribution of relevant materials to organizations, institutions and individuals on gender and development issues.

(4) Strengthening the Water and Environment Centre of Sana’a University and its Programme in Integrated Water Resource Management (WEC). The project, from August 2004 to July 2008, was funded by Nuffic and implemented by seven actors: the Water and Environment Centre at Sana’a from the recipient side and Wageningen University, Delft Technical University, Cairo University, MetaMeta, Arcadis-Euroconsult and WaterWatch from the donor side. The overall idea and objective of the WEC project was to support the reform process of Yemen’s major water and environmental sector organisations, mainly in human resources development. This was needed to strengthen their institutional and technical capabilities in order to facilitate effective water resources management in the country. The WEC was undertaken to help alleviate the water crisis in Yemen and to help achieve ‘water security’, which has been identified as Yemen’s second priority next to national security. The specific objectives of the project were to: promote the capacity of WEC-
SU to be a sustainable centre providing diploma course in IWRM for mid-level managers and policy decision-makers; to create an MSc course in IWRM with different specializations for BSc students; and to foster high-level consultancy services that will act as an information centre on IWRM developments. The emphasis in all of this was on the sustainability of the centre in providing these services, which meant that a sound business approach was necessary. In other words, the project was not only about education but also about creating an entity that is able to continue to provide services – and even expand these - beyond the closing date of the project. Strengthening skills in the water management of the Yemeni government has become very important. There is a severe lack of capacity in water resource management, planning and participatory development – but also in some of the supporting technical fields. In developing this program, it was proposed to liaise closely with employers of the course participants, such as government staff, to ensure that the course fulfilled the requirements. It was also proposed to include a practical element in the training course inviting mid-level managers to study or to review a real-life problem related to IWRM in their organizations (Wageningen University, 2004).

This research is an empirical enquiry and theoretically-guided examination of selected small-N cases. John Gerring (2007) argues that in case study methods with small samples the randomization in selecting cases is problematic. He argues that the cases will be unreliable and the researcher will not avoid the selection bias many scholars have warned against (see King et al, 1994; Achen and Snidal, 1989; Collier and Mahoney, 1996). The author’s proposed solution is to use purposive (non-random) selection procedures that help to pick cases that are appropriate to particular research objectives. Consequently, the selection of cases in this study was guided by what is known as a Most Similar Case Selection strategy, exemplifying and vindicating the case study methodology used in this study (Przeworski and Teune 1970; Gerring, 2007). According to the basic logic of this method, the cases under study generally deal with similar circumstances, but have experienced significantly different outcomes. This logic of research design allows us to select cases based on the dependent variable (success or failure of aid programmes) and then working backwards to see if the independent variable (incorporating local knowledge) has the right or predicted value of the dependent variable (Cook and Campbell, 2002; Tarrow, 2010: 244):
Table 5.1 Research Design

<table>
<thead>
<tr>
<th>Funding Agency-NL</th>
<th>Host Public Sector-YEM</th>
<th>Direct Donor Contractor</th>
<th>Direct Host EC(s)</th>
<th>Aid Projects</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>NL-Nuffic</td>
<td>YEM-HEM</td>
<td>ROI &amp; Leiden Univ</td>
<td>Sana’a Univ MPA (CPAD) &amp; NIAS</td>
<td>Establishment of an executive Masters in Public Administration (MPA) Degree at Sana’a Univ</td>
<td>Failure</td>
</tr>
<tr>
<td>NL-Nuffic</td>
<td>YEM-HEM</td>
<td>MSM &amp; ISS</td>
<td>Sana’a Univ MBA (CBA)</td>
<td>Establishment of an executive Masters in Business Administration (MBA) Degree at Sana’a Univ</td>
<td>Success</td>
</tr>
<tr>
<td>NL-Nuffic</td>
<td>YEM-HEM</td>
<td>Wageningen University</td>
<td>Sana’a Univ WEC</td>
<td>Strengthening the Water and Environment Centre of Sana’a University and its Programme in Integrated Water Resource Management (WEC)</td>
<td>Success</td>
</tr>
<tr>
<td>NL-Nuffic</td>
<td>YEM-HEM</td>
<td>Wageningen Univ</td>
<td>Aden Univ WRTC</td>
<td>Developing a diploma programme in gender at Women’s Research and Training Centre(WRTC)</td>
<td>Failure</td>
</tr>
</tbody>
</table>

The research design provides for a sufficient basis to test the hypotheses, as it has effectively controlled for country-donor side-affects and reduced many entries. The case study includes two projects that achieved desirable outcomes, as specified by the evaluation records; and two other projects which have not achieved the desirable development results and impact, as was specified by the evaluation records and the researcher’s participant observations. The viability of these cases lies in the fact that they were implemented by the direct beneficiaries: governmental bodies and higher education organizations. Moreover, there were two programmes (MBA and MPA) implemented at the same sub-organization, namely the Faculty of Commerce and Economy at Sana’a University, of which the MBA was more successful. The same argument applies to the donor side: the Dutch Nuffic aid agency was responsible for designing and selecting the Dutch contractors to implement all the selected aid programmes. Furthermore, the majority of Dutch contractors were organizations in higher education. They have either individually carried out these aid programmes or in cooperation with other Dutch specialist organizations within the Dutch higher education system. This confirms that the research population has similar aspects, as we have adapted similar case selection method as a prerequisite of small-n design, including smaller government organizations selected from one policy area or one
public sector, namely Yemen’s higher education sector.

The selected cases share a number of important similarities and they are sufficiently comparable to allow for a reasonably controlled comparison. The projects have either ended or are in their final stages, thus making it possible to study causal explanations. The projects allow for a complete picture of the various stages. In other words, they are recent enough for the respondents to remember the project, but sufficiently advanced (or even completed) to allow for a proper analysis. These four cases were implemented approximately in the same time-frame, a ten-year period from 2003 to 2011. They all focus on public sector reform, especially strengthening and upgrading capacity of the local education organizations and institutions; an effort that is also referred to as ‘institution building’.

5.3 Forms of Enquiry

To gain an understanding about the effectiveness or failure of the projects in achieving their overall and specific objectives, a survey questionnaire was used as the main data source of evidence. In this stage of enquiry, we adapted a quantitative method for the analysis. The quantitative analysis allows the researcher to meet the first research objective and provides important links making conclusions about the effectiveness of the selected projects. The question why two cases might be perceived as more effective than the other two is left to the second stage of quantitative analysis, in Chapters 7 and 8. This second stage of quantitative analysis describes the extent to which the selected projects were able to build capacity in the targeted organizations of the higher education sector. This is in order to determine whether these projects promoted public sector reform or achieved the overall objective of the Dutch NPT program in Yemen (see Chapter 9).

The third stage of our enquiry investigates why two cases were more effective than the other two in building capacity in the Yemeni higher education sector. This question is investigated through qualitative analysis, using the LKS explanation as the independent variable that can vary among the selected projects. At this stage, the underlying causes of the barriers and constraints that limit the effective incorporation of local knowledge in policy design and/or implementation of aid projects are consider. Multiple sources of evidence will be employed for this purpose; mainly primary sources such as key qualitative and semi-structured interviews and participant observation, as well as secondary sources (see Chapters 9 and 10):
This study adopts a mixed methods analysis. The roots of mixed methods can be traced back to the multi-trait, multi-method approach (Campbell and Fiske, 1959, cited in Teddlie & Tashakkori, 2009, p.31). It is considered a relatively new methodology, which has evolved its key philosophical and methodological foundations and practice standards since the early 1990s (Harwell, 2011: 151). The mixed methods research merges qualitative and quantitative methods and draws on the strong aspects of both traditions (Harwell, 2011). Therefore, the qualitative case study approach is critical to this dissertation, because it employs an intense investigative process that compares, classifies and contrasts events and subjects to provide decision-makers and policy-makers with the information needed to improve program implementation and policy choices. Qualitative approaches are generally useful to study the particular context in which the participants act and the influence this context has on their actions. The qualitative approach is useful in interpreting not
only the physical events and behavior that is taking place, but also the way the
participants make sense of the events and how their understandings of their mission
influence their behavior in policy design and implementation of aid programmes
(Bredo & Feinberg, 1982; Geertz, 1973; Rabinow & Sullivan, 1979, in Maxwell, 2008).

Qualitative approaches typically study relatively small-N case studies of individuals
or situations and preserve the individuality of each of these in their analysis, rather
than collecting data from large samples and aggregating the data across individuals
or situations. Thus it enables us to understand how events, actions, and meanings are
shaped by the unique circumstances in which these occur (Maxwell, 2005:22). In
other words, qualitative approaches can help to develop a causal explanation for what
outcomes of aid programmes occur under different values. Although qualitative
research is not unconcerned with outcomes, a major strength of qualitative studies is
their ability to target the causal inference that leads to these outcomes (Cook, &
Campbell, 2002: 389). In this study, the main concern is to investigate the causal
effect of incorporating local knowledge in outcomes of selected aid programmes
rather than simply demonstrating regularities in the relationships between variables
(Maxwell, 2005 & 2008).

The research is empirical in that direct observation of the world is seen as the way
to generate and validate truths. Only through ongoing interaction between theory
and data can we expect to generate a knowledge base that is solid and enduring.
Qualitative analyses tend to be grounded in the empirical world and the problems
apparent in the empirical world often require different approaches in order to gain
understanding and reach conclusions (Patton 1990). Furthermore, “quantitative
methods of data analysis can be of great value to the researcher who is attempting to
draw meaningful results from a large body of qualitative data. The main beneficial
aspect is that it provides the means to separate out the large number of confounding
factors that often obscure the main qualitative findings.” (Abeyasekera, 2008:2).

The focus of the research is on the process, implementation and development of a
program and its participant’s qualitative approaches are used. To complement the
qualitative approach, a quantitative evaluation was used for evaluating the effectiveness
of the selected projects, i.e. the dependent variable, in capacity and capability building
processes in a specific skill area. This was done by a survey among the direct
beneficiaries of the aid projects within the Yemeni government and the private sector.
Additional quantitative measure indicators supplement the survey to analyze whether
foreign aid projects succeed in poor policy environments, which is a theoretical
challenge to the PCP approach.

The justification for mixed methods research in this study tallies with the assertion of Johnson and Turner (2003), that the fundamental principle of mixed methods research is that multiple kinds of data are collected with different strategies and methods in ways that reflect complementary strengths and non-overlapping weaknesses; thus providing insights not possible when only either qualitative or quantitative data are collected. Put another way, mixed methods research allows for the “opportunity to compensate for inherent method weaknesses, capitalize on inherent method strengths, and offset inevitable method biases” (Greene, 2007: xiii as cited in Harwell, 2011: 152).

5.4 The Sources of Data

As was mentioned earlier, the adopted mixed methods research merges qualitative and quantitative methods. Therefore, primary data was collected for the literature review, by applying a cross-sectional data collection method using survey questionnaires. The data collection was arrived at through the secondary data method and the primary data method. The first method includes published data such as papers, documents and other literature, for example reports related to the NPT program and its related projects, and research published online. The primary method is related to the empirical work and includes three main streams: questionnaire, interview, and observation:

Figure 5.2: Data Collection for Present Research
5.4.1 The Secondary Sources

Most of the secondary data gathered for this research originates from the selected policy setting (the higher education sector in Yemen) and the organizations involved in the study, as well as from publications of the Nuffic and MinBuza. For secondary data, it is necessary to recognize, categorize and analyze all documentation related to the selected projects, including: policy declarations and strategic plans for donor and recipient, project documents, funding proposals, progress and evaluation reports, press releases, archival records, memoranda and timetables.

For an understanding of the donor perspective, I first analyzed those documents that identified and described the objectives of aid programs at all levels, including the budget justification. Project documents are produced by donor agencies, often in cooperation with the recipient government. These documents describe objectives, a plan of implementation, management arrangement, time frame, terms and conditions and confirm expected outcomes. The middle agency donor (Nuffic) or the project directors, mostly the implementers or the contractors of the aid project, deliver progress reports and evaluation reports. To understand the recipient side, I gathered information by looking at the strategic plans for improving higher education sector and development needs. The funding proposals from the relevant government institutions or ministries were analyzed and compared to the initial objectives, plans and expected outcomes.

Looking at the secondary data is important to provide blueprints and evidence on what was planned and intended. It is also important to consider the way LKS is incorporated through institutional landscapes and organizations of public sectors that aid programs aim to reform. Documentation of aid programs can also help us to ensure that we have selected the relevant staff for conducting interviews.

5.4.2 The Primary Sources of Data

• Semi-structured Interviews

To provide an insight into what was actually happening during the process of implementation, considering the matter of incorporating local knowledge, this study investigates aid policies, practices and management, mainly through the experience of the people involved in planning, design and implementation. Qualitative interviews provide the researcher with the opportunity to have an intimate, recurring and prolonged involvement in the life and community of the respondents. Typically, the interviews are conducted with an unstructured or minimally structured format. The
questions tend to emerge as the researcher is sensitized to the meanings that the participants bring to the situation (Mertens, 1998: 321). For the interviews in this study, the researcher chose to provide some structure to assist with the gathering of specific knowledge, pertaining to the barriers and constraints of incorporating local knowledge in policy design and implementation of the selected aid projects.

Qualitative interviews are conducted with different categories of participants to capture a range of relevant viewpoints, at all levels of the administrative process of aid programs. The first category of people are actors who were responsible for setting up policies and aid program plans at the policy making level, to discover if and how they incorporated local knowledge to execute their tasks. The second category consists of participants from the middle management level who carried out the design and implementation of aid programs. The third category includes key participants who were responsible for the actual execution of aid programs. They are the players at the heart of the aid game: intermediaries who are connected to recipient countries, donor countries, agencies, and beneficiaries. Their comparative advantage is to solve the local knowledge problem.

The participants selected for the donor's perspective, included officials from top and middle management who represent the donor organizations in Yemen and who were responsible for setting and implementing aid policies. In addition, in order to be able to measure process directly against results, we included experts, consultants, field officers, coordinators and trainers from donor organizations to allow us to triangulate the collected data, and capture the different aspects of the intervention throughout the process of design and implementation of selected aid programs.

From the side of the recipient government, participants were selected from all administrative levels, including senior government officials and policy makers from the related ministry and institutions, top and intermediate management officers, experts, consultants, coordinators, field officers and workers. Furthermore, project directors, contract and direct-hire staff were also selected, in addition to the private sector companies that were directly involved in the projects and grassroots beneficiaries.

The actual number of participants selected depended mainly on the scale and complexity of the particular aid programs. For example, in some cases the donor or recipient had only one person responsible for project management and design. We tried to conduct the first interview with the project director, who in turn helped us to identify the other managers, consultants and personnel involved in the process of
designing and implementing aid programs. To make the process of conducting interviews clearer and easier, I started the interview process with one pilot case study from the selected cases. The idea behind this was to ensure that all those involved with project design and implementation processes were interviewed.

To assess the reliability of the information that was collected from the interviews, the following points were considered. Firstly, to ensure that the direction of a certain topic was not lost, and to prevent the interviewee from digressing to a different topic not relevant to the objective of the interview, the researcher developed an interview guide to help the interviewees understand the main argument of the research. Secondly, the researcher used the content analysis structure to evaluate the data. The expert interviewees were informed and the common representative statements, secondary data, relevant structures, constructs of reality and interpretations were thus analyzed. Thirdly, the study analyzes similarities and differences through a so-called thematic comparison based on typical statements, and it stresses thematic focus (Katharina, 2009: 26).

There are other steps to ensure the reliability of the information collected from the interviews. As we mentioned before, the interviews on the donor side dealt with the different policy levels: policy-making, middle management, and operational level. At every level, questions were asked about the role of the interviewee and their evaluation of the other management levels in the Dutch context. For additional checks, the donors were asked about the role of the participants in the whole process. These different checks served as a technique to avoid any ethical issues, as asking the participants about their role might often prove sensitive. The respondents were given the space to speak freely about their ideas on the subject matter. This procedure was also followed for the recipients. In addition, two conferences, five group discussions and six expert meetings were organized. I organized and arranged the raw empirical material such as photos, observations, recorded interviews and transcriptions by writing memos of my thoughts and by defining, categorizing and coding data and constructing meanings. All of this was in line with the working procedures of the theoretical perspective in my study (Charmaz, 2003, as cited in Voeten, 2012). I worked in the academic year 2008-2009 as a teacher in the MPA project and was also involved with some activities in the MBA project as they are in the same building. This increased my understanding of the training processes in these projects, and provided me with reliable information about the problem that I am dealing with.
• Observation

Qualitative observation should take place in naturalistic settings without predetermined categories of measurement or response (Patton, 1990). The data collected through observation should be regarded reliable and valid only if the methods for collecting them are value neutral (Kawulich, 2005). The level of participant observation is what is defined in Mertens’ text as moderate participation, whereby the researcher attempts to balance the insider and outsider roles by observing, and participating in some, but not all of the activities (Mertens, 1998: 317-318). Through the use of observational reporting, the researcher will be able to cross-check findings and eliminate inaccurate interpretations. Through observations in different settings, at different times of the day and year, the reliability of the data will be improved. This is in line with the suggestions made by Adler and Adler as presented in Mertens (1998).

In the present study, the observation constituted a very important method of collecting data, as the researcher was part of the plan for implementation of at least two of the selected projects as case studies (MPA and MBA) at Sana’a University. This offers greater opportunity for the researcher not only to adopt a direct observation approach during moments of fieldwork, but during the whole process. It is a structured method that is pest described, and valuable for mapping any informal organizations and institutions of a specific time and place. In addition, the nature of the problem regarding local knowledge and the effectiveness of foreign aid needs to be studied by a researcher who is preferably familiar with such local knowledge. The fact that I am a Yemeni researcher with more than six years’ experience working as a teaching assistant in the political science department at Sana’a University, is very advantageous to the study. My cultural and social background opens up greater opportunities to gain information about the formal and informal institutions and organizations in a specific place and time.

• Questionnaire

The questionnaire is a convenient and an efficient technique of gathering data and measuring the impact of the selected projects in reality. It is difficult to use a qualitative technique to measure the impact of selected projects, because of the large population in the selected cases involved and the difficulty to control or cover all target groups. The written questionnaire is typically more efficient and practical and allows for the use of a larger sample, a method that is widely employed in educational research (Ary, et al, 1972). The questionnaire was designed to obtain specific data to measure the level of
effectiveness of the selected projects in achieving their overall and specific objectives. For this purpose, it was divided into four categories: (I) information related to the demographic profile and work background (II) Evaluation of the capacity building projects in the education sectors (III) Evaluation of the transfer of training skills to the work setting. (IV) The level of effectiveness. This section (IV) could reflect the obstacles that these Yemeni higher education organizations face in the process of capacity building of their participants and in the process of transferring the trained skills back to the respective organizations.

It is difficult to fully claim that we have ready and available variables that are useful to measure the level of effectiveness of the selected aid projects in achieving the overall objective of the NPT program in Yemen. Instead, we used variables or proxies that explicitly capture a fragment of the level of effectiveness achieved through these aid projects for reforming the public sector.

One variable of the survey is the demographic profile of the participants, which provides critical information about institutions. The analysis of the demographic profile shows something about the success and failure of the selected project in accomplishing the overall objectives of the NPT program. The enrollment process of participants (students) should be based on the priorities of the human development plan, identifying its vision and goals of “quantity” in promoting the public sector reform agenda. This human development plan should be responsive to the needs of current and future public sector reforms agenda. Based on the literature reviewed in earlier chapters of this study, and the researcher’s own knowledge and experience, the project will better reflect the local need for public sector reform (1) when there is orientation towards the promotion of female enrolment, (2) when there is a geographical and an organizational prevalence in the participation process across different organizations and cities ,(3) when the enrolment of the participants is primarily focused on the age group 30-35 years and below, and (4) when the enrolment of the participants focuses on the low-career and (mid) level civil servants to develop a new breed of leaders and managers.

Another variable is capacity building of the participants. This variable can be linked to the first variable, because as we select appropriate participants, the chance of developing their capacity and capability will be higher and vice versa. In Chapter 4 we used the “ideal model” for capacity building in public sector in Yemen, that reflects the managerial skills that are needed to develop the capacity within human resources in Yemen. The managerial skills introduced at the end of Chapter 4 helps to evaluate the extent to which the selected project will effectively develop the capacity of the
participants. The managerial skills are: (1) Strategic management (2) Human resources (3) communications and (4) leadership skills.

The effectiveness of this training is related to the status of the strategic transfer of skills back to the office (Kirkpatrick, 1967). In this respect, the third variable has been developed to measure the efforts made by the selected projects to transfer the developed skills of their participants back to their work settings. Several predications are used as proxies to measure this variable, and are included in the following table:

**Table 5.2: The Developed Variables and Measured Items to Evaluate the Level of Effectiveness of the Aid Projects in Achieving their Overall Objective (Promoting the capacity of the Public and Private Sectors)**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Measured items</th>
<th>Alpha (number of items in scales)</th>
<th>Place in the questionnaire</th>
</tr>
</thead>
</table>
| 1-The demographic profile of the participants and the requirements for a capacity and capability building process in Yemen | • Involvement of female participants  
• Involvement of participants from outside the main cities  
• Involvement of young participants  
• Involvement of participants with low-level jobs | 0.682 (4 items) | Demographic Profile Section |
| 2- The achieved level of capacity and capability building among the participants (Managerial skills developed) | 1. Strategic management skills developed:  
• Developing and analyzing policies for your organization  
• Developing a strategic plan for your organization  
• Developing a Plan of Action and a basic budget based on the strategic plan  
• Developing indicators to measure the successful implementation of the strategy  
• Organizing the organization according to the strategy | 0.886 (5 items) | Section One |
|                                                                          | 2. Human resources management skills developed:  
• Defining the functions and job descriptions of the organization  
• Introducing and implementing an evaluation system of functions and jobs  
• Setting criteria for evaluating the performance of the employees in the present situation  
• Developing an employees’ upgrading programme  
• Identifying the shortcomings of the present employees to fulfil new jobs and functions  
• Setting up and implementing a monitoring and controlling system of change processes  
• Establishing reference group to discuss and defend your visions within the organization  
• Establishing a reference group to discuss and defend your visions with the beneficiaries  
• Making and executing effective and participatory decisions  
• Introducing and maintaining ethics and values such as transparency and accountability | 0.931 (10 items) | Section One |
### 3. Communications skills developed:
- Setting up and implementing a communication system within the organization
- Developing meeting skills
- Setting up a reporting system
- Developing reporting skills

### 4. Leadership skills developed:
- Getting a better understanding of English
- Promoting technology literacy
- Reading and comprehending information
- Being creative in solving problems
- Being integrative in thinking
- Being a critical/analytical thinker
- Setting up and implementing a reward penalty system within your organization
- Managing conflicts and effective dispute resolution
- Stimulating and implementing team building exercises

### 1. The selection process of the participants
- The method of involving in the Education Centres (ECs) as participants
- The method of paying the fees of the master degree
- Involving in the training process as part of the upgrading plan of the related organization
- The competition policy of selecting the participants within the specific organizations
- The criteria of selecting the participants by the related organizations
- The criteria of selecting participants by the selected ECs.

### 2. The teaching methods used
- Lecture technique
- Case Study
- Role Playing
- Business Games

### 3. The relevance of the curriculum and course training for the Yemeni context
- Overlapping in course materials
- The courses are relevant to the special needs of the Yemeni related sector
- The master degree is focusing on how the manager should manage rather than what management is in general
- Good balance between theory and practice
- Case studies presented in the courses are based on the daily practice of the Yemeni sectors

### 4. The application of training knowledge and skills in the job-setting:
- Describe your position in your organization before attending the education center.
- Describe your position in your organization after attending the education center
- Promotion of function in an organization.
- Receiving a salary increase after obtaining the master degree.

### Table 1: Variables Measured items

<table>
<thead>
<tr>
<th>Variables Measured items</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receiving a salary increase after obtaining the master degree.</td>
<td>0.884</td>
</tr>
<tr>
<td>Setting criteria for evaluating the performance of employees to fulfill new jobs and functions</td>
<td>0.932</td>
</tr>
<tr>
<td>Setting up and implementing a reward penalty system within your organization</td>
<td>0.707</td>
</tr>
<tr>
<td>Using the skills, knowledge, and competencies to manage rather than what management is in general</td>
<td>0.886</td>
</tr>
<tr>
<td>Introducing and implementing ethics and values such as transparency and accountability</td>
<td>0.682</td>
</tr>
<tr>
<td>Introducing and maintaining ethics and values such as transparency and accountability</td>
<td>0.931</td>
</tr>
<tr>
<td>Making and executing effective and participatory controlling system of change processes</td>
<td>0.837</td>
</tr>
<tr>
<td>Developing a Plan of Action and a basic budget</td>
<td>0.932</td>
</tr>
<tr>
<td>Developing and analyzing policies for your organization</td>
<td>0.708</td>
</tr>
<tr>
<td>Defining the functions and job descriptions of the employees in the present situation</td>
<td>0.771</td>
</tr>
<tr>
<td>Setting up and implementing a monitoring and evaluating system</td>
<td>0.884</td>
</tr>
<tr>
<td>Setting up and implementing a communication system within the organization</td>
<td>0.886</td>
</tr>
<tr>
<td>Developing meeting skills</td>
<td>0.884</td>
</tr>
<tr>
<td>Setting up a reporting system</td>
<td>0.932</td>
</tr>
<tr>
<td>Developing reporting skills</td>
<td>0.707</td>
</tr>
<tr>
<td>Getting a better understanding of English</td>
<td>0.886</td>
</tr>
<tr>
<td>Promoting technology literacy</td>
<td>0.884</td>
</tr>
<tr>
<td>Reading and comprehending information</td>
<td>0.932</td>
</tr>
<tr>
<td>Being creative in solving problems</td>
<td>0.707</td>
</tr>
<tr>
<td>Being integrative in thinking</td>
<td>0.886</td>
</tr>
<tr>
<td>Being a critical/analytical thinker</td>
<td>0.884</td>
</tr>
<tr>
<td>Setting up and implementing a reward penalty system within your organization</td>
<td>0.932</td>
</tr>
<tr>
<td>Managing conflicts and effective dispute resolution</td>
<td>0.707</td>
</tr>
<tr>
<td>Stimulating and implementing team building exercises</td>
<td>0.886</td>
</tr>
<tr>
<td>The method of involving in the Education Centres (ECs) as participants</td>
<td>0.708</td>
</tr>
<tr>
<td>The method of paying the fees of the master degree</td>
<td>0.932</td>
</tr>
<tr>
<td>Involving in the training process as part of the upgrading plan of the related organization</td>
<td>0.707</td>
</tr>
<tr>
<td>The competition policy of selecting the participants within the specific organizations</td>
<td>0.886</td>
</tr>
<tr>
<td>The criteria of selecting the participants by the related organizations</td>
<td>0.884</td>
</tr>
<tr>
<td>The criteria of selecting participants by the selected ECs.</td>
<td>0.932</td>
</tr>
<tr>
<td>The role of education centers for promoting the workforce in order to boost training back to the workplace in order to promote the public sector reform process</td>
<td>0.707</td>
</tr>
<tr>
<td>The demographic profile of the participants and the profile of the general employees</td>
<td>0.771</td>
</tr>
<tr>
<td>Involvement of participants with low-level jobs</td>
<td>0.886</td>
</tr>
<tr>
<td>Involvement of young participants</td>
<td>0.884</td>
</tr>
<tr>
<td>Involvement of female participants</td>
<td>0.932</td>
</tr>
<tr>
<td>Involvement of participants from outside the main Yemeni sector</td>
<td>0.707</td>
</tr>
<tr>
<td>Setting up and implementing the Education Centres (ECs) as participants</td>
<td>0.932</td>
</tr>
<tr>
<td>The method of paying the fees of the master degree</td>
<td>0.707</td>
</tr>
<tr>
<td>Involving in the training process as part of the upgrading plan of the related organization</td>
<td>0.886</td>
</tr>
<tr>
<td>The competition policy of selecting the participants within the specific organizations</td>
<td>0.884</td>
</tr>
<tr>
<td>The criteria of selecting the participants by the related organizations</td>
<td>0.932</td>
</tr>
<tr>
<td>The criteria of selecting participants by the selected ECs.</td>
<td>0.707</td>
</tr>
<tr>
<td>The relevance of the curriculum and course training for the Yemeni context</td>
<td>0.886</td>
</tr>
<tr>
<td>The application of training knowledge and skills in the job-setting:</td>
<td>0.932</td>
</tr>
<tr>
<td>The role of education centers for promoting the workforce in order to boost training back to the workplace in order to promote the public sector reform process</td>
<td>0.707</td>
</tr>
<tr>
<td>The demographic profile of the participants and the profile of the general employees</td>
<td>0.771</td>
</tr>
<tr>
<td>Involvement of participants with low-level jobs</td>
<td>0.886</td>
</tr>
<tr>
<td>Involvement of young participants</td>
<td>0.884</td>
</tr>
<tr>
<td>Involvement of female participants</td>
<td>0.932</td>
</tr>
<tr>
<td>Involvement of participants from outside the main Yemeni sector</td>
<td>0.707</td>
</tr>
<tr>
<td>Setting up and implementing the Education Centres (ECs) as participants</td>
<td>0.932</td>
</tr>
<tr>
<td>The method of paying the fees of the master degree</td>
<td>0.707</td>
</tr>
<tr>
<td>Involving in the training process as part of the upgrading plan of the related organization</td>
<td>0.886</td>
</tr>
<tr>
<td>The competition policy of selecting the participants within the specific organizations</td>
<td>0.884</td>
</tr>
<tr>
<td>The criteria of selecting the participants by the related organizations</td>
<td>0.932</td>
</tr>
<tr>
<td>The criteria of selecting participants by the selected ECs.</td>
<td>0.707</td>
</tr>
<tr>
<td>The application of training knowledge and skills in the job-setting:</td>
<td>0.886</td>
</tr>
<tr>
<td>The relevance of the curriculum and course training for the Yemeni context</td>
<td>0.932</td>
</tr>
<tr>
<td>3-Transferring the training back to the workplace in order to promote the public sector reform process</td>
<td>0.707</td>
</tr>
<tr>
<td>Section Three</td>
<td>n.a</td>
</tr>
</tbody>
</table>
Results may differ among the selected projects in achieving the overall objective of the NPT program (promoting the capacity and capability of the respective sectors). This is based on the above-elaborated variables and the predications or proxies used in every variable. To explain the variation among the selected projects in achieving their overall objective, we developed other variables that measure the level of the effectiveness. The specific objectives as described in the table are related to the capacity building of higher education organizations. To develop the capacity of the organizations, as we mentioned before, the Dutch NPT projects aimed at strengthening the staff base and supporting the development of the curriculum and the educational infrastructure, developing course and training materials, training administrative and teaching staff, implementing scholarships to augment staff and equipping libraries and information centres with relevant materials. In parallel, organizational development was foreseen, giving the higher education organizations the tools to provide training services for the respective sectors.

Based on the development validity approach of evaluating aid projects by focusing on the results- in other words the success of development projects attributed to the inherent value of the project, (see for example Scott-Smith, 2013 and chapter 2)- we developed three constructs to measure the level of effectiveness among the selected projects in achieving their specific objectives. These variables, as outlined in Table 5.3 below are:

1. the quality of the organizational and administrative settings, 
2. the quantity of the staff members, and 
3. the quality of the staff members.
Table 5.3 The Constructs to Measure the Level of Effectiveness Among the Selected Projects in Achieving the Specific Objectives (Building the Capacity within the Higher Education Organizations)

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Measured items</th>
<th>Alpha (number of items in scales)</th>
<th>Place in the questionnaire</th>
</tr>
</thead>
</table>
| 1. The organizational and administrative settings | • Effective organization and management  
• Everything is easy to carry out your study  
• Effective operation procedures  
• Admin functions and duties are clear  
• The rules and regulations do not cause delay  
• The admin staff within the education center deal with all students in the same manner  
• The academic staff within the education center deal with all students in the same manner  
• The tuition fee is reasonable considering what the students learn from the master degree  
• The particular wishes of the students in terms of the time to follow the master programme are considered | 0.920 (9 items) | Section Two |
| 2. The quality of the staff members | • There are qualified teachers who are able to teach the characteristics particular to Yemen  
• Sufficient teachers who are able to teach the characteristics particular to Yemen  
• There are sufficient teachers to supervise the practical work or thesis of the graduate students | 0.878 (3 items) | Section Two |
| 3. The quantity of the staff members | • The attitude and working style of teachers are unique in Yemen  
• Teachers have almost the same level of understanding, skills and practical experience of education and research; work as a team in the EC  
• Teachers have the same awareness and agreement about common vision of the EC’s functions  
• Teachers are very committed and eager to work for the education center  
• Teachers have enough time to be available to the students  
• Teachers have access to the latest theoretical and practical developments in related subjects  
• Teachers are working to the highest standard  
• Teachers show that they have experience of working with the relevant sectors in Yemen  
• Giving skills and attitude-based training rather than information-based training | 0.700 (9 items) | Section Two |

The inter-item consistency and reliability: nothing in quantitative investigation is more important than reliability (Bernard, 1995: 38) due to the fact that if the reliability of measuring instruments is weak it could jeopardize drawing good conclusions. This study started measuring the level of the inter-item consistency reliability of the different factors including the study questionnaire. Reliability is the way to provide
consistent results (Gatewood and Field, 1990). The key measure and most commonly used statistic for multi item scales reliability is Cronbach's coefficient alpha (Bagozzi et al., 2003; Dean & Sharfman, 1993; Snyman & Drew, 2003, in Abdeldayem, 2003). This coefficient alpha can show the correlation average between all the entered items in order to structure the measure instrument. The values of the measurement range from zero to one. When analyzed the data indicated a higher correlation, the reliability venture is higher.

The nature and purpose of the scale are the main conditions that limit the different levels of reliability, as was shown in Tables 5.2 and 5.3 above. The most commonly used significance alpha value, which is the minimum level and recommended, is 0.70 (Nunnally, 1978). It is important to note, that the values of Cronbach's coefficient alpha are responsive to the number of entered items for analysis: the more items entered for analysis, the more reliable the test values will be. Thus, when there are a small number of items in the scale (fewer than ten) Cronbach's alpha values are relatively small. If alpha is less than 0.70, Pallant (2001) recommends removing items with a total correlation of less than 0.30. This is because an item with low values in the total correlations scheme indicates that the item is measuring something different from the scale (in Al Banna, 2010).

The creation of the questionnaire has taken into account many particular considerations. For example, it was designed to achieve the objectives by avoiding any ambiguous questions or statements. A fair variety of thoughts and judgments of the participants in the selected projects were targeted, because there is no former research in Yemen to build on (Nachmias and Nachmias, 1996). Five versions of the questionnaire were produced, first in English and then translated into Arabic, because most of the respondents in the selected projects only speak Arabic, apart from the MBA/CBA students who also speak English.

As suggested by Bagozzi (1994), there are three parts in the critical review process of the questionnaire. The researcher himself should perform the first critical review. A group of experts, who are knowledgeable about the theory and the general methods used in the study, should perform the second review. The third part is a pilot test of the questionnaire on a small number of respondents (as in al-Banan 2008: 178).

The questionnaire used for this study followed the same process. It was firstly critically reviewed. Secondly, some scholars from Leiden University and Sana'a University peer-reviewed the questionnaire and gave advice. For the Arabic version, a leading professor in statistics in Yemen reviewed it. A pilot meeting with a group of
participants was held, to make sure that the questionnaire was understandable. This was arranged through the organizing of semi-official workshops, which included participants from the different NPT-projects. The questionnaire had nearly 30 participants.

As a result of such validity reviews, many things changed in the questionnaire survey. For example, at the beginning, the Likert-scale questions had a three-response choice. However, some respondents became confused because both three-point Likert-scale questions did not give them the opportunity to answer without restraint. Therefore, in the second stage, I unified the number of scale items to a five-point scale throughout the questionnaire, which gave more options for the respondents to choose from and to answer freely.

It is worth noting here that selecting a sample is a fundamental aspect of the questionnaire design. Some issues which the researcher needs to decide about are the target population, the sample type, sample size and informants and the questionnaire administration. Those issues are addressed in Chapter 7 of this study. It is also relevant to mention that the target population for this study was limited to the participants (students) in the selected projects, who were administered into four groups:

1. Randomly selected public officials who were involved in the MPA & NIAS projects at Sana'a University
2. Randomly selected employees from the private sector who were enrolled at the MBA project at Sana'a University
3. Randomly selected students from the Women's Research and Training Centre (WRTC), Aden University.
4. Randomly selected students from Sana'a University, Water Center (WEC).

This population consists of participants with high-level career positions, mid-level career positions, and low-level career positions from different sectors such as the civil servant sector, the private sector, the water sector, and non-governmental organizations who are related to the WRTC project. This validates the several predictions that are used as proxies to measure the contribution of the NPT program in promoting the public sector in Yemen.

5.5 Method of Quantitative and Qualitative Data Analysis

The discussion now turns to the analysis methods of the collected data. Any scientific study demands appropriate analytical procedures to analyze the collected data. To analyze the level of effectiveness of the selected NPT projects in terms of
achieving their overall and specific objectives, we used both descriptive and inferential statistics for analysis. This means in this study we prefer not to use hard statistical measures to perceive value of effectiveness among the selected projects. It is wise to keep the statistics simple so they are understandable for the target groups in our investigation and for others. The descriptive analysis was used to “summarize, organize, and make sense of a set of scores or observations. It is typically presented in tables and histogram graphs, or as summary statistics (e.g., an average)…For example, the number of times each individual fidgeted is not all that meaningful, whereas the average (mean), middle (median), or most common (mode) number of times among all individuals is more meaningful. Tables and graphs serve a similar purpose to summarize large and small sets of data” (Carlson and Winquist, 2013). Descriptive analysis was used in this study to display the quantitative data and draw and verify conclusions among the selected projects (different groups). A one-way ANOVA test was used to interpret the results.

Frequency tables were used to show how often different values occur in the data. They can also be used to obtain summary statistics that describe the typical value and the spread of the observations. Tables display all the code numbers that occur for each variable and the number of cases in each of the code-number categories, including the value, the frequency of that value, and the percentage of the total responses for that variable. The valid percentage takes into account missing variables and the cumulative percentage (Mallah, 2003). In addition, we used the cross-tabulation as part of the analysis. The purpose of using the cross-tabulation procedure is to indicate whether there is variation among the selected projects in certain variables and measured items.

The histogram graph is a relative frequency distribution as it shows the manner of the data distribution. It is very useful to provide information to describe the population. An important point to note here is that different samples from the same population will result in different sample histograms, even when class boundaries remain fixed. However, it is expected that “sample and population histograms are similar, [and] the degree of resemblance will increase as more and more data are added to the sample. A histogram can be used in any situation in which it is reasonable to group adjacent values. It presents data in pictorial or graphical forms.” (Mendenhall and Beaver, 1991, cited in Mallah, 2003).

The mean and standard deviation were used to determine the measure of central tendency as well as the average distance of scores away from the mean. An ANOVA
was used to further analyze the means between and within the selected aid projects. An ANOVA is an inferential statistic used as a test of statistical significance, resulting in an F-value (Timothy Hatcher, 2011). It is helpful in making comparison of two or more means, enabling a researcher to draw various results and predictions about two or more sets of data. Thus a one-way ANOVA test was conducted to examine any differences in opinions among the participants in the different aid projects. These differences could be important because some aid projects may have more participants who are satisfied with the training programmes than others. The .05 level of significance was used to evaluate the association between aid projects and opinions of the participants about their actual impact.

The respondents were asked to indicate their agreement in the questionnaire on five-point Likert-type scales: 1= strongly agree, 2 = agree, 3 = neither agree nor disagree, 4= disagree, and, 5= strongly disagree. The number of the responses for each of the five points was then multiplied by its respective level of effectiveness, rating, 1 = strongly effective, 2= effective, 3 = neutral, 4 = ineffective, and, 5 = strongly ineffective. The result of the mean scores in the selected projects was averaged to construct a composite measure of the level of effectiveness of the aid project in developing managerial skills of its participants (doing cross case analysis). In other words, in calculating a mean score, a response of “strongly disagree” should be given a mean score of 5, and “strongly agree” should be given a mean score of “1”. The neutral point is “3.0”, and similarly, a response of “agree” should be given a mean score of 2 and “disagree” should be given a mean score of “4”. This composite measure reflects an aggregate view of each aid project based on the values of the total mean scores.

Based on the structure of the data, we have combined similar categories of response and removed the neutral category when presenting the frequency distributions as one analysis method. Thus, we would expect 50% of the responses in each of the two categories, assuming that those categories make up 100% of the responses. For some questions, this may give us a clearer picture of differences in the response. We have combined the categories of “Yes” as one group and” No” as another group for answers to some questions.

**Method of qualitative data analysis**: the current study adopted a small-N case study methodology, which “is often opaque, and must therefore be supplemented by another form of analysis that has come to be known as process tracing” (Gerring, 2007: 173). Alongside Gerring, other scholars like Tarrow (2010) and George and
Bennett (2005), assert that process tracing is an essential method for data analysis, in a study that selects case studies based on a similar case research design, which can be referred to as “controlled comparison.” The form of process tracing adopted in this dissertation can be described as an analytic narrative, engaging attempts to identify the intervening causal process – the causal chain and causal mechanism – between an independent variable and the outcome of the dependent variable (George and Bennett, 2005: 206). Thus, process tracing provides supporting evidence that helps to verify data obtained through other methods and sources. It is more convincing when multiple links are explicitly formalized and each micro-mechanism is proven. By generating and linking many observations to explain a case, the process tracing method contributes to hypothesis testing research. It is particularly useful for addressing the problem of equifinality by showing different outcomes for the same cause (Gerring, 2007:18; George and Bennett, 2005: 223-224).

The method of process tracing serves to examine the topic under investigation by collecting, clustering and analyzing data from a case at policy decision-making level and implementation level, from the perspectives of both donor and recipient. This examination consists of generating and linking observations to test the three main hypotheses of the study. I will trace the following steps in the form of a causal chain:

**Step 1:** The first step of process tracing is to examine policies and delivery mechanisms which comply with a donor’s thoughts, assumptions and stated objectives. At this stage, the donor’s policies and strategies were obtained from publications and documents. These were scrutinized in order to investigate the extent to which they incorporated local knowledge during the process. With regard to policy design, the focus was on the donor’s main mission statement and strategic objectives, its general policy and its specific policy towards Yemen, its regional approach, development targets and priority areas, country strategy papers, and white papers on development. With respect to the specific selected projects, this documentation was supplemented by material containing the particular project objectives and outlines.

**Step 2:** The second step of process tracing is to investigate, through interviews with policy makers and senior officials on the donor side, the written data provided on policies and the donor’s approaches in dealing with the project. This provided an opportunity to explain the disparity in the policies applied on the ground with the ones that are stated in documents.
**Step 3**: It focuses on the government’s own analysis of to what extent their missions, priorities, and national policies are reflected in the formal and informal institutions and organizations that promote public sector reforms. Some of these documents were not available however.

**Step 4**: Is semi-structured interviews with relevant government officials to get their views on their interactions with donor contractors. In addition, this step was also needed to investigate if there is a process that ensures a systematic incorporation of local knowledge, and reflects their objectives, policies and priorities that leads to applications for foreign funding.

**Step 5**: Compares the data obtained from the interviews with decision-makers with what is stated in the available official documents about development policies, project planning, design, structure and management. The data from documents and interviews is compared with the actual implementation in order to determine the extent of efficiency, consistency and ‘truthfulness’ of aid policies for both the donor and recipient, and to check if there are any changes during the implementation stage that incorporate local knowledge.

**Step 6**: This is a cross-case analysis addressing the objectives of the research in the light of the findings drawn from the integration of the case studies and the foreign aid literature. The different aspects of each objective were addressed through a cross-cutting analysis of evidence from interviews, documentation and mainly the findings of all case studies. By seeking confirmation from various data sources we get more reliable outcomes.

5.6 Validity Considerations

Validity considerations are still a major issue for researchers who are adopting mixed method approaches. Creswell and Clark (2007) noted that “the very act of combining qualitative and quantitative approaches raises additional potential validity issues.” (2007: 145) Some of the questions raised by Creswell and Clark are (1) how validity should be elaborated in mixed methods research. (2) How and when to report and discuss validity for qualitative and quantitative strands of mixed methods research. (3) Whether researchers should follow the “traditional validity guidelines and expectations; and (4) how to minimize potential threats to the validity related to data collection and analysis issues in mixed methods research?” (As cited in Venkatesh

---

4 - Regarding to these steps of data analysis, I partly adopted the ideas of Jaradat, 2008.
et al, 2013:14). To overcome these problems, Teddlie and Tashakkori (2003, 2009) proposed a solution by using the term “inference quality” to “refer to validity in the context of mixed methods research”. In contrast, Creswell and Clark argued that, because the term validity is “extensively used in quantitative and much qualitative research, it may be used in mixed methods research and thus making new terminology not essential” (as cited in Venkatesh et al, 2013, p15). This study is consistent with Creswell and Clark’s argument and uses validity in the way it is traditionally used in the qualitative and quantitative strands of research.

Accordingly, the validity considerations in this study should be related to the quantitative investigation and the qualitative investigation. Finlay (2009) argued that “qualitative researchers contest and reject the criteria used by quantitative researchers when evaluating their work: those of reliability, validity and generalizability. Instead, qualitative researchers have developed alternative criteria responsive to their specific research ideals. These criteria encompass various dimensions of ‘rigour’, ‘ethical integrity’ and ‘artistry’.” (2009: 1)

Before turning our discussion to these two views of validity considerations, it is worth defining the concept of validity. Campbell and Cronbach (2002) viewed the term validity as referring to the approximate truth of an inference or a property of inferences: “When we say something is valid, we make a judgment about the extent to which relevant evidence supports that inference as being true or correct. Usually, that evidence comes from both empirical findings and the consistency of these findings with other sources of knowledge, including past findings and theories. Assessing validity always entails fallible human judgments” (2002: 34).

The concept of validity is related to the process to which research truly measures what it is meant to measure. This criterion rests upon the assumption that the phenomenon being investigated possesses ‘reality’ in an undisputed, objective sense:

> Qualitative researchers in general view this as inappropriate. Given the diversity of the social world, they argue, it is erroneous to assume the existence of one unequivocal reality to which all findings must respond. They ask instead: whose reality is the research addressing? Moreover, qualitative research – by definition – involves subjective interpretations (often delivered by both participants and/or researchers). “If one accepts that interpretation cannot be excluded from the research process, it follows that any one analysis can only be presented as a tentative statement opening a limitless field of possible interpretations” (Churchill, 2000: 164 as cited in Finlay, 2009: 6).

Indeed, maximizing the validity consideration is a real problem that researchers face, as there are no straightforward tests or frameworks for quality measurement
(Patton, 2002). There is a wide range of evaluative criteria available to assure the quality and trustworthiness of the qualitative research. The test of validity is usually further divided into four types: internal validity, external validity, face validity, and construct validity (Dey, 1993; Lecompte & Goetz, 1982; Yin, 1994, 2003). These four kinds of validation provide a basic framework for conducting and evaluating the validation of the qualitative approach in this study. Thus we look at them in detail below.

5.6.1 Internal Validity

Internal validity is an approximate truth about inferences regarding causal relationships in a scientific inquiry. Therefore, internal validity is only relevant in studies that try to establish a causal relationship. It is for instance not relevant in most observational or descriptive studies. However, for studies that assess the effects of social programs or interventions, internal validity is perhaps the primary consideration to resort to. It is the degree to which a casual relation can be established, whereby specific conditions are proved to lead to other conditions (Yin, 1994: 35). The key question in internal validity is whether observed changes can be attributed to the program or intervention and not to other possible causes, sometimes described as “alternative explanations” for the outcome (Shadish et al. 2002).

Internal validity is related to a specific study problem or question. In other words, it cannot be generalized. According to Shadish “all that internal validity means is that you have evidence that what you did in the study (i.e., the program) caused what you observed (i.e. the outcome) to happen. It does not tell you whether what you did for the program was what you wanted to do or whether what you observed was what you wanted to observe -those are construct validity concerns. It is possible to have internal validity in a study and not to have construct validity” (Shadish et al. 2002:34)

Voeten (2013) reviewed two definitions of the concept internal validity. The first definition was cited from Sumner and Tribe (2008) who viewed internal validity as critical to the credibility of the research design and its findings. They argue that internal validity concerns the explicit correspondence, coherence and consistency between the data, interpretation and the writing-up. The second definition of the concept is taken from Benz and Newman (1998) who proposed several strategies for maintaining internal validity, including evaluating the reliability and quality of the data and the internal line of reasoning towards conclusions (2012: 17).

It is clear that our study intends to provide an insight in the ‘why’ question; it is a
causal investigation. Different essential aspects of internal validation have been realized in this research. The first validity consideration is that our study adopts case analysis methods, which leads to improvement of internal validity (Gerring, 2007). However, a specialist reader may note that because a small-N case design was adopted internal validity is at stake. This is partly overcome by adopting process tracing for data collection and analyzing stages (see previous section). Process tracing allowed this investigation to have multiple types of evidence for the verification of a single inference (incorporating local knowledge) and pieces of evidence that embody the unit of analysis on different levels (policy-making and implementation levels). The information from observations is gathered by unique populations from different levels and institutional landscapes (Gerring, 2007). As I am studying in The Netherlands, part of the population is from The Netherlands. By doing this, I gained further experience about the social, organizational and cultural aspects of Dutch society. My Yemeni background/heritage affords me insight into organizational and cultural aspects of Yemen, which is helpful for understanding the second Yemeni study population group in this study. The different backgrounds of the involved actors in the networks prevented those actors from having effective interactions, from transferring ideas and knowledge that could be useful in the processes of designing and implementing aid projects. I found that many of these helpful practicalities were reflected in the literature on logistics and research practice (Murray and Overton, 2003).

After presenting here most of the internal validity issues and their solutions, it is useful to apply Mofakkarul's (2007) overview of the different studies that describe the threats to internal validity which qualitative researchers face, and advocate possible remedies to circumvent the threats (Eisenhardt, 1989, 1999; Lecompte & Goetz, 1982; Lincoln & Guba, 1985; Miles & Huberman, 1994; Yin, 1989, 1994, 2003). Mofakkarul summarized these threats and applied them in his research. I have applied them in my study, as shown in Table 5.4 below:
Table 5.4 Important Threats to Internal Validity of Qualitative Research, Tactics Advocated in the Literature to Minimize Threats, and Tactics Applied in this Study

<table>
<thead>
<tr>
<th>Threats</th>
<th>Tactics to minimize threats</th>
<th>Applied in this research?</th>
</tr>
</thead>
<tbody>
<tr>
<td>History and maturation effects</td>
<td>• Long term residence in the field</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Retrospective tracing of phenomena</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Time-sampling procedure</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>• Revisit sites at subsequent intervals to verify time-dependent phenomena</td>
<td>N/A</td>
</tr>
<tr>
<td>Threats relating to observer effects or misinterpretation</td>
<td>• Triangulate:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Data source</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Method</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Theory</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Investigator</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Gradually disengage from informal relationships, periodically withdraw from field (i.e. don’t ‘go native’)</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Extended fieldwork to understand people and setting</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Code participant responses according to situations expected to elicit contrived responses</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Be reflective, that is, engage in self-critical reflection, monitor own bias and try to control</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Have case summaries checked by participants</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Use verbatim when writing research reports to let readers understand subjects’ viewpoints and emotions</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Perform pattern matching</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Select and study negative cases</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Peer review (e.g. discussion with colleagues)</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Be specific and look for alternative explanations; enlist as many competing explanations as possible</td>
<td>YES</td>
</tr>
<tr>
<td>Mortality</td>
<td>• Identify and record loss or gain in group membership; collect baseline data to compare activities and events over time</td>
<td>N/A</td>
</tr>
<tr>
<td>Reactive effects</td>
<td>• Compare with already documented cases</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Record reactive effects on respondents</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>• Interview subjects on reactive effects</td>
<td>YES</td>
</tr>
</tbody>
</table>

Source: entirely adopted from Mofakkarul (2007: 106) and applied in my study.

5.6.2 External Validity

This study conceives external validity as a “determination of whether the study was credible, accurate, and useful to others outside the participant group (Roberts et al., 2008). Making explicit statements and acknowledging subjective judgments ensured a study’s validity (Ashworth, 2008)” (as cited in Odemba, 2010: 81). This definition is also confirmed by Cooper & Schindler, (2008), Leedy and Ormrod (2005), and Bacal (2008) who regarded external validity as the transferability process of the study’s conclusions and findings: the capability to deal with the same problems in other situations. Ray (2003) and Churchill et al. (1998), as cited in Voeten (2012), argued that, in qualitative methodologies, “it is important that data collection and the line of reasoning that leads towards the conclusions can be validated by external
reviewers. Transparency in the line of reasoning is essential. Churchill et al. (1998) suggests that this can be judged by the coherence criterion; the coherence between the raw data and the identified patterns and conclusions. The cases should be described in a common format that facilitates identification of the patterns described” (2012: 20).

It is worth noting that according to Leedy & Ormrod (2005), generalizations of a qualitative study can be limited. The uniqueness of the study might hinder a direct replication in other contexts.

Denzin (1982), Guba & Lincoln (1985), Lecompte & Goetz (1982) and Yin, (1994, 2003) have presented the key threats to external validity and the tactics used by qualitative researchers to minimize those threats (cited in Mofakkarul, 2007), which we present in Table 5.5 below. This study tried to deal with such threats during different stages of research, especially as I am trying to construct a theoretical generalization that is related to the LKS.

**Table 5.5 Important Threats to External Validity, Tactics Advocated in the Literature to Minimize those Threats, and Tactics Used in this Study**

<table>
<thead>
<tr>
<th>Threats</th>
<th>Tactics to Minimize Threats</th>
<th>Applied in this Research?</th>
</tr>
</thead>
<tbody>
<tr>
<td>General threats:</td>
<td>Provide a detailed description of the context, characteristics of the people and groups and other necessary information needed to demonstrate typicality</td>
<td>YES</td>
</tr>
<tr>
<td>Comparability and translatability</td>
<td>Use theoretical sampling</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>In a multiple-case design</td>
<td>YES*</td>
</tr>
<tr>
<td></td>
<td>In a single-case design, use rival theories</td>
<td>YES</td>
</tr>
<tr>
<td>Selection effects</td>
<td>Collect evidence from different perspectives (e.g. participants versus non-participants)</td>
<td>YES</td>
</tr>
<tr>
<td>Setting effects</td>
<td>Beware of over-saturation of settings and avoid subjects that have been under repeated investigation</td>
<td>YES</td>
</tr>
<tr>
<td>Historical effects</td>
<td>Describe influence of historical factors on target groups</td>
<td>YES</td>
</tr>
<tr>
<td>Construct effects</td>
<td>Explicitly define constructs</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Avoid use of idiosyncratic constructs that are difficult to compare across groups</td>
<td>YES</td>
</tr>
</tbody>
</table>

Source: Mofakkarul (2007). *I used a multiple-case design within a country. All my small N or cases form the same country setting, which can represent a threat for the replication logic. However using rival theories is a solution to this threat.

The study addresses general and external validity threats by providing a detailed description of the context of the case study, people, and organizations. Furthermore, a description of the main causal mechanisms of the local knowledge syndrome at the
policy design and policy implementation level is provided to serve other researchers, who can compare their own work with the findings of this study (Chapter 3).

When we see the case-study design in a global perspective, this study has a single case nature. On a country level, our case design is a multiple-case design that increased the external validity within the Yemen context, because comparing four cases or organizations from different areas in Yemen ensures external validity. Sana'a University is located in the north of Yemen and Aden University is located in the south. The design included two cases that succeeded and two cases that failed. Three of the cases belong to the same organization (Sana'a University). That means the research’s findings can be generalized at the Yemen level.

Generalizability also refers to possible implications of this study’s findings to other settings outside of Yemen. Despite the single-case design, this study can be generalized. To overcome the generalizability or external validity threat, this study builds on multiple theories such as the explanations of LKS from the viewpoint of rational choice institutionalism, political anthropology, economic and development theories, network perspective, including the actor network approach, policy design and implementation theories, and those originating in the field of aid project management (Chapters 1, 2 and 3). To review the conceptualization of LKS, this study analyzed the problems that came up in different studies related to foreign aid ineffectiveness in recipient countries that have more or less similar economic, political, social, and administrative structures as Yemen.

This study also addressed other specific threats to external validity such as the effects of selection, setting effects, historical effects and construct effects. To minimize selection effects, data was collected both from people who were directly involved with the policy design and implementation, and those who were not involved in the process, such as the external evaluators of the NPT, some experts and senior officials from the Yemeni government and university scholars. The setting effect was clearly evident in some of the projects investigated in this study. Some of the projects were not doing well and the teams were afraid of giving the right information especially during the interviews. This threat was minimized by asking the recipient about the role of the donor contractors and by asking the donor contactors about the role of the counterparts in the recipient side. In addition, the findings of a number of subjects in the study were crosschecked.
5.6.3 Face Validity and Construct Validity

According to Dey (1993), “face validity refers to the degree to which there is a fit between the observations of a researcher and the concepts used by the person. Construct validity, on the other hand, is the degree of fit between the definitions of the concepts used by a researcher and the “previously established” and “authoritative” definitions of the concepts in the literature” (1993: 255). Face validity and construct validity thus “refer to the adequacy of the definition of a concept and how a concept is embedded, or can be traced back, in a network of theoretical associations” (Dowling, 1986, as cited in Voeten, 2012: 17). However, “fuzziness is a frequently recurring issue in the social sciences. Once a definition or construct has been proposed, it must be validated as a theoretical concept or as a fundamental unit of thought with a meaningful role within existing theoretical systems.” (Voeten, 2012: 17) To ensure face and construct validity of qualitative research in this study, we adopted the techniques presented in Table 5.6 below, based on Dey, 1993; Yin, 1989, 1994, 2003:

<table>
<thead>
<tr>
<th>Techniques in the literature</th>
<th>Applied in this study?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use multiple sources of evidence</td>
<td>YES: This study adopted a mixed study method, qualitative and quantitative, with multiple sources of evidence: interviews, questionnaire, observations, external evaluation reports of the projects based on empirical work and related project documents. I also used secondary resources such as books, articles, and internal reports related to the subject.</td>
</tr>
<tr>
<td>Use concepts from the literature</td>
<td>YES: These were explained in Chapter 1, and at the beginning of every chapter, I have explained the related concepts. The LKS is a new concept in this study that I have linked to many theories and different explanations. The operationalization of the LKS is presented in Chapter 3.</td>
</tr>
<tr>
<td>Maintain a chain of evidence</td>
<td>YES: I have done this by using the tracing process of collecting data and analysis. It was explained in this chapter.</td>
</tr>
<tr>
<td>Demonstrate the criteria by which the data was categorized and exhibit how the connections between the concepts were identified and made explicit</td>
<td>YES: discussed in this chapter</td>
</tr>
<tr>
<td>Provide both typical as well as negative or extreme instances</td>
<td>YES: the design of this study adopted two negative examples and two positive examples of cases</td>
</tr>
<tr>
<td>Have key informants to review draft case study reports</td>
<td>NO: but interview summaries were discussed with some of the key informants of NUFFIC and the ECs involved with the study as well as with my supervisors</td>
</tr>
</tbody>
</table>

Source: Mofakkarul (2007) according to Dey, (1993); Yin, (1989), (1994), and (2003), and reapplied in this study.
The main concern in this study is conceptual fuzziness, because many of the terms used in this study are broadly applied and have multiple definitions, such as institutional development, capacity and capability building, public sector reforms, administrative reforms and poverty reduction. This problem is minimized by linking the theoretical elements to the empirical side of the aid project, namely policy design and implementation. The analytical process is provided in Chapter 10. However, conceptual fuzziness is a problem within any social research and it is difficult to solve completely.

I used multiple sources of data and established a chain of evidence through the comparison of data from various interviews and documents. I crosschecked information and conclusions to ensure constructive validity (Leonard-Barton, 1990 and Yin, 1994) and also internal validity (Flick 1992). According to Yin (1994), “the use of multiple sources of evidence in a case study makes the conclusions more convincing and precise. It is a question of developing convergent lines of research. The triangulation also makes it possible to improve constructive validity, because the different sources contribute to additional measurements of the phenomenon” (as cited in Leonard-Barton, 1990: 12). In this study, a triangulation technique was used to compare data from different sources including observations, interviews, secondary resources, survey-questionnaire and official and semi-official reports published by related actors from both the Dutch side as well as the Yemeni side. We have to check if the interview guide is fully reflective of the information that was needed from the field. In the next step, the findings were linked to the theoretical perspective of this study. I formulated interview guides for the conduction of interviews with participants.

5.6.4 Objectivity of the Researcher

Before turning to a discussion on the problems I encountered during the investigation, it is important to briefly address the objectivity of the researcher. Objectivity is critical for any type of research. However, no research is value free. Every researcher has some personal motives, issues or bias toward the subject being studied. In qualitative research there is a term for this: conformability. This describes the extent to which the researcher has controlled the intrusion of his/her personal values (Voeten, 2012: 17). When I was part of the MPA project, it was logically to be biased towards the outcomes of the project, but one year after the closure of the project and having discussions with former long-term manager of the NPT program in Yemen, I felt I had enough distance to take an objective point of view on the
research idea.

I hope that as a researcher from the global south conducting this research I show how we can realize an institutional change in the poor nations by solving the local knowledge problem. I have put my focus on input from the government of the developing country that is receiving the foreign assistance, and not merely on development theories that are presented by researchers from the developed world or the donor countries. Most of these western theories of development are used by donor agencies to inform their development approaches, while these theories are actually downplaying the local knowledge and the local perceptions in the recipient countries.

5.7 Encountered Problems

There were several problems related to the interviews. Due to the sensitivity of the subject, and as a result of the political and personal interests of some participants, it was difficult to record a few interviews without affecting the participant's interest. Many of the participants who function in high government positions at both donor and recipient sides asked me not to quote them directly. The Yemen specific context also makes this increasingly difficult. I held interviews with different academic staff of several departments who belonged to the same institution. Conflicts of interests seemed to be always present, because of the Dutch project. Fear of relative deprivation seemed pervasive. The researcher’s position as a colleague also made it sensitive for them to expose their ideas freely. However, this problem was overcome by asking respondents for information about the role of the different levels. For example, I would ask the policymaking level about the role of the middle management and the operational level officials. I asked the middle management about the role of the policymaking level and operational level. I asked the operational level about the role of the middle management and policymaking levels. At the end, I asked all the levels about the role of the donor contractors and policymakers in the process. The problem was treated carefully by taking written notes during interviews in some interviews (but the other interviews I used an audio recorder), thus protecting anonymity and confidentiality, as well as making the respondents more forward in their answers.

Another difficult problem was making appointments with contractors, counterparts and policymakers on both donor and recipient sides. On the donor side, contractors are generally busy and some of them reside in different recipient countries while working on other projects. Often, government officials and policy makers were
difficult to contact (certainly for interviews) due to scheduling and travel restrictions. In many cases, the researcher had to go through intensive government procedures of writing letters and overcoming bureaucratic obstacles to get to the officials. The Yemen specific context proved an additional challenge during the field study, as Yemen is (still) suffering from political, social and security unrest, which prevented the researcher from reaching potential respondents. Despite the difficulties, I managed to meet almost 85% of the participants, mostly in informal or personal settings (e.g., during the Qat\textsuperscript{5} sessions or in their homes). In one year of field study, this research engaged 65 respondents of interviews, from both donor and recipient sides and presentations at 2 conferences in Yemen.

Another problem was related to the availability of the official documents produced by the involved actors. In most cases, there was little if any documentation on projects provided by the recipients. Donors usually generate more of the documented material, as they are obliged to report to the stakeholders who invest in their aid programmes. Yet Nuffic and MinBuza made it very difficult to get such documents, indeed it was necessary for my doctoral supervisor to get involved to acquire the solicited documents. This problem was also partly circumvented by informal ways of getting the documents through colleagues in Yemen, and by placing more weight on the interviews with respondents on the recipient side, which proved to be a more productive source of data than documentation.

\textbf{5.8 Summary and Conclusion}

In this chapter I describe the research strategy that investigated the reasons why aid projects aimed at promoting public sector reform in Yemen succeeded or failed. This question demands a theoretical framework that explains how and if unsuccessful aid programs occur as a result of a failure to incorporate local knowledge in policy design and/or implementation. The theoretical perspective of the LKS is a new approach that challenges the PCP approach.

In this research, the case of Yemeni-Dutch cooperation was studied to challenge this main argument of the PCP approach that aid interventions in poor nations are ineffective. The PCP approach claims that that the existence of multiple egoistic goals of aid assistance may be the main reason. The Netherlands also has strategic interests to give aid to Yemen: for security and stability. In the next chapter I assess if these

\textsuperscript{5} - Yemenis spend almost six hours a day chewing Qat in special places called Diwans. Qat is a white-flowered evergreen shrub, Catha edulis, of Africa and Arabia, whose leaves have narcotic properties. It is traditional for most people in Yemen to chew the leaf.
development efforts led to improvements in terms of development outcomes that have meaningful impact on security and stability.

A comparative case study approach with small-N design was used to develop an in-depth understanding. The selection of cases was guided by what is known as a “Most Similar Case Selection Strategy” that exemplifies and vindicates the case study methodology used in this study. According to the basic logic of this method, the cases under study possess generally similar circumstances, but have nevertheless experienced significantly different outcomes. This logic of research design allows us to select cases based on the dependent variable (success or failure of aid programmes) and then working backward to see if the independent variable (incorporating local knowledge) has the right or predicted value of the dependent variable.

This study adopted a mixed methods research that combines qualitative and quantitative methods and draws on the strengths of both traditions of inquiry. As I showed, the quantitative analysis allows the researcher to meet the first research objective and provides an important link making conclusions about the effectiveness of the selected projects and the question of why cases differ in their (in)effectiveness of completing their mission is dealt with qualitatively. Qualitative analysis is based upon the LKS explanation as the independent variable in this study. A qualitative and theory-building or theory elaboration approach was also adopted for the research because there has been no prior theory developed on the explanation of the LKS for aid projects’ ineffectiveness. The adoption of the mixed methods research allowed us to use multiple data collection methods: interviews, focus group discussions, document analysis, and participant observation. The questionnaire was the main method of quantitative data collection.

Following the logic of enquiry in this methodology chapter, the next chapter presents the macro analysis of Dutch development efforts in Yemen, in order to show that the multiple goals of donors are not the main cause of ineffective aid projects, but that applying local knowledge is important to reach the set goals. Then in Chapters 7 and 8 I conduct a microanalysis of selected Dutch NPT projects to conclude which projects were most effective. The answer to the question of why certain projects were more effective than others were, the relation to the LKS and the response to the PCP approach is presented in Chapters 9 and 10.