

The theory and practice of analysing institutional development in evaluation

Final report to the ECG

Strängnäs 30/3 2001

**Kim Forss
Andante – tools for thinking AB**

Executive Summary

The present study is part of the joint methodological development of the multilateral development banks. Work to improve the accountability to owners and feedback of lessons from performance evaluation are some of the concerns of the Evaluation Cooperation Group (ECG), which consists of the evaluation units of the multilateral development banks. The ECG designed the present study with the following specific purposes in mind:

1. To present an overview of how the institutional transition and development dimension has been reflected in recent studies and evaluations
2. To analyse how the banks currently apply institutional development criteria at performance evaluation of completed operations and in sector- and country oriented impact studies.
3. To present recommendations for improved methodology and reflection of the institutional development impact.

The study builds on a review of policies and procedures in the six different banks, plus an analysis of a sample of 31 evaluation reports. These reports were chosen by the ECG members because they were seen as interesting examples of institutional development analysis, hence they represent some form of best practice. They are not typical of the banks' evaluation work.

Institutional development has come to occupy a prominent position in development economics. It can be argued that it is the hard core of analysis of what constitutes development and how development comes about. The field is evolving rapidly, and there is a growing body of literature on aspects of institutional development, as well as tried and tested research methods.

Institutional development refers to activities at various levels of society. But there are not always clear cut borders between neither levels nor activities - and the broader concepts encircle the more narrow. It is useful to distinguish:

- (1) Human resource development (training and education) which are concerned with how people are educated and trained, how knowledge and skills are transferred to individuals, competence built up and people prepared for their current or future careers.
- (2) Organisational development which seeks to change and strengthen management systems in specific organisations in order to improve performance. There are variations between theories and strategies, but they have in "pure" form the following characteristics: (a) focus on individual formal organisations and particularly their internal functioning, (b) less attention paid to external contextual influences on performance, (c) concern with internal organisational changes, and (d) activities include education, training, advice, design of structures and processes.
- (3) Systems development is a broader concept than organisational development. In addition to a concern with human resources and the development of particular organisations, it includes an emphasis on linkages between organisations and the context or environment within which

organisations operate and interact. It is useful to make a distinction between three system levels:

(i) *The network and linkages* among organisations, which include the network and contact between organisations that facilitate or constrain the achievements of particular tasks.

(ii) *The regulatory environment*, which refers to the policy environment of the public, private and civil sector that constrains or facilitates organisational activities and affects their performance, including e.g. laws, regulations and policies.

(iii) *The value framework* - including the economic, social, cultural and political milieu in which organisations operate, and the extent to which conditions in this broader environment facilitate or constrain the functional capacity of organisations.

This description of institutional development underlines the interaction between micro (internal) and macro (external) factors determining how organisations translate their capacities into actual performance. Generally speaking, institutional development thus refers to change processes that improve the capacity of a social system to achieve its goals and objectives.

With this definition in mind, the study turns to an analysis of the evaluation reports. The first question is what they actually analyse when they analyse institutional development. It was found that the most common aspect is organisational development; that is, most evaluations limit their analysis of institutional development to the impact on one particular organisation, which is often the object of intervention. They analyse organisational networks, but not that frequently. They seldom address issues of the regulative environment, and very seldom issues of values and attitudes. The evaluations thus focus mostly on one aspect of institutional development, while leaving many others unattended.

Evaluations can be said to take either a summative or a formative approach to their subject. The analysis of institutional development follows both approaches. However, those evaluations that take a formative approach convey more useful information. They present data that is reliable and valid. Those that take a summative approach rarely convey any data to support their conclusions, and their overall utility must be questioned.

The choice of methods is restricted, interviews are by far the most common approach to data collection. The different levels of institutional assessment require different methods. Surveys, structured observation, working with documentation, are examples of methods that need to supplement interviews. When evaluators address more complex systems, they need to consider sampling issues. They need to structure their analysis by formulating hypotheses to guide data collection and analysis. Few of these evaluations explicitly work with hypotheses, but those who do so often present more coherent analysis and better structured reports.

A third of the evaluations use indicators when they assess institutional development, however, sometimes the word indicators is used more in the sense of a topic to be researched, or a broad field of inquiry. Ranking scales differ, but there is a need to consider their design (how extensive the scale is, where it begins and ends, whether it should have a middle point), as well as the choice of words that define impact.

The recommendations of this study those focus on:

(a) Conceptual clarity; namely to agree on a joint definition. It is necessary to steer the work in this field with conceptual clarity. It is thus suggested that the participants in ECG agree on the term and issue a joint statement of how institutional development is to be defined in the context of evaluation work.

(b) Overall strategy; that institutional development is primarily treated formatively. Whatever evaluations say about institutional development must be meaningful for project managers, partners in developing countries/transition economies, and desk officers. If the information cannot be put to use by practically oriented people, then it may not be worthwhile collecting and disseminating.

(c) Developing indicators; to ensure that evaluation teams possess the skills to work with indicators, to sensitise them so they do not forget to develop indicators, and to guide them by referring to manuals, best practice examples, etc.

(d) Measurement and ranking; in that the same ranking scales were used by the banks. As there is a separate consulting assignment going on, in parallel to this, it is likely that a joint ranking system will emerge. Needless to say, the comments in chapter 4 in respect of the choice of words, the basis for ranks, and inclusion of possibly negative outcomes, are put forward as an input to this work.

(e) Methods to assess institutional development; there is a difference between evaluation theory, which is expected to be a systematic application of social science methods in an inquiry, and the actual practice of several evaluations. Due to time constraints, they often make do with a limited application of methods (read unstructured interviews), and with meagre sources of information (read project managers and other highly biased sources of information). The topic of institutional development is too difficult to handle without scientific rigour, or without systematic approach.

Practically speaking, it may be useful to distinguish between practices of evaluation and of peer review, where the latter is a less pretentious way of arriving at conclusions. Many times, it describes more accurately what actually happens. Sometimes, the etiquette of "evaluation" is merely a false label. It is better to give the study a name that describes what kind of a product it really is.

But it is also possible to really undertake evaluations, in which case there is a need to plan the inquiries well in advance, to search for locally qualified teams, to emphasise methodological competence. It would also be prudent to focus the analysis of institutional development. It is too wide a topic to be comprehensively covered in most evaluations. It is better to cover one aspect well, than many less accurately.

(f) Finally, these recommendations illustrate a general request for increased coverage, variety and sophistication in reporting on institutional development.

Table of contents

EXECUTIVE SUMMARY	2
CHAPTER 1. INTRODUCTION	7
Background	7
Purpose	7
Methodology	8
A guide to the reader	9
CHAPTER 2. TOWARDS A POLICY OF INSTITUTIONAL DEVELOPMENT	10
Institutions and the explanation of development	10
Towards a definition of institutional development	12
Contested issues in institutional development	16
Institutional development and capacity development	16
Organisational development is a more limited aspect of institutional development	16
Institutional cooperation is one of the means towards institutional development	17
Institutions and markets	17
Institutional development; the private sector and public administration	18
Mission and policies in respect of institutional development evaluation	19
CHAPTER 3. COVERAGE OF INSTITUTIONAL DEVELOPMENT	22
The focus in institutional development analysis	22
The blank fields in coverage of institutional development	24
CHAPTER 4. MEASURING INSTITUTIONAL DEVELOPMENT	26
A strategic distinction in how evaluations cope with institutional development	26
Using indicators to assess institutional development	30
The use of scales and the measurement of impact.	33
CHAPTER 5. THE METHODOLOGY OF EVALUATING INSTITUTIONAL DEVELOPMENT	36
Evaluation as a systematic inquiry.	36
Models and hypotheses	37

The choice of methods.	38
Developing the methods to evaluate institutional development	40
CHAPTER 6. EXAMPLES OF GOOD PRACTICE.	43
CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS.	50
Conceptual clarity	50
Overall strategy	51
Developing indicators	51
Measurement and ranking	52
Methods to assess institutional development	52
Increased coverage, variety and sophistication in reporting on institutional development.	53
REFERENCES	55
ANNEX 1. TERMS OF REFERENCE	59
ANNEX 2. LIST OF EVALUATIONS USED IN THE ANALYSIS	63
ANNEX 3. ANALYTICAL TABLES USED IN THE REPORT	65

Chapter 1. Introduction

Background

Several events over the past decades have led economists to consider institutional development. The boom and subsequent depression in East Asia, and the collapse of the planning economies in Eastern Europe, are only two well-known examples of institutional change, and how significant these are for the economies where they occur. In fact, a theory of development, which does not adequately cover the nature, and quality of institutional change would today appear meaningless. Institutional development is a core feature of development, and hence of crucial significance for all actors in development cooperation.

But institutional development is a complex subject. Even though significant research has been devoted to institutional development over the past decade, the state of the art is anything but coherent. Consequently many of the actors most concerned with global development have had problems bridging the gap between theory and practice. One part of the problem is manifest in the practice of evaluation; the methods to explore institutional development within the framework of conventional evaluation is at best uncertain, and the practice ranges from the excellent to the miserable. There is a need to take stock of the state of the art, and to ascertain that evaluation systems provide information so that actors can learn from experience and be held accountable for what they do in terms of institutional development.

Purpose

Irrespective of whether and how it is understood, the activities of multilateral funding agencies are bound to have an institutional development impact - and there is a need to forecast that impact and to verify whether it happened or not. The present study is part of the joint methodological development of the multilateral development banks. Joint work to improve the accountability to owners and feedback of lessons from performance evaluation are some of the concerns of the Evaluation Cooperation Group (ECG), which consists of the evaluation units of the multilateral development banks: the European Bank for Reconstruction and Development (EBRD), the World Bank, the International Finance Corporation (IFC), the European Investment Bank (EIB) and the three regional development banks; the Inter American development Bank (IADB), the African Development Bank (AfDB) and the Asian Development Bank (AsDB). The ECG has designed the present study with four specific purposes in mind:

1. To present an overview of how the institutional transition and development dimension has been reflected in recent studies and evaluation guidelines by member multilateral development banks (MDB's), taking into account their mandates, strategies and operational policies.

2.To analyse how each MDB currently applies specific institutional development criteria at performance evaluation of completed operations and in sector- and country orientated impact studies.

3. To present a comparative analysis and to reach conclusions.

4. To present recommendations for improved methodology and reflection of the institutional dimension in evaluation criteria and practice by ECG-members.

Hence, this study is mainly descriptive and exploratory. It will be necessary to address issues of definition, to assess whether institutional development is understood (needs to be understood) similarly by the cooperating partners from the banks. Evaluation guidelines, and criteria for assessment, would be confused, if they do not rest on a sound, clear and precise definition of the object they are applied to. The study is a first step, and will in due course be followed by a second phase, which is meant to contain recommendations concerning best practice and processes in evaluation. The terms of reference for this study are enclosed in annex 1.

Methodology

The study builds on a review of policies and procedures in the six different banks, plus an analysis of evaluation reports. The review of policies and procedures builds on manuals of operation, policy statements, and information obtained through interviews and mail correspondence with the members of the Evaluation Cooperation Group. The information thus obtained describes the intentions of the banks. But intentions are one thing, actual practice another.

The actual practice was analysed through reviews of a select number of evaluation studies. A total sample of 31 reports was selected. These represent only a small share of total evaluations, and they were selected because they represent particularly interesting approaches to the evaluation of institutional development. Annex 2 contains a list of the evaluations that were reviewed.

The analysis of evaluations was undertaken with the help of a model which distinguishes between the contents in terms of entry points to institutional development, methods in use to assess impact, as well as conclusions in respect of that impact. In other words, the model describes what the evaluations analyse, how they undertake the analysis, and what results they arrive at. The study refers to the definitions of quality advanced by the American Evaluation Association, and as published by Joint Committee on Evaluation Standards (1994).

This report does not contain any analysis of typical patterns of institutional development analysis, and hence it does not reflect the actual practice of the banks. Instead, it is a step in the development of improved methods, and more streamlined procedures that, when further developed, should increase the accountability of the banks, and facilitate sharing of lessons learned, and making use of common experiences.

Towards the end of the study, it was found that the 31 evaluations analysed missed out in respect of several dimensions of institutional development, much because many were project level evaluations. In fact, institutional development is likely to be analysed in more aggregate

evaluation reports; in country studies, thematic studies, etc. Hence, six more evaluations were incorporated, not to be analysed comprehensively, but merely to pick out examples of good practice. These additional evaluation reports are listed separately in annex 2.

A guide to the reader

It would be vain to start this review without some attention to the definitional issues. The next chapter (chapter 2) thus outlines why and how the research community in development economics came to be preoccupied with institutional issues. This leads to a definition of a framework for the analysis of institutional development, and discusses such issues as the distinction between institutions in transition and institutional development, private sector versus public administration, and the market as an institution - and its institutions. This chapter also reviews the thinking of the multilateral development banks in respect of institutional development, and describes how they have operationalised these issues in evaluation policies and guidelines. The chapter thus discusses the framework for evaluation, and it compares the differences as well as similarities between the different banks.

Chapter 3 analyses what kind of institutional development is assessed in the evaluations, and describes the common patterns between the banks. The chapter uses the 31 evaluations to exemplify typical assessments, point at particularly clever approaches to the task, and also points at the white spots on the map, that is, aspects of institutional development that are seldom addressed.

Chapter 4 looks at what the evaluations actually say about institutional development. The chapter is thus concerned with indicators, measurement, and scales. Evaluation is eventually an inquiry into the worth or merit about something, in this case institutional development. The indicators and measurements, and the philosophy behind them, are critical to understand in order to develop the accountability systems regarding institutional development.

Chapter 5 finally looks at the quality of analysis, that is, are the value statements credible? Can they be trusted, are they good enough to lead onwards to well informed lending decisions. Which methods need to be applied, and which resources are needed to make informed judgements about institutional development?

Chapter 6 contains a discussion of good practice in relation to the use of indicators, structuring an inquiry with the help of hypotheses, and selecting appropriate methods for different levels of intervention. The paper concludes with recommendations for future to develop the evaluation of institutional impact (chapter 7).

The reader should remember that institutional development can be analysed at the level of a project, at the programme level, in respect of sector evaluations, thematic evaluations and country evaluations. The focus in this report lies on the project level, particularly in the good practice section, which only provides examples from project level evaluations.

Chapter 2. Towards a policy of institutional development

Institutions and the explanation of development

When Douglass North was awarded the Nobel prize for his contributions to economic history, a wave of interest in institutional development followed. This may seem strange, as the work he was awarded the prize for had appeared many years earlier, and many other researchers had for long been working on institutional development. However, even though fashions come and go, it is probably correct to say that nobody, either in academia or as practising development economists, can ignore institutional development any longer (Adelman and Taft Morris, 1997).

But what does it really mean? In their seminal work on institutional development, North and Thomas (1973) focus on *“those institutional arrangements which enable units to realise economies of scale (joint stock companies, corporations), to encourage innovation (prizes, patent laws), to improve the efficiency of factor markets (enclosures, bills of exchange, the abolition of serfdom), or to reduce market imperfections (insurance companies). Such institutional arrangements have served to increase efficiency. Some could be created without changing existing property rights, others involved the creation of new property rights; some were accomplished by government, others by voluntary organisation.”*

North and Thomas were above all interested in explaining growth, and hence they tended to focus on aspects of institutional development that make the markets for economic transactions effective, and in particular their concerns focus on the role of property rights. At times they speak of markets as one of the fundamental institutions of society; hence institutional development also implies the development of markets, as well as of governmental functions. In North's later work, institutional analysis, very broadly defined, moves to centre stage. *“In our daily interaction with others, whether within the family, in external social relations, or in business activities, the governing structure is overwhelmingly defined by codes of conduct, norms of behaviour, and conventions”* (1990, p.36). He then argues that these *“institutional frameworks”*, are *“the critical key to relative success of economies”*.

Aid agencies in general, and perhaps the international agencies in particular, became more manifestly interested in institutional development in the early 1990s. It was considered surprising, and path-breaking, when the Managing Director of the World Bank, at the Bank's annual meeting in 1991, said that: *“The East Asian NICs and their successful emulators are a powerful argument that a more activist, positive governmental role can be a decisive factor in rapid industrial growth”*. The statement signalled that the bank was ready to shift its perspective on the state's role in development.

Knowledge about the role of institutions in building societies, were of course not new, but its rediscovery in leading circles of economists showed that the political hegemony of neoliberalism had passed its zenith. Two studies in particular are worth mentioning, as they played important roles in redefining the conditions of development. Amsden (1989) and Wade (1990) consider the construction and execution of projects based on a symbiotic relationship

between the state and industrial groups. The state's contribution to such shared projects is crucial, but they also require developmentally engaged partners on the societal side.

Amsden argues that "late industrialisation" requires state intrusion. The first industrial revolution was built on *laissez faire*, the second on infant industry protection. In late industrialisation the foundation is subsidy, which includes both protection and financial incentives. The allocation of subsidies has rendered the government not merely a banker, but an entrepreneur, using the subsidies to decide what, when and how much to produce. In addition, the state must impose performance standards on the interest groups receiving public support. The combination of incentives and performance does not just shape the behaviour of existing industrial operations; it enables the state to coax into being a set of entrepreneurial groups that can serve as the societal side of a joint project of industrial transformation. Amsden studied Korea; Wade similarly concluded that "Taiwan's industrial success lay in the governed market, a series of policies that enables the government to guide market processes of resource allocation so as to produce different production and investment outcomes than would have occurred with free market policies".

In spite of the resurgence of institutional analysis in the 1990s, the starting point goes further back. The work of Weber, Polanyi, Gerschenkron, Braudel and Hirschman are the places to start. Polanyi (1957) provides a keen sense of the degree to which the markets depend on state action. Weber (1968) offers a powerful hypothesis as to what kind of internal organisation is likely to give states the capacity to construct markets and promote growth. Gerschenkron (1962) and Hirschman (1958) make it clear that state - society relations, particularly those that link states and entrepreneurial elites, are as important as markets. Braudel (1980) brings in the long-term historical relationships, and point to the increasing sophistication and globalisation of the market mechanisms.

More recently, Evans (1995) advanced a sociological theory on how economically underdeveloped states successfully industrialise. Evans asserts that the more state agencies approximate Weber's ideal type bureaucracy (e.g., insulation from external interference, a merit based system of recruiting experts, an ample security and reward system) the more they can successfully spur economic development. But Evans further suggests that bureaucratic autonomy is insufficient. Bureaucracies must also become embedded in concrete social ties that bind them to society. How bureaucracies provide institutional channels for continual negotiation of goals and policies with external groups is crucial for development.

Evans identifies a combination of four major roles that the state plays in economic development. As custodian, it provides protection, policing and regulation of infant industries, as midwife it attracts private enterprise into new sectors by subsidies, tax breaks and other devices; in the husbandry role, it teaches, cultivates, nurtures and prods entrepreneurial forces that have been activated; as demiurge, it becomes directly involved in productive activities that complement private investment, only to denationalise later when industries are established. Evan's analysis thus provides a broad spectrum of examples of the meaning of institutional development, as the functions that sustain the state's activities in these four roles. As we can see, the difference is that where Evans concentrates on the state's role in institutional development, other researchers emphasise the market, and the private actors (as for example the previously quoted Braudel (1980) and Hirschman (1958)).

However, institutional development does not necessarily have to be confined to the institutions that generate growth. As an example, Putnam (1989; 1999) provides a rich expose

of institutions that sustain democratic development, including the institutions of civil society. The term institution may itself be closely linked up with the notion of society, and it may also be described as an emergent phenomenon, an illusive and subjective component in the social construction of reality (Castoriadis, 1987). It is a concept which is closely linked to the abstract formation of society (James, 1996).

The prevalent view is that institutional development is closely linked to the concept of development itself. In fact, taking a broad look at the development literature, it seems as if institutional development not only explains development, but is synonymous to development. When practitioners and policy makers in development cooperation nowadays turn to institutions, they are going back to the core issues of the reasons for aid.

Towards a definition of institutional development

Even though there is a widespread agreement at present on the importance of institutional development, a closer look at the literature (and praxis) reveals that the terms are widely disputed and quite problematic. Many use the term as synonymous to organisational development, others focus on markets, legislation, or even norms, values and aspects of culture in general. As this study has a practical task, we need a set of definitions. I suggest that the following working definition, which establishes a framework for analysing institutional development be used. This is obviously not easy to achieve since there are a large number of concepts and definitions around. A review of donor literature and current use of terms for institutional development present a relatively large number of alternatives, but the underlying processes - or rather the main concerns within institutional development share important similarities - despite considerable semantic pluralism. The aim is not to invent a new set of terms, but to stay as close as possible to what key and influential actors use.

It seems that most concepts refer to and can be classified along two variables: (1) the intended level of intervention (from individual and organisational to sectors and systemic levels), and (2) the type and composition of activities (from training and organisational development activities to establishing/enforcing rules and regulations for effective functioning of sectors and systems). This means that the concepts refer to particular set of activities at various levels of society. But there are not always clear cut borders between neither levels nor activities - and the broader concepts often encircle the more narrow. The terms express direction and intention - more than conceptual stiffness. Using these variables, one could thus distinguish:

(1) Human resource development (training and education) which are concerned with how people are educated and trained, how knowledge and skills are transferred to individuals, competence built up and people prepared for their current or future careers.

(2) Organisational development which seeks to change and strengthen structures, processes and management systems in specific organisations in order to improve organisational performance. There are variations between theories and strategies, but they have in "pure" form the following characteristics: (a) focus on individual formal organisations and particularly their internal functioning, (b) less attention paid to external contextual influences on performance, (c) concern with internal organisational changes, and (d) activities include education, training, advice.

(3) Systems development is a broader concept than organisational development. In addition to a concern with human resources and the development of particular organisations, it includes an emphasis on linkages between organisations and the context or environment within which organisations operate and interact. It is useful to make a distinction between three system levels:

(i) *The network and linkages* among organisations, which include the network and contact between organisations that facilitate or constrain the achievements of particular tasks.

(ii) The *regulatory environment* referring to the policy environment of the public, private and civil sector that constrains or facilitates organisational activities and affects their performance, including e.g. laws, regulations and policies.

(iii) *The value framework* - including the economic, social, cultural and political milieu in which organisations operate, and the extent to which conditions in this broader environment facilitate or constrain the functional capacity of organisations.

This description of institutional development underlines the interaction between micro (internal) and macro (external) factors determining how organisations translate their capacities into actual performance. Institutional development refers to change processes - located at any of these three levels - that improve the capacity of a social system to achieve its goals and objectives.

The question is if the definition contains too many levels? In the tables that I am going to use in latter parts of the report there will be five columns to categorise aspects of institutional development assessment. That is quite a lot, and it would seem desirable to reduce the number. On the other hand, if any further merging of the different categories only mean that they will have to be broken up again in order to explain and distinguish important aspects, then the simplification has no long run effect. Let me briefly consider the alternatives.

There is no doubt that human resource development often is an aspect of organisational development. Perhaps these two categories could be merged? When analysing organisational development, one would always include an analysis of the human resource development of the organisation. However, human resource development could be a number of other things than organisational development. Many projects of the banks are geared towards general human resource development, for example through school systems, vocational training and the like. The impact will be analysed with other methods than organisational development would be analysed.

On the other hand, an analysis of organisational development would often include an analysis of the organisational environment. So perhaps the two categories of organisational development and networks/linkages could be merged? I would not like to see them merged, as the criteria to evaluate a network are normally quite different from those applied when evaluating an organisation. The health and well being of the network is a distinctly different question from that of the health and well being of any one organisation that takes part in it. To focus an understanding of networks as part of the assessment of one organisation could be very misleading. A market system of competing and cooperating enterprises would look different if analysed from the point of view of an existing firm, than it would if analysed on its own merits.

Yet another possibility would be to merge the two categories "legislation" and "norms/values". There is a conceptual connection between the two; the legislation in society can be expected to reflect the norms and values of the citizens. On the other hand, there a number of norms and values that are not expressed in legislation and some of the tension and dynamism in institutional development arise when the two are not the same. It is worth noting that when an evaluator is to assess legislative development, he or she would use entirely different methods than when assessing norms and values.

Even though a broad concept such as "enabling environment" would seem to cover legislation, norms and values, and perhaps even organisational networks, it will always be necessary to distinguish them for practical reasons. An enabling environment is not a practically researchable concept; it needs to be operationalised. In the process of operationalising the concept, one of the first steps would be to distinguish aspects such as those mentioned above. The question is how much of a help it has been to have the broader concept, if it must immediately be abandoned?

It is much a matter of preference and intellectual temperament how many categories one should work with in the definition of institutional development. The table below illustrates how the different conceptual schemes relate to each other. On the right I present the five categories that I personally prefer to work with, and to the left of them a few different schemes. Those to the left incorporate the categories of the five-dimensional scheme in different ways.

Table 2.1 Schemes to represent institutional development

Scheme of two dimensions	Scheme of three dimensions "I"	Scheme of three dimensions "II"	Scheme of five dimensions
Organisation	Organisation	Organisation	Human resource development
			Organisational development
Rules of the game	Laws and rules	Linkages	Networks and linkages
		Laws, norms and values	Legislation, rules and standards
			Values, norms and attitudes

One criteria of what constitutes useful and workable conceptual categories would be to see how popular writers discuss institutional development. The Pulitzer Prize winning book by John Dower (1999) "Embracing Defeat: Japan in the Wake of World War II" is an exceptional analysis of institutional change and development. The scope of the discussion can be illustrated by the following quotation:

"They (the Japanese) lived in a world of tumultuous institutional renovation that touched their lives directly. Old laws were abrogated and new laws, civil as well as criminal, introduced. Basic structures of authoritarian control were undermined in fundamental ways. Land reform all but eliminated exploitative landlordism and rural tenancy. Electoral reform strengthened the bicameral legislature as a major political actor. Constitutional reform established the principle of popular sovereignty for the first time, guaranteed a more extensive range of human rights than even the U.S. Constitution, and set anti-militarist ideals at the very centre of the national charter. Labour reforms gave workers basic rights hitherto unknown. Educational reform liberalised curricula, promoted coeducational egalitarianism, and broadened access to the elite track of the universities. Reform of the civil code eliminated the legal underpinning of the patriarchal family system and strengthened the position of women in critical areas such as divorce and inheritance."

The quote illustrates that a common use of the word institutional development is complex, multifaceted and covers all the dimensions suggested above. It also illustrates that a best-selling book aimed at a general public can employ a sophisticated use of the concept, and be read by a wide audience.

In the course of this study, the ECG meeting in February 2001 took part of the draft version of this report, and devoted some time to the issue of a joint definition of institutional development. The following text was agreed:

Institutional Development Impact (IDI) Criterion: *the extent to which a project helps improve the ability of a country or region to make more efficient, equitable and sustainable use of its human, financial, and natural resources through: (a) better definition, stability, transparency, enforceability, or predictability of institutional arrangements; or (b) better alignment of the mission and capacity of an organization with its mandate, which derives from these institutional arrangements. IDI includes both intended and unintended effects of a project.*

So far, there has not been any common definition suggested within the field of development cooperation. Each of the bilateral agencies use their own definition, which is more or less established in development economics, and more or less spread in the organisations. It is interesting to note that the OECD/DAC, which has otherwise done important work to standardise evaluation terminology has not set any definition of institutional development. In the process of this study, I looked at the OECD webpage for definitions, in the OECD glossary on terminology, and interviewed staff of OECD headquarters, but could not find any definition. Perhaps the OECD/DAC would be the partner in development cooperation, which is in the best position to provide an authoritative definition?

Contested issues in institutional development

Having established a framework that illustrates what institutional development is, I will now explore what institutional development is not - and how the term should relate to a number of other concepts that are used in the discourse on development cooperation.

Institutional development and capacity development

Capacity development is also a term used in connection with institutional development. Indeed, many of the UN organisations identify capacity development as the product of institutional development. This is a typical question of what comes first - the chicken or the egg. The approach taken here is that the two concepts can be treated as synonymous. When institutions develop, the actors increase their capacity to pursue objectives and goals. Capacity refers to the ability of individuals, organisations and systems to perform their functions efficiently, effectively and sustainably, and capacity development includes all processes which seek to improve and strengthen such abilities at individual, organisational and systemic levels (Kruse et al, 1998). In this sense, capacity building could be seen as a term encompassing all kinds of institutional development. However, outside the world of the UN agencies, capacity would usually have a more narrow meaning.

Most of the time people use the word capacity they refer to some very specific skill, knowledge or aptitude. An automobile may have capacity to take 6 passengers, a diver can have the capacity to go down 60 metres below sea level, a workshop may have the capacity to do repair work on a Jaguar. Taken in this sense, specific capacities are resources that are needed to accomplish specific tasks. As institutions develop, they encompass more and more of such resources - and hence, capacities can be seen as the building blocks of institutional development. Adding capacity to some organisational entity leads to institutional development, which in turn enhances that institutions capacity to act.

Organisational development is a more limited aspect of institutional development

In the definitions above a line is drawn between organisations and institutions - a line which in practice is often blurred. It is for example, not quite sure whether North (1990), Amsden (1990), and others refer to particular organisations when they spoke of institutional development - as for example courts, property right registrars, etc. In this paper, and for the purposes mentioned above, I believe it is important to keep a distinction between the two terms; that is, between organisational and institutional development. Different methods would normally be employed to analyse organisational development, than could be used to assess institutional development.

One of the earliest researchers who used the term institutional development in development cooperation was Arturo Israel (1987). Identifying institutional weaknesses as one a major roadblock to development, he analysed the World Banks experience by sector and by country. Whereas economic progress requires institutions with the ability to make effective use of the human and financial resources available, developing countries government agencies, public administrations, and enterprises (private and public), often lack that ability. In this sense, the word institutions is used for all kinds of organisations, private or public, in a country.

Perspectives change. At present most agencies refer to different though interdependent processes underlying organisational and institutional development. Influenced by institutional economics several reports advocate that institutions represent "the rules of the game" in society - the norms and values which guide and constrain the behaviour of individuals and organisations and shape human interaction, while organisations are the actors or "players". The purpose of the rules is to define the way the game is played, while the objective of the players is to win the game by a combination of skills, strategy and coordination following the set of rules, or through efforts to change them.

Yet another tradition emphasises the network characteristics of organisations. Castel (1996) describes the development of organisational linkages as a distinguishing characteristic of the present, indeed as one of the main driving forces of social evolution. In words of the framework above, the institutional development is caught in the density of networks. Cooperider and Dutton (1998) focus on the extent and type of linkages in networks, and discuss the effectiveness of different kinds of networks. The idea of design variables in networks is also elaborated in Forss et al (2000).

Another more sociological orientation perceives an institution as an organisation infused with values (Selznik, 1957). An institution represents and is infused with symbolic values in addition to solve certain technical tasks. Putnam (1999) uses the term for values and attitudes that sustain social cohesion, the social construction of meaning, and the political culture of nations. There is thus a clear trend to apply the term institutional development for aggregate phenomena. At present it refers to the framework within which specific organisations are placed, in terms of values, regulations, and other organisations (networks). It is a wide and holistic concept, which implies an open-systems view of organisations.

Institutional cooperation is one of the means towards institutional development

Institutional cooperation comes in another category than the concepts above, and is one of several tools to be used in order to strengthen organisations by job-specific skill upgrading, procedural improvements, training and advice, etc. Institutional cooperation represents formalised long-term cooperation between two similar organisations to achieve mutual capacity strengthening. When institutional collaboration is successful, it results in increased organisational capacities in one or both of the organisations that take part. Whether this should also be seen as institutional development is a question, the answer to which depends on your definition. According to the definition used here (reflecting the 5 levels of change) it certainly is, but it is only one aspect of institutional development. Institutional cooperation may, for example, have strengthened the customs authorities in a country, but whether it has also strengthened networking capacities or legislation is another matter.

Institutions and markets

Is the market an institution in its own right, or is the market as an institution more properly seen as a network of organisations? To pursue the discussion, I need to return to the definition of institutional development. In their seminal work on institutional arrangements conducive

to growth, North and Thomas (1973) focus on "those arrangements, which enable units to realise economies of scale, to encourage innovation, to improve the efficiency of factor markets, or to reduce market imperfections". So it seems as if we have to turn to a market analysis to discuss the institutional development. This may seem highly theoretical; but then again, it is a theoretical question so it needs a theoretical answer.

The market is in essence a meeting, when forces of production meet those of demand, of consumption. However, it does not make sense to describe any meeting of buyers and sellers as a market. It is only when that meeting is sustained, when it takes place repeatedly, when it is confirmed by values and attitudes, and supported by an organisational framework, that it has become an institution. Customs authorities, litigation mechanisms, the judiciary, the price supervisions, the standardisation bodies are organisations, and as such part of the network which defines the institution of the market economy. The market is not only the meeting of buyers and sellers, but also that network of organisations that have different roles to play in facilitating and regulating the activities of buyers and sellers.

The capital market is one of the most important parts of the market economy. Even that system consists of a number of separate organisations, ranging from the Ministries of Finance, the public bodies such as tax authorities, central banks, bank inspections etc, and onwards to economic research, consulting firms, public and private banks, the financial press, and so on. Any one of these institutions can be the object of organisational development, but it is quite clear that the link between those organisations improved performance, and the performance of the network, is at best weak.

One of the key issues in development lies in finding out when things "get better". What characterises a more - or better - developed market? This is not an easy question to answer, but some indications can be suggested. First, it is competitive. A more developed market has more actors, and actors who compete with each other. Second, it is an open market, new firms enter and others go bankrupt and exit. There is turbulence in a more developed market. Third a more developed market is more diversified market. Finally, change is generally taken to be a sign of a more developed market. Schumpeter wrote of the creative destruction of the capitalist entrepreneurial system, whereby market economies generate increased wealth through the forces of change. Analysing the health or progress of a market is thus a different issue that analysing the development of any of the organisations that act on the market.

Institutional development; the private sector and public administration

It is not uncommon to see misconceptions about how institutional development relates to ownership. There are those who see institutional development mainly as a question of finding the right role and mandate for the public administration. There are also those who see institutional development as an aggregate term for privatisation and transition from centrally planned economies to market economies. In either case, the issue is more complex. Looking at the above mentioned dimensions of institutional development, the issue of ownership could in some cases be of importance, in others simply irrelevant.

Let us first look at interorganisational networks. Networks are not owned as such. The organisations in the network are owned, but not the links between them. Ownership is still interesting, and a network between public organisations would probably be quite different

from one between private organisation. In any given network, it would be interesting to know who controls the participating organisations (the public sphere, cartels, mafia-like structures - to mention some possibilities) or to measure how fragmented ownership is. Many of the organisations that are usually considered as part of institutional development can be owned privately, publicly, or have mixed ownership. It is an open question which solutions work best. If there is any common ground between researchers, this could be formulated "any best solution depends on circumstances and varies from one social setting to another".

One of the dimensions of institutional development was defined as the regulatory environment; that is, the judiciary, the standard and norm setting bodies, and possibly also policy formulating bodies. It can of course be contested whether policy formulation in all instances belongs with government; in fact, much of the political debate revolves around whether citizens should confer this right to a central government. Even if civil law is often seen as a part of the public administration, other legal developments can be vested in private bodies, or in cooperative mechanisms under joint public and private authority.

Values, attitudes and social norms are beyond public or private ownership. But values are formed, and they can be formed by private, commercial interest, or by public bodies. Many public organisations actually have a mandate to shape values and attitudes among citizens. In a broader sense, the political arena is increasingly concerned with shaping values and attitudes (Beck, 1996). However, even though both public and private actors have roles to play in this respect, the issue of institutional development in this dimension is not first and foremost one of public administration - or the private sector.

In conclusion, institutional development must be concerned with issues of ownership, but it goes beyond the question of how specific organisations are controlled. An investigation of institutional development would relate ownership to other issues, such as the function of markets, the social dynamism, the emergence and formulation of values. Institutional development is rooted in an older and more holistic concept of social studies than is either the tradition of public administration, or neo-liberal economics.

Mission and policies in respect of institutional development evaluation

It is quite obvious that the multilateral development banks have a mandate to pursue institutional development. For all practical purposes, institutional development is synonymous to development itself; explaining institutional development as a factor in development very soon turns into a discussion of which comes first - the chicken or the egg. The founding charters of the organisations, the proclamations in strategy and policy papers, presentations of the mandate, all bear witness to a concern for institutional development impact.

The question to pursue here is not whether that mandate exists or not, but rather to analyse whether and how the mandate has been extended to the evaluation function.

The policy and strategy for evaluation of institutional development can be taken to consist of a number of steps. These steps can be seen as the design variables of strategies and policies, that is, points at which there is a choice on how to approach the subject matter.

The first step is concerned with the definition of institutional development. The intellectual bend visible in communications and discussions suggests that there is a high degree of commonality in practice, but I have not found any common definition of institutional development yet. The World Bank's draft guidelines of May 2000 contain a definition, but other guidelines (such as those of AsDB, of August 2000) do not develop a full definition.

The second step in policy and strategy would be to establish the importance of institutional development in comparison to other issues and criteria that must be attended to by an evaluation team. There are no common approaches to this, but it is implicit in evaluation guidelines and in terms of reference that institutional development often is a secondary priority. The AsDB Guidelines, for example, describes the basis for an overall assessment based on five building blocks; relevance, efficacy, efficiency, sustainability, institutional development and other impacts. Our subject thus falls into a category of "others" at the very end of the list of issues to be covered in evaluation. In the overall ranking system, which is suggested in the guidelines, this category is accorded a weight of 15%. Even though institutional development is recognised as important, this operationalisation relates it to something of secondary and minor importance. The EBRD, on the other hand, puts "transition impact" at the top end of what evaluators should look for, which illustrates how institutional development get real and practical priority.

The third step in the development of policies and strategies would be to suggest a framework of indicators, which can be used in evaluation. So far, I have not seen any development of this sort, and I will turn to the concept of indicators in chapter 4 of this report. Suffice it to say, that an evaluation strategy would be a tool to define the scope and utility of indicators, and to suggest how indicators could be practically used by evaluation teams.

The fourth step in policies and strategies would be to develop the outline of an aggregate reporting format. This will be the subject of chapter 4 as well, where I will turn to a description of what the banks do in this respect. The World Bank's draft suggests a ranking (or rating, the terms are used synonymously) scale, so do the AsDB's guidelines, and the EBRD also employs a common ranking system for institutional development impact. Among the other banks, the policy is rather to tailor-make the rating, as well as the use of indicators, to each evaluation. Whether it is an implicit or an explicit policy is not quite clear.

A fifth step in policies and strategies would be to consider the sharing of lessons learned, possibly of sharing basic conclusions, methodological approaches and the like. However, I have not found any evidence of policies or strategies that apply specifically to institutional development. The practices in this respect relate to evaluations as such, and there are no structures or processes in place that could single out one dimension of all those that are applied in an evaluation.

In sum, the work of the evaluations units naturally focuses on the work processes that constitute an evaluation; from the very beginning of formulating a problem, to the end when the results of the evaluation are put to use. Institutional development impact is one of the dimensions in this respect, and though all recognise its importance, the practical focus often lies on the more direct and immediate concerns.

Evaluation resources are scarce, and to analyse institutional development impact is something, which often requires considerable time. In practice, the result is that other subjects get priority and the evaluators respond to demand from desk officers, project managers, and sometimes

executives and the board - for knowledge of short-term results. An important function of policies and strategies would be to define the real priority of institutional development, and to set the standards for how resources should be allocated among the different subjects that arise during an evaluation.

Chapter 3. Coverage of institutional development

In this chapter, the entry point to the subject of institutional development will be analysed. A sample of 31 evaluation reports (see annex 2), produced during the last ten years, have been submitted by the evaluation units of six of the seven MDBs. The first question that will be addressed is: What do these evaluation reports assess? The framework that defines institutional development, which was introduced in chapter 2, will be used to answer the question. Each one of the evaluations has been read, and I have then listed which level of institutional development it reports on.

This procedure makes sense as part of this particular study, as it will help us to map out the entry point to institutional development. However, the reader must remember that this analysis does not reflect the quality of the evaluations themselves. I have not looked at the terms of reference, so I do not know whether omitting a certain level of institutional development means that the evaluation team did not find any evidence in this respect, or whether they were never meant to pursue such questions. The subject of the loan agreement may in fact not be relevant for a full coverage of institutional development issues.

In one sense it is thus quite an "unfair" review of the evaluations. I am using a terminology, a map of institutional development issues, which was not available to the evaluation teams, or to those who commissioned and received the evaluations. On top of that, the assessment builds on an understanding of institutional development, which is not necessarily shared by the MDBs, at present it is merely forwarded as a possible framework to identify the issues of institutional development.

However, I am confident that the framework as such is broadly confirmed by the present state of the art in thinking on institutional development. It is indeed a broad subject, which includes human resource development, organisational development, and "systems" development (networks, legislation, values). But it must be emphasised that this review of entry points should be understood in descriptive terms, the fact that an evaluation does not cover that many aspects of institutional development does not mean that it is a poor evaluation, nor that it would have been better with another coverage. With these caveats in mind, let us now turn to the evaluations.

The focus in institutional development analysis

Table 1 in annex 3 summarises the assessment of what these 31 evaluations cover. However, in the process of reading the evaluations, I found it difficult to simply note whether an issue was treated or not. Several times, a more finely tuned scale seemed to be called for. As a consequence, three different modes of treating an institutional development subject were distinguished. When an issue of institutional development has been treated fully, it means that several pages, perhaps a chapter, was used to describe a situation, to analyse shortcomings, achievements, and/or options for the future. When an issue has received some treatment, it has

been mentioned, but there is not a fully argued case either for how that aspect has affected the bank's operations, or what the impact has been. When there is no reference to that aspect of institutional development that means the issue is not referred to at all; there is no mention of it. This is not a very finely tuned scale, but it will serve as a starting point for the analysis.

As an example, the evaluation of *Telecommunication Legislative and Regulatory Development in Lithuania* is primarily devoted to describing and analysing the sector licensing policy, basic regulations as well as regulatory instruments, tariffs and other service charges. This is also the focus of the intervention, and hence it would be strange if the evaluation did not cover these issues. It is a good example of where an evaluation in detail, and with a comprehensive approach, deals with one of the aspects of institutional development.

The evaluation of *Effectiveness and Impact of ADB assistance to the Reform of Public Expenditure Management in Bhutan, India, Kiribati, and Lao People's Democratic Republic* can be taken to illustrate when it seems appropriate to say that an issue has received some treatment. As for the human resource development, the evaluation specifies (p. 15) for example; "some skills were developed for analysing state owned enterprises"...or "skills were developed in macroeconomic modelling and the potential exists for NAO to function as an independent office". But as the evaluation does not say much more than so, it is not really a full treatment of the subject. Again, I would like to emphasise that it is not certain more was needed; the focus of the evaluation was on other issues (in this case, efficiency and effectiveness).

Table 3.1 presents a summary of the entry points to institutional development analysis. Half of the evaluations contain a full analysis of organisational development issues; there is no doubt that organisational development is the topic, which is most commonly and most comprehensively treated in the evaluation reports. This is not surprising, as the loan or grant assistance usually targets one particular organisation. Hence, evaluations tend to analyse the status of that organisation - be it a ministry, a bank, a public authority, or a firm, a hospital, or indeed a number of health centres, or vocational training schools.

Table 3.1 Summary of entry points to institutional development (n=31)

	Human resource development	Organisational development	Networks and linkages	Laws and rules	Values and norms
Full treatment	5	15	9	3	-
Some notes	8	13	15	13	6
No analysis	10	3	7	15	25

It is only a small share of the total of 31 evaluations that contain a "full" treatment of any of the other aspects of institutional development. 9 of the evaluations (27%) contain extensive

analysis of organisational development issues. These tend to be loans or grants that have had a heavy emphasis of technical assistance and training. The evaluations devoted a significant share of time and space to describing and measuring the technology transfer which was expected, and to some extent delivered. Somewhat fewer evaluations contain a full analysis of the regulatory environment or of human resource development. None gave the subject of norms and values any full treatment.

That organisational development is the most commonly analysed aspect of institutional development is further emphasised by looking at the number of evaluations that have "some" analysis of this subject. Another 13 evaluations do treat organisational development, which means that it is only 3 out of the 31 evaluations that contain "no" analysis at all of organisational development issues. Otherwise, the entry point to institutional development, which is most commonly treated in this manner, is the level of networks and organisational linkages. 24 of the evaluations, just above 70%, have at least "some" form of analysis at this level. It is also quite many, 18 evaluation that contain "some" analysis of the regulatory environment.

The blank fields in coverage of institutional development

The information contained in table 3.2 can be used to illustrate the main entry points to institutional development. A simple scoring of the different levels (0 for no analysis, 1 for some notes, and 2 for full treatment) yields a list such as presented in table 3.2. The dominance of organisational development as the entry point to institutional development is clearly visible. It is equally visible that the evaluations seldom treat values, norms and attitudes.

It is interesting to remember that a full coverage of institutional development issues in all the selected evaluations would have yielded a total score of 62 for that particular subject. The use of the scoring instrument thus points at a strong focus on one aspect of institutional development, and also indicates that there is little information about the other aspects.

Table 3.2 Ranking the entry points to institutional development

Entry point	
Organisational development	43
Networks and linkages	33
Laws and rules	19
Human resource development	18
Values, norms and attitudes	6

At this point it may be useful to look closer at what an evaluation might have to report on values, norms and attitudes. The evaluation of *Bank Technical Assistance for Capacity*

Building in Indonesia was one of those, which contained some treatment of values, norms and attitudes. The evaluation discusses attitudes in cooperating organisations before technical assistance is fielded, it discusses values and attitudes in respect of ownership, commitment and motivation to change, and measures these with the help of a questionnaire which was disseminated to a large number of respondents. An evaluation, which decides that values and attitudes should be analysed, can certainly do so, and in this case the analysis is very relevant for the topic of evaluation.

The question is if these 31 evaluations are typical for the MDBs? That question cannot be answered now, but my guess is that they are quite typical. In all probability, the standard evaluations contain less analysis of institutional development, and in particular I would think it is uncommon that other evaluations go beyond the topics of human resource development and organisational development. However, it is perhaps more common for evaluations to contain a full treatment of human resource development issues than is evident here. In this sample, there could be a shift from these "basic" aspects of institutional development, to the more complex analysis of networks and regulatory environments.

Yet another question is if it is desirable, indeed if it would be relevant, for the evaluations to contain additional entry points to institutional development? On the one hand, there is a need to focus evaluations on the issues at hand that are the most relevant. If networks, regulatory environments, values and norms, were not targeted when the intervention was designed, then it is probably not very helpful to cover these issues at the stage of evaluation. On the other hand, there is a difference between analysing impact from an intervention, and explaining what happened during the course of implementation. It is common that the performance of projects are more affected by institutional development in the environment than one estimates when a project is started. In order to learn the lessons why things happened as they did, a more thorough assessment of institutional development may turn out to be necessary.

Chapter 4. Measuring institutional development

The purpose of this chapter is to look at what the evaluations actually say about institutional development. The chapter is thus concerned with indicators, measurement and scales. Evaluation is an inquiry into the worth or merit of something, in this case institutional development. The indicators and measurements, and the philosophy behind them, are critical to understand in order to develop the accountability systems regarding institutional development.

A strategic distinction in how evaluations cope with institutional development

The World Bank Performance Audit Report *"Burkina Faso Agricultural Research Project and Agricultural Extension Project"* rates institutional development impact. The ratings are presented on page iii of the report together with other principal ratings. The report covers two projects and both were rated. Page 2 of the executive summary presents the ratings, and notes that the ICR rated institutional development as "substantial", but the PAR rated it as modest *"on the grounds that, although some of the structures and processes established are important, it was too much oriented towards a top down system."*

In the main text of the report, chapter 5 covers ratings. Paragraph 5.5 (in respect of the research project) says: *"The audit agrees with the ICR that institutional development impact was substantial. Institutions and associated processes are now significantly stronger."* Paragraph 5.10 (in respect of the support services project) says: *"Institutional development is rather modest, indicating that the project increased the country's ability to effectively use human, organizational and financial resources to a limited extent. In some respects it could be rated substantial, but there are questions about the relevance - whether the approach was the appropriate form of institutional change to introduce, even by the standards of the time. Thus, there are institutional opportunity issues to be considered - the counterfactual. Nevertheless, some of the institutional structures and processes put in place would still be appropriate for other extension services in the future"*

The above quoted sentences are essentially what the PAR has to say about institutional development. It is certainly everything, which mentions institutional development as a subject. The section on findings and lessons learned contains a brief discussion of land policy issues, privatisation, legal protection of contracts, etc. However, these are mentioned as subjects that hold importance for extension services in the future, but there is no analysis of what the consequences are, or how they relate to the projects, or what action should be taken.

The point I wish to make is that this report illustrates the rating of institutional development. According to the Bank's policy, the task of the evaluators has been to rank institutional development impact on a scale from unsatisfactory to substantial. The motivation for the ranking is no less, and no more, than the sentences quoted above. The 26 pages of the report thus have relatively little to say about institutional development.

The World Bank's Country Assistance Review "*Philippines: From Crisis to Opportunity*" presents another - and a different - example of how institutional development may be handled in an evaluation report. In this study there is no rating, neither of institutional development impact nor of anything else. That does not mean the report is not evaluative; on the contrary, rubrics such as "Assistance Strategy: Satisfactory, but Uneven and Below Potential" clearly reveal the evaluators opinions on the programme.

The 55 pages in total contain references to various aspects of institutional development on almost every page. Specific chapters deal with policy weakness, peaceful revolution, reforms, instruments and partnerships, to take some examples. The main thrust of the analysis is to explain the context of the Banks assistance strategy, and to explain what the Bank has done in terms of institutional development, and how institutional development has affected the Bank's lending operations.

Where the country report discusses how the institutional environment has affected the banks lending operations, the audit report discusses how the projects have affected the institutional environment. The country report has a lot to say on its subject, the audit report little. Reading through the Philippines report gives insights into current development issues, and a good understanding of why the Bank's assistance strategy met with limited success. However, the audit report from Burkina Faso did not really tell anything about institutional development issues in that country (or in relation to the project).

The comparison is not quite fair, and should not be seen as a criticism of the reports per se. The point is that institutional development can either be treated as a broader context within which projects and programme can be understood and their performance explained, or it can be treated as a specific impact category. In the latter case, the question is whether the project had an impact on its institutional environment. In Annex 3, table 2 there is an assessment of whether institutional development is treated as a factor to be understood affecting project performance, or whether it is primarily understood in terms of a created impact. The annex table is summarised below. The differences I have talked about above can be related to the major distinction between formative and summative evaluations. A summative evaluation is usually defined as one, which is conducted after completion of a program, and for the benefit of some external audience or decision-maker. A formative evaluation is typically for the staff of a program, with the intent to improve, and it is often conducted during the development or improvement of a program. Though there are differences between those terms and the above-mentioned treatment of institutional development, I think the differences are aptly illustrated by those terms.

Table 4.1 Differences in treatment of institutional development in evaluation reports (number of evaluations)

Formative approach	Summative approach	Formative and summative
9	6	16

In order to illustrate the difference, I will quote at length from two evaluations. The AfDB thematic study of Water Supply is an example of how the evaluation team discusses the context of the projects, the institutional environment, and uses characteristics of these environments to explain the problems encountered in implementation:

”5.1.2 Institutional weakness in RMCs was assessed as one of the main reasons for poor water supply and sanitation service delivery in utilities. Efforts made at institutional development and training in Study projects are noted in Annexe 9. Inefficient operation, weak commercial and financial management, poor maintenance, lax billing procedures, and excessive political interference and a lack of commitment to reforms and autonomy were common maladies noted in the Study projects. The legislative, regulatory and organisational frameworks in the sub-sector were generally weak. Even though there were some notable exceptions, the governments owned and operated most of the water and sanitation infrastructure, and private sector participation was limited. Those charged with the responsibility of delivering water supply and sanitation services seldom had managerial and financial autonomy to function in an environment conducive to efficiency. There were compulsions to deliver services below costs. On the other hand the managers of the publicly owned and managed utilities were not accountable and there were no clear performance indicators to measure their performance.

5.1.3 The few components for institutional development that were included in the reviewed projects largely comprised staff training components for staff directly connected with implementation or operational maintenance of assets. A general lack of overall manpower development or staff training policies was noted and “bits and pieces” training programmes were invariably linked to and dependent on external funding. Institutional autonomy was also identified as an objective of some of the projects. For example, reforms for achieving institutional autonomy were undertaken in some projects such as the Bauchi Township and the Niger State Water Supply projects in Nigeria and the Nouakchott Water Supply project in Mauritania but failed to achieve the desired results and impact, in the absence of political commitment, a practical road map and inadequate Bank follow up. Similarly, staff development and training in areas such as accounting, billing and revenue collection, computerisation of customer accounts and billing and tariff reforms, to achieve a reasonable degree of financial viability, formed part of the objectives of most projects but the results were mixed.

5.1.4 Healthy and effective sector institutions are the key to sustained operations in the sector. Institutional capacity for project implementation and system operation depends on organisational autonomy, leadership, management and administration, commercial orientation, consumer importance, technical capability, staff development and retention and linkages with external institutions. Conversely, weak institutional capability is seen in all sectors, including the water supply and sanitation sub-sector, as an inhibiting critical factor in implementation and sustainability of projects. Weak institutions also hinder effectiveness of development assistance. Efficient absorption of the assistance is jeopardised.”

From the same Bank, there is the evaluation of credits to Nigerian Agricultural development. The quote below indicates that this is an evaluation, which primarily discusses what impact the project has had on institutions, rather than the other way round, as the quote above illustrated:

"4.4.1 The injection of the external resources provided under the two LOCs into the Nigerian agricultural sector has had a significant impact on that sector and on the economy as a whole. Based on interviews held with, and questionnaires administered on the randomly selected beneficiaries during the field visits of the PPER mission, the objectives of the two lines of credit were, generally, attained, considering the fact that NACB outreached thousands of small-holder farmers, particularly women in rural areas. Access to rural credit has always been a constraint to development in the sub-sector. For that reason, the project had the institutional development impact on national capacity, and was rated satisfactory."

It is obvious that the reports cater to different audiences. The difference between the World Bank PAR from Burkina Fasso and its Country Assessment from the Philippines also reflects that they have different objectives, and cater to different audiences. So the treatment of institutional development follows logically from the intended readers. But the fact remains that within the broader title of "evaluation" there are strategic differences to the subject of institutional development. The reports have different uses. Could it be said that one approach is more useful than any other is?

If I am allowed to present a personal opinion, the evaluations following the summative approach actually tell me little about institutional development. They rank an impact, but they do not provide the flesh and bone of what has happened or why. They do not contain anything, which might provide ideas for project design, or guiding action, on other projects. I would suppose that they contain nothing new for project managers, task officers, or borrowing agencies. Others might agree or disagree with the ranking, but there is no more to it, as the judgements are only supported by a few highly general observations. There is not any substantive discussion. But there could be; it is not hard to imagine how the evaluation team could have presented the information that led them to the judgement in the first place, and that they could have taken the reader through the analysis to allow him or her to arrive at similar (or different) conclusions.

Evaluations that take a formative approach, or that combine the two, hold far more interest (which is, again, a personal reflection). These evaluations enter into a substantive description and analysis of institutional development. It is possible to learn from this, to draw you own conclusions, and to find ideas that can be of use in other contexts. The EIB evaluation of financial markets in Spain, is an example of a very comprehensive treatment of markets, institutions and legislative developments. It can be read for its own sake, by anyone who is interested in the institutional development of capital markets.

But there are different uses. Those evaluations that take a formative approach to institutional development cannot be used to present results to management, to a board, or to an external audience. The summative approach, and the related ranking systems employed by some banks can be used for aggregate reporting on how many projects that have a satisfactory institutional development impact, or a transition impact. There is thus a level of utility above and beyond the contextual information provided in the reports.

Broadly speaking, the approach illustrated by the formative evaluations generate information that can hold much interest for task managers, project staff, borrowers, and a more general audience with an interest in development issues. The summative approach focuses on providing aggregate data for performance management primarily directed at senior levels in the organisation, and the board, as well as control by these and external stakeholder groups. Is it possible to combine the two approaches and thus to maximise the utility of evaluation?

Once the basic parameters of institutional analysis are put in place there is nothing to stop the evaluators from continuing to rank impact. In theory this is simple, but in practice it may be harder. The reasons are philosophical and methodological. With a rich contextual description of the interaction between a project and its institutional environment, the ranking exercise may become harder, in spite of the prevalence of empirical data. More data generally make things more complicated, and consequently there would be a strong tendency for evaluators to opt for the middle ground on ranking scales.

Using indicators to assess institutional development

Irrespective of the approach taken, evaluations need to develop indicators of institutional development. As the subject is so wide, evaluation teams are bound to use indicators, whether they say so or not. In order to make sense of reality; they use indicators, but that use may be highly implicit. An implicit use of indicators can make good practical sense, but raises issues of credibility and validity.

The first question to address is what an indicator is all about. Indicator is the noun derived from the verb "to indicate". To indicate means (1) to point out or to, and (2) to state briefly, show indirectly (The Merriam Webster Dictionary). In management use, an indicator is an item of aggregated data, which easily and briefly provides a lot of information about an issue. An indicator is a measuring device that allows the evaluator to clarify and measure a concept. Indicators make a concept more tangible, gives something to measure, and allow measuring to take place over time. This being said just to make sure that the concept is understood in the same way by all.

When looking at some of the indicators mentioned in the evaluations I get a bit confused - or rather, I fail to understand how the issue described as an indicator can be defined as an indicator. As an example, I would like to point at the evaluation "*Assistance to the Reform of Public Expenditure Management*". Annex 1 lists a number of indicators in respect of issues around capacity building components, such as:

"Role of Government in designing TA; Availability; Quality; Release process; Etc...."

My point is that such broad concepts are quite a long way from practical utility as indicators. In a very simple context, quality may for example be assessed through a variety of indicators, but quality as such is normally too complex a concept to be an indicator of something. On the contrary, quality is usually measured through a set of indicators (if we were speaking of a car, some of these indicators could be safety, comfort, acceleration speed, design...). The issues mentioned above are more examples of topics to be analysed than indicators. The latter would need to be far more specific and concrete. It is not uncommon that evaluations speak of indicators, when in fact they identify areas that need to be researched.

Indicators can be quantitative or qualitative. The difference between the two is sometimes subtle, and they are often mixed together. In general, quantitative indicators are numeric representations of a concept (for example, the number of doctoral dissertations written as a result of a research project). Qualitative indicators are less tangible. They are often

individual's perception of a situation and are not always easy to count (for example, descriptions of the ways people found a research project useful).

Box.4.1 List of the evaluations that have used some form of indicators to assess institutional development.

Effectiveness and Impact of Asian Development bank Assistance to the Reform of Public Expenditure Management in Bhutan, India, Kiribati, and Lao People's democratic Republic (annex 1, p.19 and 20)
Effectiveness and Impact of Training in Education projects in Indonesia (annex 5, page 27)
A Preliminary Overview of the IDB Experience with Sector Lending: First Phase (in particular pages 9,19,20,25,26,27,29, 30)
Summary Evaluation Report IDB activities in Support of Public Sector Reform (Annex 1)
Social Investment Funds in Latin America (page 93 and 94)
Mexico: Evaluation of the Mexican Health Services Upgrading Program (indicators throughout, e.g. page 30)
Western Siberia Oil and Gas Rehabilitation (Annex 11 and 12)
Vneshtorgbank. Standby Facility (pp 18 1and 19)
Mining Privatisation (annex 2)
Small Municipalities Environment. Estonian Water Company (annex 3)
M1-M15 Toll Motorway project (Appendix 9)

Indicators can be either direct or indirect. A direct indicator measures the concept itself - if you want to know staff size, for example, you count the number of staff. Indirect indicators measure something that happens as a result of the concept. An indicator of impact following dissemination of information will tend to be both indirect and qualitative, but some may be indirect and quantitative. It is prudent to remember that indicators often point out a paradox or a dilemma. For example, an indicator such as "diversification of funding" could be seen as a positive or negative. On the one hand, it is a sign that the organisation is financially viable and not reliant on one donor. On the other hand, having several donors can lead to fragmented programs, lack of coherent visions, and cumbersome administrative procedures. Indicators can lead in different directions, but often there is a balance to be struck, and management needs to consider the pros and cons of diversified funding. The indicators may help both to define the issue and to allow decision-makers to define the right balance.

Box 4.1 lists the evaluations in the sample that use some form of indicators in their analysis of institutional development. It is actually only 11 of the 31 evaluations that make explicit use of indicators. The EBRD has introduced a format of indicators that is used to assess transition impact, and this is applied in a modified form on most evaluations. Otherwise it is interesting to note that the formative approaches use indicators more frequently than those evaluations that take a summative approach. Neither the World Bank nor the AfDB evaluations contain indicators to assess the institutional development impact. On the contrary, it is the more qualitative and formative discussion in the AsDB and IADB evaluations, which make extensive use of indicators. This is rather odd, as the literature on indicators and performance management suggests the opposite.

Box 4.2 Indicators to assess transition impact. From the EBRD evaluation reports.

Type of impact	Short term verified impact	Longer term expected impact	Justification
<i>Step 1. Project Affected Change at Corporate Level</i>			
Private ownership			
Know How			
New standards for Business Conduct at Enterprise			
<i>Step 2. Transition Impact at Industry Level and in the Economy as a Whole</i>			
Enhanced competition			
Frameworks for markets, institutions, laws..			
Skills transfer and dispersion to the industry and economy as a whole			
Demonstration effects			
Setting new domestic standards			
Market expansion			
Overall rating			
Justification			

The evaluations of the AsDB and IADB that use indicators do not follow any standardised format. In these reports, the evaluators have designed indicators that will be of use in their particular inquiry. These indicators are mainly quantitative, even though of different quantitative natures. The indicators used in the IADB evaluations are based on statistics; e.g., numbers of hospital beds, inhabitants per medical personnel, and the like. The data to substantiate the indicators are fetched from other reports, from statistical sources and the like. The indicators used in the AsDB evaluations, on the other hand, are often based on surveys in the evaluations. The indicators reflect responses to questionnaires and interviews. When the majority of the respondents express that they view the performance of technical assistance personnel as satisfactory, this is taken to be an indicator of success.

The EBRD indicators are mainly qualitative. They are brief, but nonetheless qualitative descriptions of an organisation, of change, legislation, or values and attitudes. As the indicators in the other reports are all tailor-made for each evaluation, I see no point in replicating them here. However, it holds more general interest to look at the indicators of transition impact. These are reproduced in Box 5.2. It should be noted that there is some difference between evaluations, but they are quite similar nonetheless.

The approach has several advantages; it enumerates several indicators, which increases the sophistication of the measurement, provides more complete information; and forces the evaluators to provide justification for the rating of each indicator, as well as for the overall rating. In the reports, these justifications usually run to some 50 words for each indicator. In other reports, where the evaluators have ranked institutional development impact, I have often wondered why they chose a particular rank. With this format of multiple indicators, and justifications at specific as well as at aggregate levels, the reader is in a better position to understand why a certain rank is given. Hence the transparency is increased, and the evaluation is more likely to inspire trust among readers.

Whereas I have not come across any other systematic use of indicators as a means to assess institutional development than that of the EBRD, there could still be many individual evaluations that have a more sophisticated use of indicators. However, the 31 evaluations indicate that there is considerable scope to make better and more systematic use of indicators in the course of evaluation, both for the purpose of formative and summative uses of evaluation.

The use of scales and the measurement of impact.

In this section I will turn to the question of how the evaluators present their conclusions in respect of institutional development. There are two approaches; some evaluators present a more or less lengthy description of the institutional development impact. These can either be quantitatively oriented (as in the AsDB evaluation of Assistance to the Reform of Public Expenditure Management) or it can be more qualitatively oriented (as in the IADB evaluation of the Policy-Based Loan Portfolio). Other evaluators proceed rapidly to rank the institutional development impact. Out of the 31 evaluations perused in the course of this study, a total of 13 use a scale to measure institutional development impact.

The word measurement stands at the centre of attention. It is of course possible to assess institutional development (either with a formative or summative approach) without measuring it. Measurement must be understood as the procedure by which the evaluators obtain symbols which can be used to represent the concept to be defined (Ackoff, 1962). The purpose of measurement is to represent the content of observations by symbols, which relate to each other in the same way that the observed objects, events, or properties are or can be. Ranking involves what is known as an ordinal scale. There are those who limit the word measurement to operations that involve an interval or ratio scale. However, it is more common to speak of measurement in a wide sense, involving both ranking and assessments on nominal scales.

The banks that encourage measurement of institutional development impact do so with rank order scales. In the past, the scales were different (table 4.1 show the scales used in the 31 evaluation reports). As a result of recent methodological developments, the MDBs now employ the same scale, which is the one in the column to the right.

The design of the scales raises a number of interesting issues. Is it better to have an odd or an even number of alternatives? Should there be a directly negative option, or is the worst possible outcome that institutional development is "negligible"? If direct value expressions are used, for whom is the outcome deemed satisfactory? What are really the differences between "high" and "substantial", or between "modest" and "negligible"? Which forms of bias in reporting are inherent in the scales used?

Table 4.1 Scales and the ranking of institutional development.

	EBRD	AfDB*	World Bank**
Highest	High	Highly satisfactory	High
	Medium	Satisfactory	Substantial
	Low	Unsatisfactory	Modest
	None	Highly unsatisfactory	Negligible
Lowest	Negative		

* The evaluation reports rank institutional development on a score from 1 to 4, with the descriptions noted above; that is, 1= Highly unsatisfactory, and so on.

** In some evaluation reports the signification "partial" is also used, which is probably synonymous to "modest".

It is only one of the ranking scales, which recognises that institutional development impact, may be negative. I am reminded of the famous history of Europe by Norman Davies, whose comments on mediaeval France seem pertinent: *"Yet Phillippe le Bel was the author of durable institutions: With the aid of his légistes or legal counsellors, he found all manners of pretexts to fleece his subjects, to institutionalise his depredations, and to cloak them in the guise of a national consensus"*. Recent studies of aid effectiveness in Southern and Eastern Africa confirm that the institutional development impact can be negative, even if intentions of the funding agencies were the best.

Sound evaluation practice would suggest that a scale should be symmetrical; that is, in the EBRD case it should actually be possible to discern between low, medium and high negative impact. Practically speaking, I do not think the development assistance community is ready to recognise, let alone systematically explore, that the impact of activities can be that diverse. It would also be quite cumbersome to have ranking scales with 7 rank orders (3 on either side of none). Nevertheless, it is important that ranking scales recognise the possibility of zero or even negative impact, because these outcomes are definitely within the scope of what is possible, and it would be important to learn from mistakes as well as from successes.

It is worth noting that there is no scientifically agreed ideal of how a scale should be constructed. Personally, I like scales with six alternatives. They have the advantage that they can be shortened to two, three or four graded scales when results are presented, and then the choice depends on what the results are, if one should point to an even distribution around a mean, to a central tendency, or to an eschewed variation around a mean. Utility is the key word, and the 6 graded scale makes it easy to emphasise the true nature of data in a presentation.

Sociological research often uses ranking scales where respondents report of something being satisfactory or not. Opinion polls of various kinds do the same, as well as studies exploring corporate cultures. It is also quite possible to assess project efficiency and effectiveness in terms of how satisfactory it is. But one should bear in mind that the question of whether a project has been satisfactorily completed or not can be debated, and defined, in relation to whether objectives have been reached, targets met, etc. The credibility of the statement whether something is satisfactory or not rests on the assumption that it is fairly clear why it is

satisfactory. There must be some norm which both the evaluator and his/her audiences can refer to. Institutional development impact often lacks such a common reference point. Institutional development is a very wide subject, and the impact goes beyond the question of whether project objectives were met. Indeed, institutional development impact is often a side-effect. The statement of whether that impact is satisfactory or not thus risks becoming very personal, and the validity of the observation can be very low.

However, one must bear in mind that a ranking scale that measures "satisfaction" can be more relevant when the assessment concerns human resource development or organisational development. It is primarily when the ranking proceeds to the more complex levels of institutional development that the issue becomes problematic, as of course when the evaluator does not explain which aspects of institutional development the ranking refers to.

Similar issues must be raised in connection to the World Banks ranking scale. As the possible ranking values are presented on a scale, the reader (and the evaluator) understands that substantial is a little bit lower than high. But if the ranking scale is not visible it is quite difficult to discern the difference. In a summary page, where the evaluators provide their assessment, the reader finds in some instances the rank "high" and yet in others "substantial". But as the categories are not explained, it appears quite confusing and you have no clue to which should be considered more successful.

In sum, it is of course quite attractive to consider a unified measure of institutional development impact. It would certainly facilitate aggregate reporting. But there is a need to refine the scales being used, to consider the categories the evaluators report in, and to ensure that the assessments are made according to common criteria. Consequently, there is also a need to develop indicators that can be used to support such scales. The most challenging issue is to ascertain that those who finally use the scales do so with some common frame of reference in respect of the objects of judgement. This requires a significant effort in developing scales, defining indicators, and making sure that evaluators use them similarly. The question is if the obstacles are insurmountable, or whether the costs of developing the methods can be justified by how the reports are being used.

Chapter 5. The methodology of evaluating institutional development

When evaluations rank institutional development impact, most of the work has probably been done. In this chapter I will look at what happens before the evaluators are ready to present their rank order, before they have summed up what their indicators point to. This chapter deals with the methods through which the evaluators collect data, process data and use data in the course of their analysis. The chapter starts with some basic arguments concerning the role of data in a structured and systematic inquiry. The next section presents the methods used in course of the 31 evaluations used in this study. The final section discusses the choice of methods to study institutional development.

Evaluation as a systematic inquiry.

Hitherto we have not actually defined evaluation. I have taken for granted that what the evaluation units of the banks do is actually “evaluation”. If they commission a study, then that study is an evaluation. If I am presented with a report which I am to analyse as part of this assignment, then that is an evaluation report.

It is now time to challenge that approach. An evaluation is commonly defined as “a systematic inquiry into the worth or merit of an object”. This definition was agreed under the auspices of the American Evaluation Association, which had appointed a joint committee of civil servants, researchers and consultants to agree on the subject and to define quality criteria. The quality criteria are published under the title “The Program Evaluation Standards” (Sanders 1994), and they are approved by the American National Standards Institute. The European Evaluation Society, as well as most of other professional organisations accept this definition.

Note that the definition says “a systematic inquiry”. What does that mean? It means that the evaluation should apply scientific method in its quest. It is not just any guesswork on project performance, nor any haphazard conversations that allow evaluators to reach conclusions. The challenge to use systematic methods means to structure data collection, to make informed choices on methodological issues, and to apply methods that yield objective, accurate, valid and verifiable answers to the evaluation questions.

If the methods employed do not live up to these standards, indeed, if no methods are employed, then we actually do not have an evaluation. I think it is plain to all that there are reports, which bear the title evaluation, but where there is no systematic method. There are also borderline cases, where the methods are loose; samples biased, response rates low, and interview formats misleading, so that even though the method may be systematic, the results cannot be accurate. Such reports do not merit being called evaluations; they do not enter into the category of the definition above.

This is a difficult distinction. I am only looking at the institutional development analysis in the evaluation reports. I have analysed what they have to say in this matter, and what methods they employ. However, institutional development is only one of the many dimensions that the evaluators have to consider. Often, it is but a secondary task. Most evaluations focus on subjects such as project design, implementation, the borrowers capacities, efficiency and effectiveness in relation to project objectives. Broader issues such as relevance, sustainability and institutional development usually receive less attention from the evaluation teams.

Hence, the evaluations may use systematic methods of inquiry to evaluate design, implementation, efficiency and effectiveness. They are indeed evaluations. But when they discuss institutional development, they do not always use any methods. Consequently they do not evaluate institutional development, even though the discussions take place under the title of an evaluation report.

Recent literature on evaluation methodology has introduced the distinction between evaluation and peer review. A peer review occurs when a team of seasoned experts observe events, and come up with an independent opinion on worth or merit. They do not use any structured approach to data collection; often they only do some random interviews, and then present their findings. The process is quite popular; it is inexpensive, quick, and often very practical. But it is not to be confused with evaluation. As we will see, the analysis of institutional development has more similarities with peer reviews than with evaluations.

Models and hypotheses

One of the most difficult challenges for an evaluation team is to structure the inquiry. The terms of reference are usually of help, but they are set to define the purpose and to facilitate commissioning the evaluation. Their primary task is not to structure the inquiry in the operational sense of that word.

The most important tool that can be used to structure an inquiry is a model of the phenomena to be studied. Models are representations of states, objects, and events. They are idealised in the sense that they are less complicated than reality and hence easier to use for evaluation purposes. The simplicity of models lies in the fact that only the relevant properties of reality are represented. I find it strange that the evaluations do not make more explicit use of models when they assess institutional development. In fact, among the 31 evaluations presented here, there are only a few (from the ADB and EIB) that use models to represent aspects of institutional development.

However, other evaluations have descriptive and analytical texts that indicate that the authors have a model in mind, which they use to structure their analysis. But the model is not made explicit. Perhaps this reflects a lack of space, or a lack of trust in the relevance of the model (any shortcomings in the choice of variables and in the nature of their connections rapidly becomes visible). Yet another possibility is that there is not much familiarity with reading and using models of institutional economics.

Following the development of models, another very useful approach to structure an inquiry is to develop a set of hypotheses concerning the development, which may be expected. This is a very useful way of structuring an inquiry; it helps the evaluator define a focus, to discuss the

relevance of that focus, and to assign priorities in the choice of methods. By formulating hypotheses early on in the evaluation, the evaluators will have an obvious instrument to assess whether their thinking is relevant and whether it will generate any new lessons. It leads to an economical use of the evaluation resources. The evaluators can then focus on whether to prove or disprove the hypotheses. And if the issue is trivial this soon becomes obvious.

The choice of methods.

When I now turn to the choice of methods, I will primarily deal with methods of data collection and analysis. In fact, several methodological choices are made even before that, as for example whether or not to develop models or to use hypotheses. There are other than strictly methodological choices, as for example whether to involve stakeholder groups in participatory evaluation. There are those who prefer to start with an overall choice between quantitative and qualitative methods. I find the latter to be of limited relevance, as the qualitative and quantitative nature of data tends to merge in the course of a practical evaluation.

In essence, an evaluator has a choice between four basic ways of collecting data; interviews, surveys, studying documentation, and observation. The aim of a discussion of methodology would be to ensure that evaluators make flexible, sophisticated and adaptive methodological choices. Table 3 in annex 3 presents the methods adopted by the evaluations in my sample. I have analysed the methods used in the 31 evaluations based on the information provided in chapters or sections devoted to methodology, or by inferring from the data presented in the report what the evaluators have done (not all write about their methods, something which I consider to be a major weakness).

Table 5.1 The choice of evaluation methods

Method in use	Number of occasions when the method was chosen (alone or in combination)	Percentage of total application of methods
Interviews	70	44%
Documentation	57	36%
Observation	19	12%
Surveys	13	8%

Interviews in this case means individual interviews as well as group interviews, and indeed participatory focus groups. At times the evaluation teams have interacted with village meetings, or student gatherings, or staff meetings in organisations. These are all covered under the caption interviews. The overwhelming majority are individual interviews. Interviews can be more or less structured, standardised or open ended. Few evaluations document how they have interviewed, or what questions they asked. The credibility of the

data is affected by the fact that so few evaluations actually refer to information provided through the interviews. When conclusions are reached, the reader seldom knows whether this emanates from the interviews, or are conclusions drawn irrespective of interview data.

Table 5.1 provides the rank order of the data collection methods used in the evaluations. I have simply added all methods used, and then counted each method's share of the total. This does not reflect the true extent of work with each method, but the information provided in the evaluations does not allow any more sophisticated analysis through weighting, or some other ranking system. It is quite clear that the most common methodological choice is interviews, closely followed by documentation. Evaluators go out and speak to government representatives, project managers, the management of other local organisations. The evaluators study project documents, previous evaluations. mid-term reviews, and the like.

Table 5.2 presents another view of the methodological choices. As table 3 in annex 3 shows, it is not uncommon that evaluations use several methods; indeed, there are only some few that limit themselves to one method. The evaluations that only use documentation are desk studies, and then there are some few that only use interviews and documentation. The latter appear to have the character of peer reviews, undertaken with little resources, under severe time constraints. Typically, they imply a few weeks' fieldwork, and a few weeks to write a report. It is questionable whether they should be called evaluations in view of the requirement for systematic inquiry mentioned above.

Table 5.2 The choice of methods on sample evaluations.

Method in use	Number of evaluations that make use of	Percentage of evaluations that make use of
Interviews	27	87%
Documentation	26	83%
Observation	13	42%
Surveys	9	29%

The more interesting question is if the evaluators make the right methodological choices in respect of their subjects. As table 3 in annex clearly shows, the most sophisticated methodological choices appear to be made in respect of organisational development. In this dimension, we find a wider variety of methods applied, and within any one evaluation there usually at least two sources of data collection.

Human resource development is not as much in focus, and often the evaluations use only one or a limited number of methods to assess this dimension. It is perhaps not surprising that a relatively large share of these assessments is made with the help of surveys. The evaluations from the ADB have, for example, made ambitious use of tracer studies, based on reaching a large number of participants in training programs, or in schools, to follow up the impact of training.

There is less variety to the choice of methods for the more complex aspects of institutional development; the networks, regulatory environment, values and norms. In particular the latter suffers from a rather simplistic choice of methods. The information in the evaluations is based

on interviews and reports, always without explicit reference to sources. It is unclear whether the conclusions regarding this dimension are based on hearsay, personal reflections, or plain guesswork.

The regulatory environment is mainly assessed through interviews and study of documentation. The evaluation teams have, through their meetings, received information on legal proposals, amendments to existing legislation, development of new institutions that regulate, for example, securities markets. In rare cases do the evaluations actually analyse sources. The documents referred to are mostly project documents, progress reports, and other documents directly relating to the projects. However, a more thorough assessment of the regulatory framework would presumably have to turn to the sources, to investigate bills, to trace budget allocations, and the like. I will return to the methods that could be adopted below.

The connections between organisations are studied with a more varied approach. Even though interview information tends to dominate here too, the evaluators make use of observation, surveys, and documentation. A common approach is to look at institutional development through the lens of a case study approach. The case study is a methodological choice in its own right, but it is also an organising framework within which the evaluators work with interviews, documentation, observation and surveys. As networks are emerging phenomena, it could be misleading to rely on only one method to explore them. Here, more than in respect of the other dimensions, the evaluators would have to apply several methods, for example through triangulation.

In sum, the evaluations use varied approaches to the subject of institutional development. However, when looking in detail at how interviews are performed, observations conducted, and documentation used as sources, several weaknesses are revealed. Interviews tend to be conducted among a small group of project managers, the evaluators seldom treat sample biases, and no formats are presented. It is not clear what information the interviews yield, nor what role it plays in the conclusions reached by the evaluators. I have not seen any example of structured use of observation, although it appears to be a significant source of information. Documentation is primarily used as a source of economic data, and seldom used for qualitative assessment. Organisational charts are sometimes used in organisational analysis. The choice of documents focuses on organisational development, but more rarely relates to other aspects of institutional development.

Developing the methods to evaluate institutional development

Institutional development impact can be a variety of changes in society. There is not any one single method to capture and describe them; the method needed depends on the type of change the evaluator is looking for. In this section, I will not discuss aspects of human resource development. Traditional evaluation theory has large numbers of handbooks on how to evaluate the effects of training, counterpart systems, scholarships, personnel development programmes, organisational learning and the like.

Organisational development is also a well-known field. It is more contested ground, which reflects that there are paradigmatic choices in organisation theory. Organisations can be evaluated as political systems, as cultural systems, as rational management systems, or as

flexible, adaptive and self-organising entities. There are many handbooks available, but the best is probably Clegg, Hardy and Nords (ed) *Handbook of Organisation Studies*, published by Sage, 1996. It covers everything.

Let me rather turn to the issues of norms and values, legal developments, and inter-organisational linkages. These areas are not so well known (in evaluation practice), and with some few exceptions, the evaluations have not been able to apply much systematic analysis to these aspects of institutional development.

Norms, values and attitudes are primarily explored with the help of surveys. The scale of the survey depends on the population, but often the sample can be much smaller than what evaluators usually consider necessary. One should remember that opinion polls could usually make quite accurate election forecasts with stratified random samples of some 1.000 respondents, out of a population of several million. If a project claims to have had a significant effect on norms and values in a larger group of people, then there is seldom any alternative to evaluation than to develop a survey format. Sage publishes a handbook entitled "Values in Evaluation and Social Research", which is very useful (House and Howe, 1999).

Information on norms and values can also be gathered from key informants, (which is actually what most evaluations in the sample did). But this is very unreliable. Dower (1999) discusses social change in Japan in the wake of the second world war, after a crushing defeat, nuclear bombing, and seven years of occupation under a military dictatorship. In 1952, Emperor Hirohito and General MacArthur independently commented on the extent of change in Japan. The former said that in essence Japan had not changed between 1945 and 1952; the latter said that change had been tremendous, from a feudal society to a full-fledged democracy. These two would probably be considered the ultimate authorities on that society!

Other information on values and attitudes can be fetched from documentation. Newspapers, journals as well as other media can be used for an informed discussion of values and how they change. The crux is that such a discussion requires local cognisance, something that a team of foreign evaluators seldom possess. Consequently, more local evaluators, or foreigners better familiar with the society and its institutions are required.

Turning to legislative development and how to assess that, one has to look for inspiration in jurisprudence and political science. These subjects in turn have much in common with historical method. In particular legal studies rely on careful examination and critical thinking about documentation (preparatory legislation, political debate and documented exchanges of opinions, draft parliamentary bills, differences between drafts and actual bills, and final legislation). It is also common to turn to biographies, key stakeholder memoirs, and investigative journalism. Some recent textbooks that can be useful are; *Research Strategies for Secondary Data* (Riedel, 2000), *Doing Research in Political Science* (Pennings et al, 1999), and *Doing Discourse Analysis* (Wood and Kroger, 2000) and *Media Analysis Techniques* (Berger, 1998).

As for organisational linkages, there are many interesting developments in network theory. Often, the basic information for a network analysis is contained in evaluations. It is gathered in the course of interviews and site visits, and through observation. It is only a small step to collect the data, systematise it, and apply it in the form of graphs and similar maps of social networks. Once the description is made, it is possible to proceed to analyse design variables and to discuss various aspects of quality of the network. By the same methods, one can

analyse how networks change. Even though the basic information is to be had from interviews and documentation, a full treatment of networks require a further elaboration. And if that further elaboration is to be credible, then the initial collection of data must be structured.

There are many handbooks available and also many examples in investigative journalism as well as in research. It is more difficult to point at really good network analysis in evaluation reports. Even though there are many useful handbooks, I would particularly like to point to Scott's (1998) *Social Network Analysis*. Nohria and Eccle's (1992) study "Networks and Organisations; Structure, Form and Action" is not exactly a handbook, but it illustrates well how networks can be studied.

In sum, this chapter has discussed the methods employed in the banks evaluations of institutional development. The chapter has suggested where further knowledge of various aspects of institutional development evaluation is to be found. Time and cost are crucial factors. Institutional development is only one of the concerns in the evaluations studies, and often it appears to be a secondary concern. Institutional development impact is difficult to analyse, and it cannot be done credibly without proper resources. That does not mean that evaluations should be transformed into research projects, on the contrary, evaluations must be practical and utilisation focused. Conclusions must arrive in time for decision-makers to consider the results. There is a balance to be tread, because the information presented by evaluations must be trustworthy. Many of the examples from the 31 evaluations in the sample show that it is possible to come up with far-reaching conclusions on institutional development impact.

Chapter 6. Examples of good practice.

It is now possible to outline a methodological approach to the evaluation of institutional development impact. Let us start with the five dimensions of institutional development; human resource development, organisational development, networks and linkages, legislation and rules, norms and values. Let us say that these constitute the vertical axis of a model.

We can then proceed to look at the different steps in the progress of an evaluation as a systematic inquiry. These are, from the beginning to the end; developing indicators, formulating hypotheses, and selecting methods. One could possibly add the design of ranking scales, but as these are now uniform among the MDBs, I will leave that subject to the side. I will let these steps constitute the horizontal axis of the model.

The model is illustrated in table form below. It can now be used to pinpoint the areas where we may look for good practice examples. Furthermore, it is of course widely different to evaluate a project as compared to evaluating a sector of assistance, support to a country over a longer period of time, or to evaluate a particular theme. The table below only covers good practice examples in relation to methodological steps for institutional impact evaluation in the context of project assistance.

When there is a more comprehensive view of experiences, it would be possible to construct a similar map in respect of each level of intervention, and to illustrate each cell in the tables with a good practice example! But I am not there yet, indeed, this assignment is meant to focus on project evaluation. Furthermore, not even with the limited subject of project evaluation is it possible to provide illustrations in respect of all the cells.

	Indicators	Hypotheses	Methods
Human resource development	Section 1	-	Section 6
Organisational development	Section 2	Section 5	Section 7
Networks	Section 3	Section 5	-
Legislation	-	-	-
Norms and values	Section 4	-	-

It is of course possible to consider many different types of indicators in respect of each dimension. The table and the text to follow is only meant as an introduction, and is based on the 31 evaluations, plus the additional six evaluations that were meant to lend substance to this analysis. In particular the column on methods should be used carefully. In fact, it would be proper to have four columns, one for each major category of methods (questionnaires, interviews, documentation and observation) plus perhaps a fifth column for triangulation, and a sixth for case studies....

When there is a cell that does not refer on to any particular section, this means that I have not found any example which merits being called good practice among these 37 evaluations. But 37 is still a small selection among all the many evaluations produced by the MDBs, so I dare not say that good practice in these respects does not exist.

Section 1. Good practice in respect of developing indicators for Human Resource Development.

The Impact Evaluation Study of the Technical and Vocational Education Projects in Malaysia, Pakistan, Papua New Guinea, and Sri Lanka provides a good example of how an evaluation team can develop indicators to assess Human Resource Development. The Human Resource Development component in this case refers both to the skills and the motivation of the staff of Vocational Training centres, as well as the skills of graduates.

The evaluation has utilised a flexible approach, whereby data was gathered via interviews, via questionnaires to teachers, industrial managers, and graduates from the schools. Furthermore, the evaluation has made extensive use of secondary data, various kinds of industrial statistics, manpower forecasts, etc.

Among the indicators, I would like to point at the following:

- job search duration, which is a statistic that comes out of the questionnaire to graduates, asking them how long it took before they got an employment. The answers could vary from "immediately" up to more than 12 months. There were four categories of time intervals. The indicator is likely to be quite reliable, it is meaningful to compare different schools, at it is meaningful to make comparisons over time. (see appendix 10, page 53, of the evaluation (annex 2 for reference))
- assessment of preparedness, which is also a statistic coming from questionnaires to teachers and graduates; that is, the two categories are asked the same question. The response signals whether the graduates are ready for future work, higher studies. To start their own business, and some other categories. The mean response also point at interesting differences between schools, projects, and groups of respondents. (see appendix 9, page 52, of the evaluation (annex 2 for reference))

Section 2. Good practice in respect of developing indicators for organisational development

It is always a matter of judgement what constitutes good organisational performance. The indicators vary widely depending on the nature of the organisation, and what aspect of the organisation that a project has sought to address. The World Bank study "Developing Towns and Cities: Lessons from Brazil and the Philippines" World Bank, OED, 1999, focuses on the capacity of municipalities to become financially competent. The indicators developed by the evaluation team focus on this aspect of municipal management, so they do not give a comprehensive picture. I think it represents good practice anyway, because they are focused, the indicators supplement each other, they are practical, and provide meaningful information.

- own revenues' share in total revenues, this is an indicator of financial autonomy, and it shows how the situation has changed over time, between different municipalities, between project participants and non-participants. (Table 3.2, p. 8)

- own revenue per capita, which illustrates the municipalities capacity to mobilise their own revenues, which is a good indicator for the same reasons as the one above. (Table 3.3, p.9)
- property tax per capita (Table 3.4, p:10)
- betterment charge per capita, which is an indicator of the municipalities capacity to recover costs for its services and its development programmes.
- surplus as share of revenues, which is an indicator of the health of municipal finances. Table 3.6, p. 12

These and a few more indicators together provide a good picture of the changes in financial management of the municipalities. The quantitative number of the indicators are easily understandable, and the process of data collection does not appear to be very complicated. (See annex two for reference to the Evaluation Report).

Section 3. Good practice in developing indicators to assess network development.

Networks are more complex systems than traditional organisational structures, hence indicators of network development would normally be even more complex than indicators of organisational development. Complexity would practically mean that you need more indicators to assess a network, that you need different kinds of indicators, and that you may need indicators of both structures and processes.

In the second chapter where definitions were discussed, the comparison between the market as an institution and as a network of firms, public agencies, insurance brokers, research bodies, etc, was mentioned. When searching for an example of good practice when using indicators to assess a market, I thought it would be relevant to refer to the EIB evaluation of capital market in Spain. The evaluation team developed a number of indicators that were used to assess the development of the capital market, and these were described over time, so as to trace whether the lending operations appeared to have had an impact. The indicators are replicated below. See the full report "The Impact of EIB Borrowing Operations on the Integration of New Capital Markets" for a closer description of how the indicators were measured and how they were used in the analytical part of the evaluation.

Otherwise, I have not found any examples where the evaluation teams develop indicators to assess networks. Still, it is quite common in sociological and political science research to analyse networks. The first step is usually to map the nodes and the links between these. Second, to measure the content of the links, which could be frequency of contacts, commercial volumes, exchanges of personnel, types of communication, and many others. An indicator of network development would usually be some form of quantitative description of the content in the channels, but it could also refer to the number of nodes in the network, the organisation of links, and possibly also the extent of changes (numerical indicators could put a value on change, innovative capacity, institutional reforms).

Indicators	Spain	Portugal	Greece
------------	-------	----------	--------

First foreign issue	1987 (August) EUROFIMA		1988 (November) EIB		1994 (February) EIB	
I. RELATIVE SIZE OF						
As % of nominal GDP:						
Gross bond issues	1988-90	6.2	1989-91	12.6	1994-96	15.3
	1996-98	12.4	1996-98	16.0	1997-98	16.5
Net bond issues	1988-90	1.3	1989-91	6.8	1994-96	7.4
	1996-98	5.4	1996-98	7.2	1997-98	12.0
Bond market capitalisation	1988	22.6	1988	8.0	1988	19.5
	1997	55.4	1998	44.3	1998	63.9
Bond market capitalisation as % of balance sheet volume of banking system	1988	16.5	1988	6.5	1989	37.6
	1997	32.0	1997	17.2	1996	88.2
II. FEATURES OF DEBT						
Fixed-rate government bonds: longest maturity in the market	1988	3 yrs	1988	2 yrs	1994	none
	1998	30 yrs	1998	15 yrs	1998	15 yrs
Share of floating rate debt in total marketable public debt					1994	91.3
					1998	77.6
Gross issues as a multiple of net issues	1988-90	4.8	1989-91	1.9	1994-96	2.1
	1996-98	2.3	1996-98	2.2	1997-98	1.4
Share of Treasury bills in total marketable public debt	1988	20.0			1995	37.6
	1998	25.1			1998	19.0
Share of government bonds in total bond issues (gross)	1988-90	53.9	1989-91	77.9	1994-96	97.2
	1996-98	76.0	1996-98	50.5	1997-98	99.6
III. SECONDARY MARKET INDICATORS						
Bond market turnover as % of market capitalisation (stock exchange data)	1988-90	n.a.	1988-90	n.a.	1988-90	0.8
	1997-98	70.4	1997-98	22.6	1996-98	0.3
IV. DEGREE OF INSTITUTIONALISATION						
As % of nominal GDP:						
All types of Institutional investors (financial assets), of which:	1990	16.0	1990	9.0	1990	6.5
	1997	56.1	1996	34.4	1996	28.5
Investment funds	1990	3.2	1990	4.3	1990	n.a.
	1997	33.7	1997	21.2	1996	13.1

Section 4. Good practice in developing indicators for the assessment of norms and values.

There are actually very few evaluations that really treat norms and values more than casually and in the passing. The example provided here is not really a sophisticated assessment of values and attitudes. But it is sharply focused, and in the context of the specific evaluation, relevant assessment of norms and values. It measures values that have changed, and this the assessment demonstrates the impact of the project.

The evaluation team assessed the values, norms and attitudes of stallholders and small entrepreneurs operating on municipal public markets. A questionnaire was developed to ask about their opinion of municipal services, their opinion on conditions at the market place, and their assessment of changes. The questions, the way it is used in the analysis, and the conclusions that can be based on the analysis, are demonstrated in the World Bank study "Developing Towns and Cities: Lessons from Brazil and the Philippines" World Bank, OED, 1999. (See also annex 2 for further references).

Section 5. Good practice in respect of using hypotheses to structure and inquiry.

Few evaluation teams use hypotheses, but the extensive quote below from the EIB evaluation of capital market development may show the value of structuring the inquiry into a complicated area by setting a number of hypotheses that may then be discussed, in turn. (see annex 2 for a reference to the evaluation):

”As a working hypothesis, it has been assumed that MDI borrowings can, in principle, help develop local bond or capital markets in a number of ways:

(i) The first MDB issue in a new bond market can have a positive icebreaker effect on investors in- and outside the country by strengthening general confidence in the stability and safety of the local financial system. Other foreign borrowers may follow suit, and the governments, investment banks and securities industry may feel encouraged to take further action to develop an efficient local market for foreign issues, thereby contributing to the international integration of the local capital market.

(ii) MDBs that return regularly to a new bond market can contribute to:

β broadening the capital demand side (i.e. increasing the availability of a range of AAA bonds) of that market. If MDBs are engaged in sizeable domestic currency lending programmes in the country concerned, they operate as domestic intermediaries, channelling domestic financial savings via the bond market into domestic use, for example to companies which are unable to tap the country’s bond market directly. This in itself contributes to the developmental role of the domestic bond market and is particularly important for financing investments that do not generate foreign exchange earnings;

β broadening the capital supply side of the bond market, by attracting additional funds into that market through an innovative design of debt instruments as well as through special marketing and selling efforts. In particular, by issuing bonds of longer maturity, they should help develop “the long end” of the market. When MDBs with important domestic lending programmes place their domestic-currency issues outside the country, they also supply foreign capital for development purposes while limiting the country’s foreign exchange risk exposure;

β developing, through swap operations, the local derivatives and swap markets. In some new capital markets that experience excessive capital inflows, MDB borrowing and swapping out of the local currency may also be beneficial in “sterilising” the liquidity effect of such inflows.

(iii) Having acquired extensive experience in a large number of countries, MDBs can transfer financial know-how to the local market place, mainly in areas such as the design of financial instruments, and selling and marketing techniques. Thus, they can contribute to the training and education of the local investment banking and securities industry in countries with new bond markets. In particular, MDBs that systematically swap issues in local currency into international currencies can train local debt managers in swap techniques.

(iv) MDBs can stimulate competition in the new issues market and other securities-related activities, thereby assisting the authorities in policies aimed at improving efficiency in the functioning of their country’s financial markets.

(v) MDBs can also influence the organisation, regulation and functioning of secondary bond markets. Since they are themselves interested in well-functioning secondary markets notably for their own benchmark bonds, it is worth their while helping the local securities industry and the authorities to eliminate deficiencies in this area.

(vi) Finally, through their contacts with the authorities and the local financial community, they can assist in developing a regulatory framework (admission rules, listing requirements, investment regulations for institutional investors, etc.), as well as

tax regulations and administrative procedures which are conducive to improving the functioning of the local capital markets generally.”

.....and so on. The development of the hypotheses covers another two pages of rather technical text. Even though evaluations seldom use hypotheses, it is becoming more and more common in all administrative and management research to do. A glance in any of the prominent scientific journals will provide examples of excellent practice, in particular any issue of *Administrative Sciences Quarterly*. Even though the ambition there is much higher than in most evaluation reports, their examples of how to hypotheses could be useful.

Section 6. Good practice in the choice of methods to assess human resource development

The five evaluations that contained a full treatment of human resource development all made use of some kind of questionnaire. Three of the evaluations had the questionnaires annexed to the report, so it is possible to look at the design of questions, response alternatives etc. The quality of the questionnaires vary. Some are quite long, which may reduce the reliability of the responses.

Some questionnaires ask a great number of questions that are not immediately relevant to the evaluation, and that are not used in the reports (as for example the World Bank evaluation “Developing Towns and Cities”. The fact that those who respond to what they think about municipal services also have to furnish information on employees, working hours, turnover, transport distances etc may affect the truthfulness of their responses. I would think that it also affects response rates. However, as none of the evaluations report on response rates, and as they do not treat sampling, it is not obvious that they really are examples of good practice. The actual research instrument could still be of interest.

Section 7. Good practice in the choice of methods to assess organisational development

The EBRD evaluation of the Vneshtorgbank – Standby facility is an interesting example of how an evaluation is mainly based on documentation as a source of information and evidence of impact. It is not clear what role any other sources of information has had, though I would guess that the role of interviews has primarily been to discuss the documentary evidence with bank management, and possibly other stakeholders. There are few evaluations that focus that narrowly on one method, but in this case, the information gathered becomes highly relevant, and it seems to convey a picture of the most interesting outcome and impact of the loan on the Vneshtorgbank. I would think that it has been a highly cost-effective evaluation process. It may not be easily replicated elsewhere, but it nevertheless an example of good practice. (See annex 2 for full reference to the evaluation report).

Regretably, there is not really any good practices to bring forward in respect of methodological choices concerning the remaining three levels of institutional development. The evaluations do at times cover institutional development at these levels, actually quite often as far as networks are concerned. But their methods are obscure. The analysis may be qualitatively interesting, but the data is not accounted for, nor are the methods of data collection described. I have noted above that few evaluations contain methodological analysis

of what they did. The sections on methods and methodology are very brief. Hence, I do not really know what they have done, and I cannot point to any good practice. The fact that the evaluators have written something about, for example, impact on legislative development, does not mean that they describe what methods they have used.

This lack of concern for methods is a major obstacle to learning. Perhaps there are many examples of best practice, but if the evaluators do not describe what they have done, how can we know? Perhaps another recommendation coming from this assessment should be to make sure that evaluations contain a clear description of how they have worked, or perhaps those commissioning evaluations should have the sense to ask for such descriptions when they think the process of evaluation would have some general interest.

Chapter 7. Conclusions and recommendations.

The basis for this assignment on the assessment of institutional development has been a review of 31 evaluations. These evaluations were very different, and they were of an uneven quality. It may be risky and premature to draw extensive conclusions on the information that they provide. However, I take the risk that these 31 evaluations do not represent the best of policy and practice among the multilateral development banks, and I proceed to make recommendations as if they really are typical of current practice. I leave caution to the reader.

Conceptual clarity

The first observation is that institutional development as a subject has evolved during the 1990s, both in the theoretical deliberations in development economics, and practically in policy studies of bilateral and multilateral agencies. The trend has been to move from a concept of institutional development which was confined to the public sector, and which concerned a single aspect of organisational performance, to a concept which has broader ramifications, which includes many specific kinds of development, but which also synthesise these at higher systems levels.

This development does not appear to have been fully recognised in the policy and practice of the banks. In practice, the banks have mandates that are fully in line with the common definitions of institutional development. In theory, the banks have not yet put forward a common understanding of institutional development, that may guide their evaluation work. The differences are probably not large, but it still remains to explicitly nail down a joint definition of what they consider to be contained in the word institutional development.

Obviously, the first recommendation is to accomplish that, to agree on such a joint definition. It is necessary to steer the work in this field with conceptual clarity. It is thus suggested that the participants in ECG agree on the term and issue a joint statement of how institutional development is to be defined in the context of evaluation work.

I have argued that there is one definition which is better suited to practical work, and which is more rigid and better supported by research than other competing definitions (see chapter 2). This is naturally the definition I would suggest be used, but there may be reasons for other definitions that I am not aware of, or where I am simply mistaken. Most important, is that the banks do arrive at their own explicit definition that can be put to use practically at the earliest possible date. Without conceptual clarity, there is not much to be done in other fields. The following are concrete, practical steps to be taken:

- create consensus among ECG members on the definition of institutional development,
- seek endorsement of this definition in the development banks (from research and policy divisions, top management and board),
- make the definitions public, together with a “statement of intent” that the definition will serve as the base for the banks’ assessment of institutional development impact;
 - (a) in the form of a joint ECG pamphlet,

- (b) in the form of a webpage linked to the ECG network, and to the banks' information sites,
- (c) as attached instructions for evaluation teams, as part of their briefing kit.

Overall strategy

There is wide diversity in respect of how the banks approach institutional development; some take a more formative approach in evaluations, others a more summative approach. Evaluation activities must be useful; indeed, the foremost quality of an evaluation is "utility". The review of 31 evaluations seems to indicate that those evaluations that have taken a formative approach to institutional development may be more useful than others. Hence, it is recommended that institutional development be primarily treated formatively. Whatever evaluations say about institutional development must be meaningful for project managers, partners in developing countries/transition economies, and desk officers. If the information cannot be put to use by practically oriented people in these categories, then it may not be worthwhile collecting and disseminating.

The recommendation is to carefully assess the utility of the analysis of institutional development, and to emphasise the strategic orientations that best serve the decision-makers in and around the projects and programmes supported by the banks. The evaluations must target their needs for decision support, in respect of institutional development, as well as in other respects. In practice this would mean that:

- (1) comprehensive analysis of institutional development impact becomes a mandatory subject in country evaluations, assessment of country- or region wide experience, both in what may be a mid-term review of a country program, or an evaluation when a period of cooperation comes to an end. The point is that a geographically defined evaluation process should always contain a comprehensive assessment of institutional development impact;
- (2) comprehensive analysis of institutional development impact should be a mandatory part of any sector- or thematically oriented evaluation exercise. This would include mid-term evaluations, evaluations at the end of a project or programme, or any other framework of a thematic or sector-wide evaluation.
- (3) mid-term reviews/evaluations should always analyse institutional development impact. Even though a full evaluation of impact is not possible, the evaluation teams must have a mandate to scan the environment for institutional development impact, and in particular to analyse how the institutional environment affects the project/programme.

Developing indicators

The review has shown that there are evaluations that make sophisticated use of indicators when addressing institutional development issues. But there are also many that make no use of indicators at all, or whose indicators are more like "areas of study", which in themselves presuppose a number of indicators (which may or may not be developed in the evaluation). The review suggests that the concept of indicators, and the practical and operational use of indicators in the course of an inquiry, are not properly understood by all evaluators.

There is a rapidly growing literature on indicators in public administration, but I have not seen any application of this work for the evaluation of development projects. However, it might be necessary to ensure that evaluation teams possess the skills to work with indicators, to sensitise them so they do not forget to develop indicators, and to guide them by referring to manuals, best practice examples, etc. In practice, this could take the form of:

- (1) developing a handbook on the use of indicators for evaluators. Such a handbook could consist of:
 - (a) a general introduction to why and how evaluation teams use indicators,
 - (b) sections that present suitable indicators in the five dimensions of institutional development discussed above, but with the emphasis on indicators at the systems level (network, legislative development, values and attitudes),
 - (c) illustrations of best practice, in each of the five dimensions;
- (2) developing a website, directed at stakeholders that are concerned with the design of projects and programmes, as well as with evaluation, that constitute a discussion forum on institutional development, a point of reference for research, and information on the application of indicators in evaluation work – with references to literature, best practice, bank experience, consultants, and specific evaluation reports;
- (3) developing an interactive handbook/website to guide evaluation teams on the use of indicators, with content such as under 1 and 2 above;
- (4) and finally, though this is a recommendation which presupposes more knowledge than may exist currently; namely that the tender announcements include as a specific criterion that the tendering consultants include a presentation of the indicators they intend to use in the evaluation – and that the quality of these are specifically commented in the evaluation of tenders.

Measurement and ranking

The banks apply several different scales to measure institutional development impact. Still, the intent behind the scales is the same, as is the philosophy of what to measure. The differences simply reflect that the development of scales have taken place in isolation from the other banks.

It would greatly facilitate the work of evaluation teams, and evaluation units, and staff members who are in contact with other agencies, not to mention projects that are co-financed, if the same ranking scales were used by the banks. As there is a separate consultancy assignment going on, in parallel to this, it is likely that a joint ranking system will emerge. Needless to say, the comments in chapter 4 in respect of the choice of words, the basis for ranks, and inclusion of possibly negative outcomes, are put forward as an input to this work.

Methods to assess institutional development

It is quite difficult to assess the overall quality of the evaluations of institutional development impact. The task is made more complex as the evaluations usually have a focus (which they should have). But as a consequence, they cover one or two aspects of institutional

development quite well. Other aspects are covered insufficiently, or not at all. As most evaluations have to work under considerable time constraint, and with limited personnel resources, it is quite understandable that they do not always have a comprehensive coverage of institutional development.

There is a difference between evaluation theory, which is expected to be a systematic application of social science methods in an inquiry, and the actual practice of several evaluations. Due to time constraints, they often make do with a limited application of methods (read unstructured interviews), and with meagre sources of information (read project managers and other highly biased sources of information). The topic of institutional development is too difficult to handle without scientific rigour, or without systematic approach.

Practically speaking, it may be useful to distinguish between practices of evaluation and of peer review, where the latter is a less pretentious way of arriving at conclusions. Many times, it describes more accurately what actually happens. Sometimes, the etiquette of "evaluation" is merely a false label. It is better to give the study a name that describes what kind of a product it really is.

But it is also possible to really undertake evaluations, in which case there is a need to plan the inquiries well in advance, to search for locally qualified teams, to emphasise methodological competence. It would also be prudent to focus the analysis of institutional development. It is too wide a topic to be comprehensively covered in most evaluations. It is better to cover one aspect well, than many less accurately. In practice, This would mean that in each evaluation the following steps were taken:

- (1) start the process with a review of the dimensions of institutional development, and carefully assess what kinds of data that is needed to document impact;
- (2) if any of the dimensions are not relevant for the particular project/programme being evaluated, this should be stated explicitly and put in an annex to the report,
- (3) exclude interviews with project/programme staff, and their counterparts in the national administration as sources of data (for institutional development impact), and design the evaluation process with this in mind.
- (4) the three steps above must be completed before actual work starts, and before the team starts working at the project site(s),
- (5) in practice, these requirements mean that evaluation teams must include team members who are familiar with extensive data collection in the country concerned, who have access to institutional environments where secondary data exists (in the form of statistics, qualitative data and so on), and who speak the language(s) of the host countries to the project properly,
- (6) in practice, it also means that evaluations must be planned to allow sufficient time for data collection. This means that the time from beginning to end of the evaluation process often must be longer than is presently common – but it must not necessarily be more expensive.

Increased coverage, variety and sophistication in reporting on institutional development.

The review of the 31 evaluations suggests that there is a need for an increased coverage, variety and sophistication in reporting on institutional development. Analysis of institutional

development tends to focus on organisational development, which is but one entry point to the subject. The methods that are used for the more complex subjects of networks, legal development, and values and attitudes, are still quite rough. There is a need to incorporate knowledge from specific fields of study that are not commonly used in the evaluation of development projects, notably legal studies, sociology, and network analysis. The methods are readily available, the problem is to recruit evaluation teams who are familiar with such methods, or to encourage evaluations teams to develop an expertise in such methods, and to apply them within their assignments. The practical steps to be taken involve the following:

- (1) ascertain that all evaluations are put on tender and that invitations to tender are sent to firms, individuals and institutions with documented competence also in institutional development analysis,
- (2) in the calls for tender, specify that the tenders will be evaluated (among other things) according to the methodological choices made in respect of the analysis of institutional development impact,
- (3) specify that innovation, collaboration with local consultants, use of hypothesis etc. will be rewarded in the evaluation of tenders,
- (4) provide technical support to the staff who evaluate tenders, so that their ability to perceive a good methodological solution to the problem of institutional development analysis can be guaranteed,
- (5) publicise the tender evaluations. The steps 1 to 5 together represent a forceful combination of sticks and carrots to increase the quality of tenders and the subsequent analysis of institutional development impact.
- (6) The evaluation units should monitor the institutional development analysis of evaluations. This could use the format presented in this report, so that the coverage of evaluations, and the methods they used, are reviewed regularly. For example, if there are few evaluations that cover the development of networks, then priority should be given to these in the coming months. If there are no evaluations that use methods of legal studies, then give priority to develop these in the coming months. With the help of a strategic review of the evaluations, the workplans should be designed to increase the variety in methods and coverage as an end in itself.
- (7) There is a need to add competence from the various fields of institutional development to the conduct of evaluation; this will among other things require:
 - (a) cooperation with research institutions in the disciplines concerned
 - (b) seeking out consultants within these fields
 - (c) encouraging network building among traditional development consultants and those mentioned above
 - (d) rewarding tender proposals which include competence from such fields (legal studies, organisational development, sociologists, network analysis, etc)
 - (e) using information technology to make methods and tools available to evaluation teams (such as the suggestions above on handbooks, webpages, etc).

References

- Amsden, A., 1989, *Asia's Next Giant: South Korea and Late Industrialisation*. New York; Oxford University Press.
- Andersson, G. and Winai, P., 1996, Diagnosis of Organisations in development cooperation. Unpublished memorandum. Stockholm, Sipu International.
- Baser, H. and Morgan, P., 1996, Review of Technical Assistance Loans in the World Bank. Washington, D.C.: Operations Policy Department, World Bank.
- Beck, U., 1995, Att uppfinna det politiska. Bidrag till en teori om reflexiv modernisering. Göteborg, Daidalus.
- Berg, E., 1994, *Rethinking Technical Cooperation: Reforms for Capacity Building in Africa*. New York: UNDP and Development Alternatives Inc.
- Berger, 1998, *Media Analysis Techniques*. London, Sage
- Berlage, Lodewijk and Olav Stokke, eds., 1992, *Evaluating Development Assistance: Approaches and Methods*, London: Frank Cass.
- Bossuyt, J., Laporte, G. and van Hoek, F., 1992, *New Avenues for Technical Cooperation in Africa*. Maastricht: European Centre for Development Policy Management.
- Braudel (1984) *Civilization and Capitalism*. Vol I – III. London; Collins
- Brinkerhoff, D., 1994, Institutional Development in World Bank Projects: Analytical approaches and intervention designs. *Public Administration and Development*, Vol 14.
- Brinkerhoff, D., 1990, Technical Cooperation and Training in Development Management in the 1990s: Trends, Implications and Recommendations. *Canadian Journal of Development Studies*, XI (I): 139-149.
- Buyck, B., 1991, *The Bank's Use of Technical Assistance for Institutional Development*. Washington, D.C., the World Bank.
- Cameron, K.S. and Whetten, D.A., 1983, "A model for teaching management skills," *Exchange: The Organizational Behavior Teaching Journal*, 8, 21-27.
- Cassen, R. and Associates, 1986, *Does Aid Work?* Suffolk: Clarendon Press.
- Castells, M., 1996, *The Information Age: Economy, Society and Culture*. Vol 1 – 3. Oxford, Blackwell
- Castoriadis, C., 1987, *The Imaginary Institution of Society*. Cambridge, Polity Press
- Clegg, Hardy and Nords (ed), 1996, *Handbook of Organisation Studies*. London; Sage.
- Cooper, L., 1984, *The Twinning of Institutions: Its Use as a Technical Assistance Delivery System*, World Bank Technical Paper #23. Washington, D.C.: World Bank,
- Cooperrider, D. And Dutton, J., 1999, *Organizational Dimensions of Global Change*. London, Sage.

- Davies, N. (1998) *Europe: A History*. London; Pimlico
- Dower, D (1999) "Embracing Defeat: Japan in the Wake of World War II", New York, New Press
- European Commission, 1996, *Etude sur le developpement institutionnel. Rapport de synthese*. Brussel, European Commission.
- Evans, P., 1995, *Embedded Autonomy. States and Industrial Transformation*. Princeton University Press, Princeton.
- Forss, K., Carlsen, J., Frøyland, E., Sitari, T., and Vilby, K., Copenhagen, 1988, Evaluation of the effectiveness of technical assistance personnel financed by the Nordic countries. Copenhagen, DANIDA
- Forss, K., Birungi, H. And Saassa, O., 2000, Sector wide approaches: from principles to practice. A study commissioned by the Department for Foreign Affairs, Dublin, Ireland.
- Gerschenkron , A., 1962) *Economic Backwardness in a Historical Perspective*. Cambridge; Belknap House and Howe, 1999, *Values in Evaluation and Social Research*. London; Sage
- Hirschman , A.O., 1958, *The Strategy of Economic Development*. New Haven; Yale University Press
- Hofstede, G., 1980, *Culture's Consequences. International Differences in Work Related Values*. London: Sage.
- Institute for International Cooperation, 1994, *The Need for a Heightened Project Management Emphasis on Institutional Development*. Tokyo: JICA.
- Israel, A., 1987, *Institutional Development: Incentives to performance*. Baltimore: Johns Hopkins University Press.
- James, P., 1996, *Nation Formation: Towards a Theory of Abstract Community*. London, Sage.
- Kruse et al, 1998, Development through Institutions? Oslo; UD. Evaluation Report 5. 98.
- Materu, J.S., Civil Society Capacity Building: Challenges and Hurdles. *Paper presented at the DAC TC Network. OECD*, Paris, December 1996.
- Mathieu, N., 1996, *Industrial Restructuring. World Bank Experience, Future Challenges*. Washington, World Bank.
- Moore, M., Stewart, S. and Hudock, A., 1995, *Institution Building as a development assistance method: A review of literature and ideas*. Sida Evaluation report 1995/1. Stockholm: Sida.
- Moore, M., Gould, J., Joshi, A., and Oksanen, R., 1996, *Ownership in the Finnish Aid Programme*. Helsinki: Ministry for Foreign Affairs.
- Morgan, P., and Qualman, A., 1996, Institutional and Capacity Development, Results-based Management and Organisational Performance. *Paper presented at the DAC TC Network. OECD*, Paris, December 1996.
- Mosley, P. , 1987, *Overseas Aid: Its Defence and Reform*. Brighton: Wheatsheaf.

- Mosley, P. , 1987, *Poverty Focused Aid: The Lessons of Experience*. Action Aid Report.
- Muscat, R., 1986, Evaluating Technical Cooperation: A Review of the Literature. *Development Policy Review*. 4: 69-89.
- Nohria, N., and Eccle, R.,1992, *Networks and Organisations; Structure, Form and Action*. Boston, Harvard Business School Press
- North, D.C., and Thomas, R.P., 1973, *The Rise of the Western World. A New Economic History*. Cambridge, Cambridge University Press.
- OECD, 1991, *Principles for New Orientations in Technical Cooperation*. Paris.
- OECD, 1995, *Participatory Development and Good Governance*. Development Cooperation Guidelines Series. Paris.
- Pennings et al, 1999, *Doing Research in Political Science*. London; Sage
- Pinto, R. and Mrope, A., 1995, *Assessing Sector Institutions: Lessons of Experience from Zambia's Educational Sector*, World Bank Discussion Paper # 297. Washington, D.C.: World Bank,
- Pio, A., 1996, Integrating Institutional and Capacity Development into Sector Investment Programs: Some Lessons from the Asian Development Banks Experience. *Paper presented at the DAC TC Network*. OECD, Paris, December 1996.
- Powell, W. and DiMaggio, P., 1991, *The New Institutionalism in Organisational Analysis*. Chicago: University of Chicago Press.
- Putnam, R.D., 1989, *Bowling Alone; The Collapse and Revival of American Community*. New York; Simon & Schuster.
- Putnam, R.D., 1999, *Making Democracy Work*. New York; Simon & Schuster.
- Riedel, 2000, *Strategies for Secondary Data*. London, Sage
- Rebien, C. , 1996, *Evaluating Development Assistance in Theory and Practice*. Aldershot: Avebury.
- Sahara, T. 1991, Donor Delivery Style, Learning and Institutional Development. *Management Development in the Third World*. Vol. 10 (6) 61-81.
- Samset, K., Forss, K., and Hauglin, O., 1993, *International Learning from Evaluations and Reviews*. *Evaluation Report 1.93*, Oslo: Ministry of Foreign Affairs.
- Scott, J.,1998, *Social Network Analysis*. London; Sage
- (Selznik, 1957). *Leadership and Administration*. Evanstone; Row and Peterson.
- Steedman, D., 1995, *The Importance of Governance in assessing the Potential Benefits and Risks of Technical Assistance for Institutional Development*, PSM Network Discussion Paper. Washington, D.C.; World Bank.
- Takavarasha, T., 1996, *Management and Institutional Performance Programme*. *Paper presented at the DAC TC Network*. OECD, Paris, December 1996.

Uphoff, N., 1992, *Learning from Gal Oya: Possibilities for Participatory Development and Post-Newtonian Social Science*. Ithaca: Cornell University Press.

Wood and Kroger, 2000, *Doing Discourse Analysis*. London; Sage

World Bank, 1991, *Managing Technical Assistance in the 1990s. Report of the Technical Assistance Review Task Force*. Washington, D.C.: World Bank, Operations Evaluation Department.

World Bank, 1993, *Handbook of Technical Assistance*. Washington, D.C.: World Bank, Operations Policy Department.

World Bank, 1994, *Improving the Management of Technical Assistance for Institutional Development*, *Organisational Design Staff*, August, 1994, Washington, D.C.: World Bank.

World Bank, 1996d, *Performance Indicators for Technical Assistance Operations*. A First Edition Note. Operations Policy Department. Washington, D.C.: World Bank.

Annex 1. Terms of reference

***European Bank for Reconstruction and Development (EBRD)* Draft Terms of Reference and Budget (Approx. Euro 40,000)**

Institutional development and transition in a comparative MDB-evaluation perspective.

1. BACKGROUND

Joint work to improve the accountability to owners and feed-back of lessons from performance evaluation of projects and programmes is the concern of the Evaluation Cooperation Group (ECG). Members are the evaluation units of the multilateral development banks (MDBs). EBRD, is represented by its Project Evaluation Department (PED) together with the evaluation units of the World Bank, IFC, EIB and the three regional development Banks (IADB, AfDB and AsDB).

The following proposal concerns a study to focus on institutional aspects of economic transition and development. This area is now perceived as a key aspect of evaluation in *ex-ante* and *ex-post* perspectives alike. EBRD's evaluation Department, PED, chairs ECG in 2000. EBRD would, in line with ECG tradition, seek TC-funds for any joint studies from its own donor network.

2. OBJECTIVES OF THE STUDY

A. Study on evaluation of institutional development impact (IDI)

- To present an overview of how the institutional transition and development dimension has been reflected in recent studies and evaluation guidelines by member MDB's, taking into account their mandates, strategies and operational policies.
- To analyse how each MDB currently applies specific institutional development criteria at performance evaluation of completed operations and in sector- and country orientated impact studies.
- To present a comparative analysis and to reach conclusions.
- To present recommendations for improved methodology and reflection of the institutional dimension in evaluation criteria and practice by ECG-members.

B. Updated exchange on recommendations concerning best practice and processes in evaluation.

- The consultant will review and extract from each member's current guide-lines and manuals evolving best practice in evaluation of public sector operations.

- Recommendations will be presented on best practice and opportunities for harmonized criteria and processes, with due regard to MDB's partly differing mandates.

The A-part of the Study is of particular interest for EBRD: The aftermath of the Russian financial crisis in August 1998 highlighted that fledgling institutions posed significant risk to the ongoing transition to functioning market economies. This concerns in particular countries like Russia, CIS and Central Asia, where the Bank invests in projects, apart from those in the Baltics and in Central and Eastern Europe)

3. SCOPE OF WORK

A. Institutional Development Aspects of Evaluation.

The intended Study requires assistance from an independent consultant (the "Consultant") in providing:

- Overview, analysis and recommendations pertaining to the institutional transition and development dimension in evaluation work by MDB's in line with the above objectives;
- Documentation and analysis of current evaluation practice of the MDBs in addressing institutional aspects in performance evaluation. This includes differences in the approach to public and private sector operations and also how the institutional dimension has been applied to developing countries as well as to transition economies in Central Europe and the CIS.
- Presentation of recommendations. These should take into account common ground in line with ECG's aims and the aforesaid ECG report of 1999 on harmonised evaluation criteria. The recommendations should also concern which tailored approaches would be required in view of the varying mandates of the MDBs.

B. Updated exchange on best practice

- Desk study of current materials in the form of guide-lines, manuals and other relevant documentation posted on the ECG Intranet and of prior ECG reports of relevance (including the so called Eriksson report of 1996).
- Analysis and recommendations.

4. IMPLEMENTATION ARRANGEMENTS

4.1 Phasing

The character of the Study and its audience requires the consultant to work in two stages: *The first phase*, comprising about 1/4 of the allocated man-days, will result in a comprehensive Inception Report.

Consultations on this report with ECG's Working Group on Evaluation Criteria and Ratings in Public Sector Evaluation (WGEC) will aim at a narrowed, more detailed scope for *the second phase*, up to and including delivery of the Final Report.

4.2 Work Methods.

Desk Study: The assignment will be performed principally as a desk study, using documents and reports as posted by the ECG members on its joint "Intranet". These will include recent evaluation guide-lines, manuals and reports of relevance for the assignment. Member MDB's will also provide lists of evaluation reports from 1999 and onwards and recommend 3-5 reports from each which illustrate best practice in the application of institutional development impact (IDI) criteria. They will also provide references to studies that may not have been posted on the Intranet.

Telephone Interviews, E-mail correspondence: Each MDB will name a contact person to respond to questions, offer explanations and references etc.

Analysis, synthesis, reporting: The final scope will be agreed upon the Inception Report while reporting will be as prescribed below.

5. DELIVERABLES

The Consultant is required to submit the Inception Report by end August 2000 to ECG members. A presentation by the Consultant to WGEC at its planned meeting on 7th September in Washington D.C. will include discussion and agreement on modified terms of reference for the Final Report with suitably narrowed scope and precise dates for Draft and Final reports.

The present outline of consultant's schedule is as follows:

Startup of assignment	June, 2000
Inception Report distribution presentation	Late August , 2000, 7 th September, 2000
Draft Final Report Distribution Consultations with ECG	October , 2000 Discussion Nov. 2000- Jan., 2001
Final Report	Febr. 2000

6. Consultant Qualifications

For Part A: A senior economist with an institutional orientation, long experience from independent performance evaluation of projects and programmes as well as a profound insight into the remits and work of MDB's in ECG.

For Part B: An economist or person with other relevant academic qualifications for evaluation orientated work or with a good insight into the work of IFI's. This person should work under guidance from the senior economist.

Annex 2. List of evaluations used in the analysis

Impact Evaluation Study of the Technical and Vocational Education Projects in Malaysia, Pakistan, Papua New Guinea and Sri Lanka. AsDB, 1999

Technical Assistance Performance Audit Report on Advisory and Operational Technical Assistance Grants To The Energy Sector in Mongolia. AsDB, 1999

Effectiveness and Impact of Training in Education Projects in Indonesia. AsDB, 1998

Assessment of the Effectiveness of Bank Technical Assistance for Capacity Building in Indonesia. AsDB 1997

Special Evaluation Study of Effectiveness and Impact of AsDB Assistance to the Reform of Public Expenditure Management in Bhutan, India, Kiribati, and Lao People's Democratic Republic. AsDB, 1999

Performance Audit Report. Burkina Faso Agricultural Research Project and Agricultural Services Project. World Bank, 1999

Performance Audit Report. Brazil Municipal Development Project in the State of Parana and Municipal Development Project in the State of Rio Grande do Sul. World Bank, 1998

Performance Audit Report. Ghana Transport Rehabilitation Project I and II, National Feeder Roads Rehabilitation and Maintenance Project. World Bank, 1999

Performance Audit Report. India Integrated Watershed Development (Hills) Project and Integrated Watershed Development (Plains) Project. World Bank, 1998

Impact Evaluation Report. Building Institutions and Financing Local Development. Lessons from Brazil and the Philippines. World Bank, 1998

Philippines: From Crisis to Opportunity. Country Assistance Review. World Bank, 1999

Summary Evaluation Report IDB Activities in Support of Public Sector Reform. IADB, 1995
Summary Report Evaluation of Policy-Based Loan Portfolio – Phase III. IADB, 1999

Social Investment Funds in Latin America. Past Performance and Future Role. IADB, 1997

Project Performance Review. Mexico: Evaluation of the Mexican Health Services Upgrading Programme. IADB, 1997

Project Performance Review. Guatemala: Decentralisation in the Health Sector. IADB, 1998
A Preliminary Overview of the IDB's Experience with Sector Lending: First Phase. IADB, 1996

XMR Assessment of Vneshtorgbank – Standby Facility. Russian Federation. EBRD, 1999

Western Siberia Oil and Gas Rehabilitation Project, Russia. EBRD, 1997

Telecommunication Legislative and Regulatory Development, Lithuania. EBRD, 2000

Mining Privatisation Kazakhstan. EBRD, 1999

Small Municipalities Environment Estonian Water Company. Estonia. EBRD, 1999

M1 – M15 Toll Motorway Project, Hungary. EBRD, 2000

The Development Impact of EIB borrowing Operations, Spain. EIB

EIB contribution to regional development in Portugal and Italy. EIB

The Impact of EIB operations on the Integration of Capital Markets. EIB

An Evaluation Study of 17 Water Projects Located Around the Mediterranean. EIB

Additional Evaluation reports that were added for good practice examples:

India: The Dairy Revolution. World Bank, 1998

Investing in Health. World Bank, 1999

Financial Sector Reform, World Bank, 1998

The World Bank Forest Strategy, World Bank, 2000

Albania, Country Assistance Evaluation, World Bank, 2001

Developing Towns and Cities. World Bank, 1999

Agricultural Extension. The Kenya Experience. World Bank, 2000

Annex 3. Analytical tables used in the report

Table 1. Entry points to institutional development analysis

Title of report	Human resource development	Organisational development	Networks	The regulatory environment	Values, norms and attitudes
Energy sector Mongolia	Full	Full	No	Some	No
Technical and vocational education	Full	Full	Some	Some	Some
Reform public expenditure management	Some	Some	Some	No	Some
Capacity building Indonesia	Full	No	Some	No	Some
Training in education projects	Full	Full	No	No	No
Toll Motorway project	No	Full	No	Some	No
Estonian Water Company	Some	Some	Some	Some	No
Mining privatisation	Some	Some	No	Full	No
Telecommunication legislation	Full	Some	No	Full	No
Western Siberia Oil and Gas	No	Full	Some	Some	No
Vneshtorgbank	Some	Full	Some	No	No
IDB sector Lending	No	Some	No	No	No
Guatemala Health sector	No	Some	Some	Some	No
Reform of tax administration	No	Full	Some	No	No
Mexican Health services	Some	Some	Some	No	No
Social Investment Funds	Some	Full	Full	Some	No
IDB Policy-Based Loan Portfolio	No	Some	Some	Some	No
IDB Public Sector Reform	No	Full	Full	Some	No
EIB borrowing in Spain	Some	Some	Full	Full	Some
EIB contribution to regional development in Portugal and Italy	No	Full	Some	No	No

Integration of capital markets	No	No	Full	Some	Some
Evaluation of Water projects	No	Full	Some	No	No
Watershed Development India	No	Full	Full	No	No
Transport Rehabilitation Ghana	Some	Some	Full	No	No
Municipal development Brazil	No	Some	Full	No	No
Local Development Brazil, Philippines	No	Full	Some	Some	No
Agricultural research Burkina Faso	No	Some	Full	No	No
Philippines Country Assistance Review	No	No	Full	Some	No
Water Supply Thematic study	No	Some	Some	Some	Some
Education sector Malawi	No	Full	Some	No	No
Agricultural Bank Nigeria	No	Full	No	No	No

Table 2. Formative or summative approaches

Title of report	Institutional development used to discuss and explain project implementation and performance	Institutional development used as a measure of succes, as an impact created by the project
Energy sector Mongolia	Yes	Yes
Technical and vocational education	Yes	No
Reform public expenditure management	Yes	Yes
Capacity building Indonesia	Yes	No
Training in education projects	Yes	Yes
Toll Motorway project	Yes	Yes
Estonian Water Company	Yes	Yes
Mining privatisation	Yes	Yes
Telecommunication legislation	Yes	Yes
Western Siberia Oil and Gas	Yes	Yes
Vneshtorgbank	Yes	Yes
IDB sector Lending	Yes	No
Guatemala Health sector	Yes	Yes
Reform of tax administration	Yes	No
Mexican Health services	Yes	No
Social Investment Funds	Yes	No
IDB Policy-Based Loan Portfolio	Yes	No
IDB Public Sector Reform	Yes	No
EIB borrowing in Spain	Yes	Yes
EIB contribution to regional development in Portugal and Italy	Yes	Yes
Integration of capital markets	Yes	Yes
Evaluation of Water projects	Yes	No
Watershed Development India	No	Yes
Transport Rehabilitation Ghana	Yes	Yes
Municipal	No	Yes

development Brazil		
Local Development Brazil, Philippines	No	Yes
Agricultural research Burkina Faso	No	Yes
Philippines Country Assistance Review	Yes	Yes
Water Supply Thematic study	Yes	Yes
Education sector Malawi	No	Yes
Agricultural Bank Nigeria	No	yes

Table 3. Methods used in the evaluations

Title of report	Human resource development	Organisational development	Networks	The regulatory environment	Values, norms and attitudes
Energy sector Mongolia	I	I,O,D	-	I	-
Technical and vocational education	I,S,D	I,O,D	I	I	I
Reform public expenditure management	I,D	I,O,D	I	-	I
Capacity building Indonesia	I,S,D	I,O,D	I	-	I
Training in education projects	S	I,S,D	-	-	-
Toll Motorway project	-	I,D	-	I	-
Estonian Water Company	I	I,D	I	I,D	-
Mining privatisation	I	I,O,D	-	I,O	-
Telecommunication legislation	I	I	-	I,D	-
Western Siberia Oil and Gas	I	I	I	I	-
Vneshtorgbank	I	I	I	-	-
IDB sector Lending	-	D	-	-	-
Guatemala Health sector	-	I,D	I,D	I,D	-
Reform of tax administration	-	I,O,D,S	I,D,S	-	-
Mexican Health services	I,O	I,O,D	I,D	-	-
Social Investment Funds	D	I,O,D,S	I,O,D,S	I,D	-
IDB Policy-Based Loan Portfolio	-	D	D	D	-
IDB Public Sector Reform	-	D	D	D	-
EIB borrowing in Spain	D	D,S	I,D,S	D,S	I,D
EIB contribution to regional development in Portugal and Italy	-	O,D,I	O,I	-	-
Integration of capital markets	-	-	I,O,D,S	I,O,D	I,D
Evaluation of Water projects	-	D	D	-	-
Watershed Development India	-	I,O	I,O	-	-
Transport Rehabilitation Ghana	I	I	I,D	-	-
Municipal	-	I	I,D	-	-

development Brazil					
Local Development Brazil, Philippines	-	I,D,O	I,S,D	I,D	-
Agricultural research Burkina Faso	-	D,I	D,I	-	-
Philippines Country Assistance Review	-	-	I,D	I,D	-
Water Supply Thematic study	-	D	D	D	D
Education sector Malawi	-	I,D	D	-	-
Agricultural Bank Nigeria	-	I,D,S,O	-	-	-