EVALUATION OF HOUSING PROJECTS

METHODOLOGICAL CONSIDERATIONS BASED ON SELECTED CASE STUDIES FROM DEVELOPING COUNTRIES

by

SUSANNE RASMUSSEN-COSTELLO

M.Sc., Civil Engineering and Planning
University of Aalborg, Denmark
(1987)

Submitted to the Department of Urban Studies and Planning
in Partial Fulfillment of
the Requirements of the Degree of
Master of City Planning

at the

Massachusetts Institute of Technology

May 1989

Signature of Author _____
Department of Urban Studies and Planning
May 8, 1989

Certified by _____
Bishwapriya Sanyal
Associate Professor
Thesis Supervisor

Accepted by _____
Donald Schön
Chairman, Master of City Planning Program

EVALUATION OF HOUSING PROJECTS

METHODOLOGICAL CONSIDERATIONS BASED ON SELECTED CASE STUDIES FROM DEVELOPING COUNTRIES

by

SUSANNE RASMUSSEN-COSTELLO

Submitted to the Department of Urban Studies and Planning
in Partial Fulfillment of
the Requirements for the Degree of
Master of City Planning

ABSTRACT

Evaluation studies can be an important planning tool but they are often poorly conceived and thus generate little knowledge that is useful to planners.

Nine evaluation reports of housing development projects, including slum upgrading, sites and services, and direct construction in various developing countries are discussed. The reports are divided into 4 general types of evaluation depending on design and modes of data collection and analysis.

The usual dichotomy between quantitative and qualitative methods misstates the problem of methodological choice. Instead it is suggested that there are two research paradigms each of which has a different emphasis on quantitative and qualitative methods. Based on these paradigms there are two evaluation approaches. These approaches are termed cost-effectiveness evaluation and illuminative evaluation.

The author's field research of low income housing projects in Nairobi, Kenya in 1987 serves as a case study of her contention that effective evaluation is more than a methodological choice. The paper argues that a proper understanding of the socio-political context, in which projects are developed and implemented, is of overriding importance if useful knowledge is to be generated.

Finally, the paper discusses the choice of evaluator and the link between the institutional affiliation of the evaluator and the evaluation approach adopted.

Thesis Supervisor: Bishwapriya Sanyal
Associate Professor

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

BDA	BILATERAL DONOR AGENCY
FSDVM	FUNDACION SALVADORENA DE DESAROLLO Y VIVIENDA MINIMA (SALVADORAN FOUNDATION FOR LOW COST HOUSING)
G	NATIONAL GOVERNMENT
GBL	GOVERNMENT OF BOLIVIA
GBT	GOVERNMENT OF BOTSWANA
GEC	GOVERNMENT OF ECUADOR
GEG	GOVERNMENT OF EGYPT
GSL	GOVERNMENT OF SRI LANKA
GZA	GOVERNMENT OF ZAMBIA
GZI	GOVERNMENT OF ZIMBABWE
UNCHS	UNITED NATIONS CENTER FOR HUMAN SETTLEMENTS/HABITAT
UNDP	UNITED NATIONS DEVELOPMENT PROGRAMME
USAID	UNITED STATES AGENCY FOR INTERNATIONAL DEVELOPMENT
WB	WORLD BANK

INTRODUCTION

Recently, I attended a lecture on techniques for monitoring and evaluation of development projects. In summing up the discussion the lecturer, a very experienced and insightful professional, recommended that evaluation should be kept at a minimum. He argued, that evaluation had, since its widespread adoption in development assistance in the mid-seventies, generated large volumes of paper and computer-tapes with data. Sometimes the data had been generated in an unusable form, and most of the time the data had remained unused, regardless of the way it had been presented. On the whole, enormous sums of money had been wasted on evaluation, and thus it should be kept at a minimum.

While the lecturer´s critique of evaluation is insightful - and there is a fairly widespread skepticism about the value of evaluation - I believe it is wrong to conclude that evaluation should be kept at a minimum. I consider evaluation the only effective way in which we can systematically learn from our experience with development projects and, in particular, the only way the knowledge generated can potentially enter into the realm of public discourse. Evaluation can thus prevent us from having to "reinvent the wheel" every time we begin a project. I

believe, it is a question of how, rather than if, development projects should be evaluated. In this paper I discuss this question, both empirically and theoretically, with reference to housing projects in developing countries.

In Chapter 1 I review nine evaluation reports of housing projects including slum upgrading, sites and services, and direct construction projects. The nine reports are divided into four general types of evaluation according to their design and modes of data collection and analysis.

Chapter 2 discusses problems and advantages of the designs used in four of the evaluations as examples of the general types of evaluation. A discussion follows, in which I argue that the usual dichotomy between quantitative and qualitative methods misstates the problem of methodological choice. Instead I suggest the existence of two different research paradigms, each with a different emphasis on and usage of quantitative and qualitative methods. The chapter concludes with a discussion of two distinct approaches to what an evaluation is, for which I have adopted the terms cost-effectiveness evaluation and illuminative evaluation.[1] The nine evaluations initially reviewed are then used as examples of the two approaches.

Effective evaluation is, however, more than a methodological issue. In Chapter 3 I discuss why I believe

---

[1]. I have adopted and adapted these terms from Howard Richards' discussion of evaluation in Richards, 1985.

that a proper understanding of the socio-political context of low income housing projects is of overriding importance in evaluation research. Here I rely on the low income housing experience from Nairobi, Kenya as the empirical base for the discussion. Much of this material is based on my stay in Kenya during the winter of 1986-87.

In the last chapter, Chapter 4, I discuss the choice of evaluator, and the link between the institutional affiliation of the evaluator and the evaluation approach adopted. I argue that many of the shortcomings of evaluation are tied to institutional constraints. The different agendas and pressures on donors and recipients often stand in the way of effective evaluation.

CHAPTER 1

THE STATE OF THE ART: A REVIEW OF HOUSING EVALUATION REPORTS.

As a starting point for my discussion of evaluation methodology I shall review a sample of evaluation reports. All the reports reviewed cover evaluations of one or more low-income housing projects in a Third World country. The housing projects fall in three groups: direct construction; site and services including construction of core-houses; and upgrading projects. The projects differ in the provision of public utilities, community facilities and services, and credit. The projects are similar in that they are all fully or partly funded by external donors. The evaluations represent 8 countries and 5 multi- and bilateral donor agencies.

The design of the sample does not fall under any of the usual definitions of sample design. Random, quota, purposive, and other sampling procedures all presuppose that somehow there is equal access to the elements of a population from which one tries to construct a sample. Getting access to evaluation reports is, however, a problem. I have discovered that in many cases donor agencies consider evaluation reports 'classified material'; they are rarely

published and seldom available to the public upon request.[2] For this reason I have only been able to construct a sizeable sample by adding reports obtained through 'the network' to those available through libraries.[3]

It is alarming that donor agencies - and perhaps governments[4] - try to monopolize the knowledge generated through evaluations or at least keep it out of the hands of the general public. On the other hand, the fact that agencies do try to reserve this information tells us something important for the following discussion of evaluation methodology. First, it seems obvious that both donor agencies and governments in general are very sensitive to exposure and to admitting failure.[5] Second, one could expect publicly available reports to be biased in favor of the involved

---

2. USAID evaluation reports constitute an exception. All USAID evaluation reports are considered public domain and are for sale on a per page fee. This is very significant in the light of the restrictions other agencies place on evaluation reports.

3. This explains why the name of the donor agency, the city and the country has been changed in one of the evaluations.

4. The donor agencies often use their counterparts, i.e. recipient governments, to explain why an evaluation report cannot be 'released'. They have agreed, they say, with the governments that evaluation reports are only for internal use within the donor and government agencies, respectively.

5. Certainly, the monopolizing of information would not be caused by a desire to hide success.

government and donor agencies.[6] I shall come back to these two questions later in my discussion.

Despite the difficulties I encountered in getting access to evaluation studies of housing projects, my sample represents a broad range of methodologies used in evaluating housing projects. Even if I cannot claim randomness or representativeness, judging from my general knowledge about evaluation methodology, a discussion of my sample of evaluation reports will reflect the state of the art in evaluation methodology for housing projects.[7]

The projects

Each evaluation report covers one or more low income housing projects including direct construction, sites and services and upgrading projects. Implementation of the first and the last project started in 1974 and 1985, respectively, and some of the projects in the sample are currently still in the implementation phase. Figure 1.1 lists the main characteristics of the projects.

The evaluations

One way of representing and comparing the evaluations is

---

[6]. There are other reasons why this would be the case even if the report was not publicly available. An analogy from the business world will illustrate this point: if you are in business you must be careful not to insult your customer - or you may never do business again.

[7]. One type of evaluation not represented in my sample is participatory evaluation, which is widely used by for example Oxfam America and other agencies. See Appendix I for a discussion of the advantages and problems of participatory evaluation.

to provide a general list of housing project aspects to see which of these aspects are covered in each of the evaluations. Figure 1.2 is a matrix listing evaluations and aspects covered.

While there are some differences in the components covered - notably concerning project impact in general and project impact on project participants and community development in particular - many of the evaluations cover similar aspects of the projects. On the other hand, the evaluations are strikingly similar in their almost unanimous exclusion of gender issues and issues relating to the socio-political context of low income housing. Possible reasons for leaving these issues out and the consequences of excluding them shall be dealt with in greater detail in a later chapter.

The initial similarity, however, concerning the aspects of the projects covered, is not matched by a similarity in data collection or in evaluation design.

Data Collection

The evaluation studies represent a host of sources for collection of data: interviews, questionnaires, participant-observations, documents and records, government statistics, and physical evidence from site visits. The data collection methods roughly fall into three broad categories: some evaluators rely on review of documents and records and interviews with project management for their data; some

evaluators add socio-economic surveys of project participants; and some move into the communities they are evaluating to collect data through long-term interaction with project participants and through experiencing what it is like to live in the project, also known as participant-observer evaluation. Figure 1.3 lists the data collection sources used in the reviewed evaluations.

All evaluation studies include quantitative and qualitative data, and there is a clear tendency to include more qualitative data the bigger the emphasis on collecting data from project participants. Notably, none of the evaluations prepared for USAID used project participants as a source for data collection. This appears to be characteristic of a majority of USAID housing evaluations.[8]

---

[8]. To test the significance of this observation I conducted a quick survey of a large number of annotated bibliographies of USAID evaluations of housing projects. Thirty annotations specifically stated data collection methods used in the evaluation. Out of the thirty evaluations eight - or slightly over 25% - had used interviews with project participants as a source for data collection.

HOUSING PROJECT PROFILE MATRIX

- characteristics of evaluated projects -

PROJECT / CHARACTERISTICS	EL SALVADOR	ZAMBIA	BOLIVIA	ECUADOR	SRI LANKA	EGYPT	BOTSWANA	ZIMBABWE	SOUTHERN AFRICA
*IMPLEMENTING AGENCY	FSOVM	GZA	GBL	GEC	GSL	GEG	GBT	GZI	G
*DONOR	WB	WB	WB	WB	AID	AID	AID	UN	BDA
*YEAR COMMENCED	1974	1974	1977	1980	1978	1978	1982	1982	1979
*YEAR OF EVALUATION	1982	1981	1982	1982	1982	1982	1988	1985	1985
*EVALUATOR	FSOVM WB	GZA WB	WB	WB	USAID	USAID	GBT USAID	GZI UN	BDA
*COST (mill. $)									
- estimated		42.1	24.5	38.6		120	c) 9.3	(23) 3.8	2.7 e)
- actual								3.8	
*SITES AND SERVICES									NA
- # plots									
- target	7,000	4,339	2,000	3,200	30,000		2,362	1,000	
- actual/current	7,000	3,665			43,137		2,351	1,045	
*UPGRADING	NA							NA	
- # dwellings									
- target		24,512	7,000	3,400	8,720		d)		2,200
- actual/current		27,691 b)			3,488		1,775		
*DIRECT CONSTRUCTION		NA	NA				NA	NA	NA
- # dwellings									
- target	3,500 a)			504	36,000	200			
- actual/current	3,500				14,016				
*INFRASTRUCTURE	X	X	X	X	X	X	X	X	X
*COMMUNITY FACILITIES	X	X	X		X	X	X		X
*CREDIT									
-materials/construction	X	X		X	X	X	X	X	X
-small-scale enterprise	X		X	X					X
*TECHNICAL ASSISTANCE	X	X	X	X	X	X	X	X	X

Figure 1.1: Housing project profile matrix.

a) included in 7,000 sites and services plots
b) as of June 1981
c) funds spent by 1988
d) approximate figure
e) does not include technical assistance and other amounts difficult to allocate per program

HOUSING PROJECT EVALUATION MATRIX

- project aspects evaluated -

ASPECT \ PROJECT	EL SALVADOR	ZAMBIA	BOLIVIA	ECUADOR	SRI LANKA	EGYPT	BOTSWANA	ZIMBABWE	SOUTHERN AFRICA
*HOUSING PROBLEM AND POLICIES	X	X	X	X		X		X	
*PROJECT EVOLUTION		X	X	X	X	X		X	X
*PROJECT OBJECTIVES	X	X	X	X		X	X	X	X
*PROJECT CONTEXT									
- socio-economic	X	X	X	X		X			X
- politico-bureaucratic		X	X	X		X			X
*LAND ACQUISITION		X						X	
*CHARACTERISTICS OF BENEFICIARIES		X	X	X				X	X
*PROJECT OUTPUTS:									
- physical:									
- design and std's	X	X	X	X	X	X	X	X	X
- house constr.	X	X	X	X	X	X	X	X	X
- infrastructure	X	X	X	X		X	X	X	X
- community/commercial facil's	X	X	NA	NA	NA	X	X	NA	
- bldg. materials	X	X	X	X		X	X	X	X
- community dev't:						NA			
- plot allocation/resettlement	X	X					X	X	X
- modes of selfhelp	X	X	X	X				X	X
- income/employment programs	X	X	X	X				X	X
- comm. activities	X	X	X	X				X	
- public services	X	X	X	X				X	X
- technical ass.	X	X	X	X	X	X	X	X	X
- fiscal/economic:									
- cost of infrastr. and facilities	X	X			X	X	X	X	
- cost/house const.	X	X	X	X	X	X		X	X
- credit programs	X	X	X	X			X	X	X
- cost recovery	X	X	X	X	X	NA	X	X	
- legal:									
- tenure/ownership			X	X				X	
- eviction/demol.							X		

(--continued next page--)

HOUSING PROJECT EVALUATION MATRIX
(--continued--)

ASPECT \ PROJECT	EL SALVADOR	ZAMBIA	BOLIVIA	ECUADOR	SRI LANKA	EGYPT	BOTSWANA	ZIMBABWE	SOUTHERN AFRICA
- institutional:									
- inst. dev't	X	X				X			X
- staffing/mgt.		X	X	X	X	X	X		X
- inter-agency collaboration		X				X	X		
*PROJECT IMPACT						NA			
- project participants:						a)			
- targeting and affordability	X	X	X	X	X			X	
- hous. cost/qual.	X	X	X	X	X			X	X
- access: services and facilities	X	X	X	X				X	X
- income generation	X	X	X	X				X	X
- comm. dev't.	X		X	X				X	
- part.satisfaction	X		X	X				X	X
- gender aspects			X	X				X	X
- housing:									
- housing stock					X				
- housing policies and standards		X			X			X	X
- city/area:									
- political impact		X	X	X					
- social impact		X	X	X					
- local finance		X							

a) project delay made it impossible to evaluate impact.

Figure 1.2: Housing project evaluation matrix.

EVALUATION DATA COLLECTION MATRIX
- data collection sources used -

PROJECT / SOURCE	EL SALVADOR	ZAMBIA	BOLIVIA	ECUADOR	SRI LANKA	EGYPT	BOTSWANA	ZIMBABWE	SOUTHERN AFRICA
*DOCUMENTS AND RECORDS	X	X	X	X	X	X	X	X	X
*PHYSICAL EVIDENCE	X	X	X	X	a)	X	X	X	X
(site visits)					a)				
*QUESTIONNAIRES									
- project management									
- project participants	X	X	X	X				X	X
*INTERVIEWS					a)				
- housing officials (government, donor, and other)	X	X	X	X		X	X	X	X
- project management	X	X	X	X		X	X	X	X
- project participants			X	X					
*PARTICIPANT OBSERVATION			X	X					

a) no source indicated

Figure 1.3: Data collection sources.

Design

The designs used in the 9 evaluation studies represent roughly four general types of evaluation designs: goals evaluation, impact evaluation, self evaluation and participant-observer evaluation. I shall briefly describe each type.

Goals evaluation is primarily concerned with measuring whether actual output matches the intended output of a project and whether the outputs were produced efficiently. Goals evaluation uses physical development and management performance as indicators of achievement of goals and efficiency in implementation.

Impact evaluation tries, as the name implies, to assess the socio-economic impact of the project. Impact evaluation uses experimental and quasi-experimental designs to measure longitudinal impact of the project on project participants. Cost-benefit analysis or other methods for assessing cost-effectiveness are applied to measure the economic impact of the project.

Self evaluation does not prescribe the use of any particular methods for data collection. Through interaction between the evaluator and project management it is established which issues are of importance for the success or failure of the project and its components. The evaluation seeks to document and organize project experiences and lessons.

The aim of participant-observer evaluation is similar to that of self evaluation, but it documents and organizes project experiences and lessons through immersion in the project as a participant. The evaluator moves into the project for an extended period of time and tries to learn about reasons for project success or failure from viewpoint of the project participants.

The four types of evaluation merely represent a rough typology, and variation within each type and overlaps between them may indeed occur. On the other hand, I believe it is entirely possible to describe the nine studies I have reviewed in accordance with this typology. Figure 1.4 lists

the four evaluation designs and how they apply to the nine studies.

EVALUATION DESIGN MATRIX
- evaluation designs used -

PROJECT / DESIGN	EL SALVADOR	ZAMBIA	BOLIVIA	ECUADOR	SRI LANKA	EGYPT	BOTSWANA	ZIMBABWE	SOUTHERN AFRICA
*GOALS EVALUATION					X	X	X	X	
*IMPACT EVALUATION	X	X							
*SELF EVALUATION									X
*PARTICIPANT-OBSERVER EVALUATION			X	X					

Figure 1.4: Evaluation designs.

In the following chapter I shall discuss the four evaluation designs, including data collection and analysis methods applied as strategies for evaluation of social action programs, such as housing projects in developing countries.

It should be added, that although I have identified four evaluation design types, if the evaluations are viewed according to their general purpose, the evaluations can be reduced to two general approaches, viz. cost-effectiveness evaluation and illuminative evaluation. This discussion will be fully developed in the following chapter.

CHAPTER 2

EVALUATION METHODOLOGY

In the previous chapter nine evaluations of housing projects in developing countries were briefly reviewed concerning aspects of projects evaluated, data collection sources used and evaluation designs used. Four general evaluation designs were identified among the reviewed evaluations. This chapter continues the discussion of the evaluation designs. Four of the evaluations - serving as illustrations of the four design types - will be discussed in greater detail. The evaluations will be discussed in more or less detail, according to their methodological scope and relative importance as examples of evaluation research.

A theoretical discussion will follow the discussion of evaluation design, addressing the application of different research paradigms to the evaluation of social action programs, i.e. low income housing projects in developing countries.

The last section of this chapter examines and compares the nine evaluation studies from the perspective of the general purpose of project evaluation.

Evaluation Designs: A Critique.

Goals evaluation: the case of Gaborone West I and Francistown IV in Botswana.

The field work for the evaluation of two USAID-financed housing projects - a site and services project and an upgrading project - was carried out in a 3-week period in 1988. The clients for the evaluation were the Ministry of Local Government and Lands in Botswana and USAID.

The evaluation design is characterized by 1) an emphasis on quantitative data and non-rigorous treatment of qualitative data; 2) a near total exclusion of data on or obtained from project participants; and 3) its selective focus on achievement of physical development and of management and training objectives.

The evaluation relies heavily on quantitative interpretations of the degree to which project objectives were achieved. Most project components are evaluated in terms of number of units planned vs. actual output, i.e. success is measured in quantities. A case in point is the building materials loan program:

> It is felt that the Building Materials Loan Program has worked well in Gaborone West I. More than 64% of the SHHA residents have received loans.[9] (emphasis added)

Qualitative aspects of the projects are covered under a rubric called "Comments". What is interesting about the qualitative interpretations is that the majority of them are unsubstantiated. It is as if the evaluators apply two different standards of rigor to the two types of data. A

---

[9]. USAID, 1988:12. SHHA stands for Self Help Housing Agency.

clear example is provided in the discussion of Gaborone Self Help Housing Agency:

> The main problem appears to be with collections. The ward office community development workers are required to visit residents and encourage payment of the levy and BML loans. Since they have no authority to impose sanctions, and probably little training in this area, it is not very effective. In addition, the workers encounter hostility from the residents who do not want to be reminded that they are behind in payments. Since an unpleasant task produces little benefit, it is not surprising that the staff is discouraged. It is also likely that the negative feelings on the part of beneficiaries adversely affects the other education and community development activities of the department.[10] (emphasis added)

It is not within the realm of my research to explain exactly why different standards appear to exist for quantitative and qualitative interpretations of the project - but I shall suggest a couple of possible explanations.

One explanation could lie in standard USAID guidelines for evaluation, which require only quantitative information and, in turn, led the evaluators only to collect quantitative data. Realizing the project could not be interpreted only through quantitative data, and not appreciating - or understanding - how to produce rigorous qualitative data, the evaluators chose the format of 'comments'.

Another explanation has to do with the politics of evaluation. The 'comments' may be a way of providing an 'insider', i.e. relevant USAID staff, with information without offending or exposing anybody. This would explain the

---

[10]. ibid., 1988:22.

very vague terms used to describe some of the project's problems.[11]

Needless to say, a narrow technical-administrative focus for the evaluation and a near total exclusion of data on and/or obtained from project participants biases the evaluation.[12]

It becomes an evaluation of certain, predetermined aspects of a project rather than of the project as a whole. If, as is often the case, the outcome of a project is related to social or political circumstances - as I shall discuss in a later section of this paper - an evaluation after this model will be of little use for future policy making and project design.

Impact evaluation: the case of the First El Salvador Sites and Services Project.

The field work for the El Salvador project evaluation was carried out in a five-year period from 1975-80. The clients for the evaluation were the executing agency, the

---

11. This argument is supported by the general nature of many USAID evaluation reports. Criticism of projects is usually subdued - one gets the general notion that USAID projects are always very successful - and criticism of host governments is practically non-existent.

12. Ironically, the following comment in the evaluation report points out that such data is necessary in 'proper evaluation': "Development standards for residential areas should be based on the needs as well as the financial capacity of the target population.... but, with proper evaluation, this will be understood before accepting the standards. (emphasis added) Such data was, however, not collected by the Botswana evaluation team. Source: USAID, 1988:21.

Salvadoran Foundation for Low Cost Housing (FSDVM), and the World Bank.

The evaluation of the First El Salvador Sites and Services Project included short-term evaluation studies, designed to provide immediate feedback to project management; medium-range evaluation of project components designed to modify plans for later phases of the project still in the planning stage; and long-range impact and policy studies, designed to evaluate overall project impact and provide input for future shelter policies for the implementing agency and national planning agencies in El Salvador.[13]

The evaluation's main focus was on long-term impacts and the achievement of project objectives. A detailed analysis of the design of the long-range studies will follow. Some general comments related to the short- and medium-range studies will conclude the section.

The long-range studies included longitudinal impact studies, a cost-benefit comparison of low-income housing options, project benefit evaluation through hedonic price analysis and an evaluation of project impact through comparisons of changes in value of the house.

Longitudinal impact studies.

A quasi-experimental design was used to evaluate the

---

[13]. Bamberger, 1982a:238. It should be added, although it is not stated in the evaluation report, that the evaluation study has potential value to the Bank as input to its policy making process.

impact of the project on the socio-economic conditions of the project participants. The choice of a quasi-experimental design proved to entail a number of difficulties for the impact evaluation. The experimental sample was formed by the 196 families selected for the project, but who had not yet moved there. A stratified control sample was constructed from three main types of low-cost housing in the city.[14]

The construction of the control sample was a major practical and methodological obstacle. Maps had to be drawn of 67 communities to construct a sampling frame for the drawing of a random sample. The sample of the mesones was affected by methodological bias[15] and the sample of the colonias ilegales potentially had an even larger bias[16]. Furthermore, major theoretical problems arose from using a

---

[14]. There are three types of low-cost housing in Santa Ana: "mesones" (tenement houses), "colonias ilegales" (extra-legal subdivisions), and "tugurios" (illegal squatter settlements).

[15]. Since Santa Ana has 1000 mesones the task of constructing a sampling frame by drawing up maps for all the mesones was too big and a sample of 50 mesones out of the 1000 was drawn at random. The 'random sample' drawn from the [50] mesones is "...technically a cluster sample but as both the total number of mesones and the number of mesones in the sample, were quite large, it is argued that for operational purposes the sample can be considered to approximate a simple random sample." Bamberger et al., 1982a:255.

[16]. All of Santa Ana's 30 colonias ilegales "[..] were visited, classified over a series of indicators, and then ranked and classified into two groups, poorer and richer. A sample of five colonias was selected in each group. ibid.

non-equivalent control group, because of significant initial differences between the project participants and the control group. Even if differences were found between "...the income of the two groups in T(3) we cannot immediately infer that the difference is due to project impact."[17]

To compensate for this situation multiple regression analysis was used. Regression results for income were computed on family earnings, earnings of head of household, and earnings of secondary workers to support the analysis of the project's impact on income.

Ironically, it seems that the laborious and statistically advanced techniques used in the impact evaluation of the project concerning income did not produce much 'hard evidence' of the actual economic impact, as measured in income changes for the participants, as the summary of findings from the evaluation report reveals:

> Our results indicate that the participants may have experienced a slightly greater increase in earnings than the non-participants and that the less well off they were, the better they performed with respect to earnings per worker. Whether one considers this a project impact is a moot point as earnings are a function of a variety of things and their connection with housing is by no means direct, nor is it a strong one. However, there seems to be some evidence of motivation at work among the participating households as indicated by a relative stability in the employment patterns among their secondary workers. (emphasis added)[18]

In other words, aside from the methodological problems

---

[17]. ibid., 1982a:257.

[18]. ibid., 1982a:270.

that arose from choosing a quasi-experimental design, the choice of a quantitative method left the evaluators without much information. Quantitative methods are - as I shall discuss in more detail in the following chapter - largely "knowledge verifying". The knowledge about the relation between housing and income in developing countries is, however, rather incomplete, as the evaluators admit. Even if the design of the quasi-experiment had proved successful, it would not be possible to relate the different aspects of the project to potential changes in income.

Cost-benefit comparison of low income housing options.

A cost-benefit comparison was made of eight housing options accessible to the low-income population of San Salvador. The information on cost was obtained from the organization or the family responsible for construction. Information on benefits was estimated by asking families how much they thought they could earn from renting their house ("imputed rent"). To generalize these findings conversion factors were used to calculate cost and benefits to the nation, and flows were estimated over both 20 and 30 year time horizons. The resulting cost-benefit ratios were used for efficiency analysis, social analysis and private analysis. Different assumptions about the land values and consumer surplus were included in the calculations.

Aside from the usual criticism of cost-benefit

analysis[19] a number of other factors affect this particular design. Firstly, imputed rent - derived from interviewing families about how much they thought they could earn in rent - may be a highly unreliable figure, especially over time. The size of the rent is influenced by how many families decide to rent, the proximity of the project to work-places and public facilities, the maintenance of the project, the implementation of other low-income housing projects in the city, the economy in general, etc. It is questionable whether families have access to this information and are able to estimate the imputed rent. Secondly, projections of cost and benefits over 20 and 30 year time horizons seem highly unreliable in light of the typically unstable Third World context.

Hedonic price analysis.

Hedonic price analysis was used to eliminate the effects of government regulation and thus obtain a 'free market' picture of project impact on housing.

Using hedonic price analysis for evaluating project impact is, according to the evaluators, considered somewhat problematic. As stated in the report:

> Care must be taken in the interpretation of the results as the estimates only refer to imputed rent and do not include the value of owning as opposed to renting. One way to interpret the hedonic estimation of rent is to assume that this indicates the amount the family would have been prepared to pay to receive this package of

---

[19]. For a critique of cost benefit analysis see, for example, Skaburskis, 1987; Tolchin, 1987.

> benefits. If the imputed rent is compared with the amount the family actually has to pay to the [implementing agency], the difference can be interpreted as indicating additional benefits the occupant obtains "for free".[20]

In other words, the hedonic price analysis is based on a number of more or less controversial interpretations of the calculations used in the analysis. This may seriously challenge the result of the analysis. Furthermore, one can question the actual value of a hedonic price analysis that eliminates the effect of government regulation of the housing market. It is rather unlikely that a free market situation will ever exist, and it is significant to World Bank sites and services projects that restrictions on selling and renting apply.

Comparisons of changes in value of the house.

As an alternative method for evaluating project impact on housing value, estimates of changes in housing value were derived and compared with changes in cost. Again, the conclusions made on the basis of the calculations are not conclusive, in fact, the evaluators concluded that they could only, "make some tentative judgement as to which type of housing produces the highest benefits for occupants."[21]

Conclusion on long-range impact and policy studies.

The methodologies used for the long-range impact and policy studies of the First Sites and Services Project in El

---

[20]. Bamberger et al., 1982a:273-4.

[21]. ibid.:277.

Salvador were all based upon the evaluators´s ability to generate rigorous and reliable data as input for the rather laborious and complicated calculations used to reach conclusions about project income - data that often does not exist in developing countries and which is hard to develop. The quasi-experimental design also proved infeasible - not an uncommon situation in a Third World context. The shortcomings of the experimental design is highlighted by the following passage from the evaluation of the World Bank project in Zambia:

> To conduct a rigorous impact evaluation it is necessary to have a control group against which to gauge the significance of the changes observed in the project areas. [But] The control group originally included in the research design was dispersed due to flooding,....Further difficulties were caused by problems of matching families who moved to overspill areas with the same families interviewed in the upgrading areas before they moved. In the process of moving, many families went to unnumbered plots which made it very difficult to match the families.[22]

Aside from practical problems with design and data-collection, the particular context of the evaluation also proved unsuited for the use of quantitative, primarily "knowledge-verifying", techniques. These techniques - adopted from natural science research - can in reality only invalidate a theory about a causal relationship.[23] They

---

[22]. Bamberger et al., 1982b:43-44.

[23]. Hudson (1975) writes: "., science can never <u>prove</u> anything, but only <u>disprove</u> ideas that are false. This is in fact a basic intention of a scientific proposition - that it can be demonstrated false if it is false - in contrast to religious or moral or

cannot prove anything nor can they provide any explanation as to the reasons for causality or lack thereof. As the longitudinal impact study of income showed, if little is known about causal relationships the results of a complicated quantitative analysis may prove of little use.

Short and medium range studies.

This section is devoted to some general comments on the methodologies used for data collection for short- and medium-range studies.

The short-range studies were concerned with estimating the potential demand for the type of housing to be implemented as the First El Salvador Sites and Services Project and evaluating the reasons for families to drop out of the project once implemented. The medium-range studies evaluated different project components, i.e. the cooperative program, the mutual help program, the housing consolidation process, the participant selection process and the economics of self-help housing.

A variety of different data-collection methods were used in the evaluations: case studies, semi-structured interviews, questionnaire surveys, participant observations, review of project records and of data on social, economic and political factors affecting certain components of the project.

---

aesthetic or other nonscientific contentions. (It is not possible to prove that God doesn't exist, or that Head Start was a failure. All that science can "prove" is that God and Head Start haven't granted certain specific wishes).", p.81.

Obtaining information from project participants on certain economic factors of the project often prove particularly difficult. Many people do not keep records of their expenditures and cannot recall how much they have spent if a long time has lapsed before they are interviewed. Furthermore, some people are reluctant to give information about their income if it differs from the income-range defining eligibility for the project. These issues were dealt with in the El Salvador evaluation by interviewing participants about construction cost at three different points in time; and by a series of consistency checks of the information such as comparing income with expenditure or with typical wage rates. If inconsistencies were found a return visit was made. These methods improved the reliability of the data collected for the evaluation.

In general, obtaining information from people in developing countries by means of surveys is encumbered with difficulty if such information is perceived as controversial like a survey of family income. Even if the income-range requirements only count at the time of application, a family may still be fearful of providing the correct information, if indeed they gave incorrect information on the application form - a common occurrence in many projects. An example of the difficulty involved in getting controversial information from surveys is provided by Thorner and Thorner in their

analysis of the notorious All-India Credit Survey[24]:

> There are fields in which the errors of response may be expected to be very large as, say, in connection with sensitive subjects on which questions may be resented, e.g., bootlegging, selling opium, sex relations, gambling, tax evasion, etc. ... [Also] debt is a painful subject, not likely to be discussed easily or candidly with anyone, let alone casual strangers carrying formidable questionnaires. ... To the peasant, the moneylender is likely to be "at once a necessity and a terror." The peasant might like to curse him, but may find it wiser to keep mum.[25]

Evaluation of project context.

As described the evaluation primarily emphasized the use of quantitative data and scientific[26] methods, although in some instances efforts were made to include qualitative information provided either by project participants or from data on social, economic and political factors. Significantly, however, no attempt was made to include the effects of the 1979 coup in El Salvador and the following period of unrest in that country in the evaluation, although the evaluation report states "...they obviously have had extremely negative effects in the area of housing."[27] In the report it is argued that most of the evaluation was over some months before the coup. The report, however, was not

---

24. The All-India Credit Survey is a 'text-book case' of how not to design, conduct, interpret and use surveys.

25. Thorner and Thorner, 1962:205.

26. The word 'scientific' is here and in the remainder of the paper used to describe that which originates from natural science as opposed to that which is systematic and exact.

27. Bamberger et al., 1982a:2.

published until almost 3 years later. While the occurrence of political coups and periods of unrest is not an uncommon phenomena in many Third World countries[28], it is interesting that no attempt was made to provide even a rudimentary analysis of the coup's impact on the project.

Self evaluation: the case of the Southern Africa Upgrading Project.

The field work for the evaluation was carried out during a three week period in 1985. The immediate clients for the evaluation are the funding and implementing agencies - a non-American bilateral donor agency and the local municipal council, respectively. The evaluator, however, refers to "..those locally involved in upgrading.." as the ultimate judges of the relevance and utility of the evaluation study.

The evaluation design is characterized by presenting a series of issues, gathered from the families and professionals involved in the project, rather than being an in-depth, item-by-item evaluation. It is furthermore focused on a self-evaluation approach, i.e. only issues that arose during discussions and were pointed out to be more critical were included in the evaluation.

Methods used for data collection include review of reports and documents; field visits to areas not upgraded and previous upgrading projects; discussions with housing

---

[28]. Bolivia, for example, experienced 189 coups during the first 155 years of independence (1825 - 1980) - averaging a coup every 10 months. Source: Information Please Almanac, 1983:155.

officials and project staff members; and a brief survey of families in selected sections of the project. The data collected is organized in four sections concerning: 1) issues for project formulation; 2) issues for project implementation; 3) issues for project administration; and 4) issues for project future.

The significance of this evaluation is the lack of a preconceived design for the inquiry and the reliance on project participants and professionals to define the issues of the evaluation.

A genuine attempt is made to include the participants' view, although the evaluation seems to value input from project participants less than from professionals[29]. The questionnaire used to survey families in the project was administered by local trainees attached to the project; and

> [t]he trainees were chosen since the assumption was made that families may be more open to their questions. The community development workers were not used since it was felt that they would be perceived as representing the City Council and responses may be inhibited.[30]

The questionnaire, however, appears to produce bias against women in the project. One question, for example, asks "What is the employment of the father (household head)?"[31] and in the summary of the sample survey it is stated, that

---

[29]. A relatively short questionnaire with few open-ended questions was given to a sample of families.

[30]. Southern Africa Evaluation Report, 1985:58.

[31]. ibid., 1985:60.

> [a]cceptable respondents [to the survey] were the husband as first choice, and the wife as default.[32]

Another example of bias against women as bread winners is found in the discussion of economic promotion. The primary targets for employment training programs are defined as 'housewives' and teenagers, but the report states that:

> ..it was interesting to note that during the period of review, the attendance of classes for housewives was very poor, explained in part that the housewives did not want to miss a sale of their vegetables on the roadside.[33]

It is obvious here, that these women are engaged in earning an income and thus are wrongly viewed as housewives.

In theory, using a self-evaluation approach where parties affected by the project ('stakeholders') define the issues to be included, should produce highly relevant information. Project participants and professionals have daily experience with the project, its problems and accomplishments. Self evaluation may, however, also produce biases. Firstly, self evaluation is likely to produce bias, if a person or group has a stake in the issues identified for evaluation. This is, for example, often the case when organizations evaluate projects at the end of a funding cycle to make decisions concerning future funding.[34] Secondly, as insiders, project participants and professionals are less

---

[32]. ibid.:58.

[33]. ibid.:23.

[34]. Interview with Joel Charny, Oxfam America. October 28, 1989.

likely to see the project in a larger context and may concentrate only on issues directly related to the project in isolation. Such bias is of course far more likely in cases of 'complete' self evaluation, i.e. where no outsiders to the project are involved in identifying issues and collecting and analyzing data.

I have no way of assessing the specific relevance of the issues included in the Southern Africa evaluation. From my general experience with similar projects the evaluation does, however, touch on many critical issues in what appears to be a balanced way - at least regarding major aspects of the upgrading project as they are described in the beginning of the report. On the other hand, the evaluation shows significant bias against women as breadwinners. It is assumed that women are not heads of households, an assumption which one might question based on data from other low income housing projects in Africa.[35] But, even if no such data existed it is up to the evaluator to prove that women are, in fact, never the only or principal breadwinners of a household, rather than assuming that they are not.

Participant-observer evaluation: the case of Guayaquil and La Paz.

As an experiment with qualitative evaluation methodology the World Bank arranged in 1982 for a consultant to live in two different housing project communities in Guayaquil,

---

[35]. See for example Nimpuno, 1986 and Obudho and Mblanga, 1988:209.

Ecuador and La Paz, Bolivia for about five months in each:

> Arrangements were made with two project agencies in Ecuador and Bolivia for me to live in project communities for varying periods, to participate in local activities, and to develop new approaches and methodologies for improving communication between the participants and those sharing responsibility for the projects.[36]

The primary clients for the evaluations were the World Bank and the managers of the two projects. The evaluation methodology used by the consultant - labelled 'participant-observer evaluation'[37] - is based primarily on qualitative interviews with samples of project participants, and on informal talk and other types of interaction in the project areas.[38] The findings from the qualitative interviews are - whenever possible - supported by quantitative interpretations of the information given by the project participants.

The World Bank initiated this experiment with participant-observer evaluation out of dissatisfaction with the ability of traditional sample surveys and discussion with

---

[36]. Salmen, 1987:5.

[37]. Observation is by some authors defined as a means of collecting information without asking anybody anything - also referred to as 'unobtrusive measures; see for example Weiss, 1972:54. Other authors, e.g. Guba and Lincoln, prefer a broader definition, defining the participant observer as "a genuine participant; that is, he is a member of the group, and he has a stake in the group's activity and the outcomes of that activity." Guba and Lincoln, 1981:190.

[38]. In one case interviews were carried out with residents in a slum area, originally chosen for a World Bank upgrading project, about which it later had been decided that no project were to be implemented.

community representatives to produce relevant and - paradoxically[39] - reliable information about participants's views. Two questions were of particular interest to the Bank:

> The question was, first, whether incorporating the point of view of project beneficiaries into management decisionmaking in this way was useful and cost-effective. If it proved successful, then the question was whether professionals in developing countries would be willing and able to cross class and income lines so as to interpret the perspective of project beneficiaries of their own society and provide useful feedback to host-country and Bank managers.[40]

The primary objective of the evaluation - and of participant-observer evaluation generally - was to provide "...project managers with useful and timely information to help them make decisions that [would] improve the project's performance."[41] A secondary objective was to facilitate people's participation in policy-making; and a final objective was to generate knowledge about development projects - participant-observer evaluation should serve as a preliminary tool for "more rigorous research", i.e. the use of knowledge-verifying methods such as advanced quantitative analyses. An important aspect of this approach is its attempt to avoid the usual dichotomy made in evaluations between

---

39. Quantitative methods are traditionally advocated because of their ability to produce objective, i.e. reliable information.

40. Salmen, 1987:7.

41. ibid.:108.

process and impact evaluation.[42]

The consultant lived in La Paz and in Guayaquil for two periods in each place, separated by an interval of about seven months. During this period he carried out a total of 5 formal surveys of samples of the project participants in La Paz and 6 formal surveys in Guayaquil. The different methods include socio-economic surveys using a closed questionnaire; questionnaire-like interview forms using a mix of closed and open questions; and unstructured qualitative interviews using open questions.

A few examples of the findings from the participant-observer evaluations shall be provided here:

- In one of the projects the community leaders, although well-intentioned, were all property owners from higher income groups and did not represent the interests of the poorer renters.
- In the unstable economic and political contexts of these two cities, property ownership offered one of the few opportunities to achieve a sense of personal security. Consequently, an extremely high value was placed on ownership.
- Living in one of the newly designed houses revealed many of its design shortcomings (lack of ventilation, poor drainage, high noise level from neighbors) in a way which was never appreciated from casual visits.[43]

The method of participant-observer evaluation proved very

---

[42]. Process evaluation is usually known as monitoring. See for example Bamberger and Hewitt, 1986:7.

[43]. ibid.:153-54

successful both in the eyes of the Bank[44] and most of the project managers[45] - some of whom initially had expressed doubts or even opposition to the approach. From the perspective of the project managers the approach brought about information that "most likely would have been ignored by a measurement-oriented quantitative approach."[46] And, of particular interest to the Bank, participant-observer evaluation proved to be reasonably fast and cheap. Indirect savings may be realized by producing quick, reliable and useful information to project management. Furthermore, the evaluation methodology itself is inexpensive regarding salaries and data processing:

> The seven evaluations done by host-country personnel in Bolivia, Brazil, and Thailand cost an average of $15,000 each, while the projects evaluated averaged $9 million in loan amounts. Almost 50 percent of the evaluations' cost in these first efforts was for guidance by World Bank personnel - a cost that will diminish appreciably over time as local institutions gain the expertise to conduct this kind of work on their own.[47]

---

[44]. For some reservations about the participant-observer method on the World Bank's part see, ibid.:154-55.

[45]. A few project managers disapproved of the approach, according to the consultant, because they were concerned "...that the information produced might be used to undermine their own authority or jeopardize the program for which they were responsible." Salmen, 1987:101.

[46]. ibid., 1987:127.

[47]. ibid.:128. These projects represent the 'second generation' of participant-observer evaluations carried out by the World Bank.

Also the attempt to find host-country nationals with both the commitment to and the skill for conducting participant-observer evaluation proved to be relatively easy.

According to the consultant, however, some shortcomings in participant-observer evaluation do exist. Inexperience of practitioners and supervisors can lead to inadequate analysis and substantiation. In some cases this had led to findings being presented to management as raw data without any analysis. Also, the incomplete understanding by project managers of the difficulty of the execution of the evaluation and the value of the findings affected the success of the evaluation.

Salmen underlines the importance of experience and mature judgments in choosing participant-observer evaluators. There are, however, other potential short-comings in the participant-observer approach that are related to this and other aspects of the methodology.

Possibly, the most controversial aspect of participant-observer evaluation is the difficulty in identifying and controlling for observation bias. Use of triangulation[48] - ensuring reliability by means of cross-checking findings - can limit such bias, if it is inadvertent. But, bias caused

---

[48]. Lawrence Salmen lists 4 different means of triangulation: using various sources of information, using different researchers, approaching the issue from the perspective of different theories, and studying a problem with different methodologies. Salmen, 1987:120. See also Richards, 1985:110-145.

by co-optation of the evaluation poses a far more serious threat to participant-observer evaluation. If the observer has a stake in the project being evaluated, whether directly or merely from an ideological point of view, the implicit neutrality of the observer is threatened. In the case of the participant-observer approach tested by the Bank, the evaluation was closely related to the needs of management. Consistent with serving the needs of the managers the approach gave them control over the selection of the evaluators and the dissemination of findings from the evaluation. In such cases, co-optation by management leading to hiding of embarrassing or unwelcome findings may indeed take place. Likewise, the evaluator may be coopted by project participants - particularly in the case of host-country evaluators - through the offerings of bribes, etc.

If project participants perceive the evaluator as a representative of project management or the government - both of which they may perceive with distrust, fear or even hostility - research may become very difficult.[49]

In other words, the gaining of trust - a prerequisite

---

[49]. A good anecdote illustrating the problem of distrust was told to me by a person who worked on a project, operated by an agency of the Puerto Rican government. The project was aimed at isolated communities in Puerto Rico during the 1960's. The agronomist assigned to one community in the mountains of Puerto Rico to help peasants improve tobacco and livestock production, spent the better part of a year gaining the confidence of the local peasants. The reason: they thought he was actually a revenue agent sent by the government sent to shut down their illegal stills.

for participant observer evaluation - may entail huge difficulties and may not be possible at all. This problem, of course, only gets accentuated in authoritarian countries and communities.

As an explanation of the attractiveness of participant-observer evaluation to most managers Salmen cites the coincidence of the approach with a political democratization process:

> Both Bolivia and Brazil were moving from authoritarian to more open, participatory forms of government during the execution of these evaluations. Political leaders saw a congruence between this operational, human inquiry and their own efforts to fashion programs responsive to people's needs and wants.[50]

If, however, the opposite had been the case - as it is, for example, in Kenya - participant-observer evaluation may not be accepted by government officials or trusted by project participants. This type of approach may also disrupt the internal power-balance of a community - a situation that may prove most unpleasant for a participant-observer, if he or she is perceived as a threat by the local leaders. Disrupting the power structure of a community can lead to a breakdown of an evaluation if leaders feel threatened in their positions of power and try to use various forms of coercion to stop further research. This may also pose a direct threat to the safety of the evaluator.

As a design option for evaluation in a Third World

---

[50]. Salmen, 1987:104

context participant-observer evaluation has the ability to take the larger political, social and cultural context into account. It is reasonably fast and fairly cheap. It also has the ability to overcome the shortcomings of traditional sample surveys and quantitative impact studies, through an in-depth qualitative understanding of causal relations. Whether it has the ability to overcome co-optation, distrust and the danger in potentially disrupting local power structures is largely a question of the specific setting, and is left at the discretion of the evaluator.[51]

Research Paradigms Applied To Evaluation Research

The previous section discussed evaluation design and methods used for data collection and analysis. In this section, I argue that the discussion whether quantitative or qualitative methods are better is a misinterpretation of the methodological choices that have to be made to carry out evaluation research. Rather, there is a choice to be made concerning which research paradigm to apply to the inquiry.

"Between a rock and a soft place."[52]

It may appear that the choice of evaluation methodology

---

51. I shall discuss the choice of evaluator in relation to different evaluation strategies in more detail in a later section.

52. "To corrupt a contemporary metaphor, we risk being caught between a rock and a soft place. The rock is rigor, and the soft place is relevance...", Bronfenbrenner, 1977 - quoted in Guba and Lincoln, 1981:66.

is largely a choice between quantitative and qualitative methods - or a choice of rigor versus relevance - rather than a choice between different research paradigms and what follows in terms of specific methods to be applied. The World Bank presents the choice this way:

> There is a continuing discussion among evaluation practitioners as to whether quantitative or qualitative methods are better. The debate is often conducted in quite heated terms as it involves philosophical and ethical issues as well as methodology....The approach adopted in the present document is that all techniques have their strengths and weaknesses and that a multi-method approach should always be used in which quantitative and qualitative methods are combined.[53]

That is, the World Bank presents different evaluation methods - e.g. experimental design and ethnographic "techniques" - merely as a set of different techniques that each have their strengths and weaknesses.

A study of aid evaluation as practiced by member-countries of the Development Assistance Committee (DAC) of the OECD reveals a similar understanding, perceiving different research paradigms as being complementary:

> Clearly, the classical model of a scientific experimental or quasi-experimental research design is very rarely if ever feasible in evaluating donor-assisted development projects. Short of this ideal, however, donors have still failed to explore the possibilities of empirical research. They have neglected [these] opportunities:
>
> a) systematic efforts to experiment with "less-than -perfect" variants of the classical model;
>
> b) the use of innovative research methods that do not rely wholly on the "expertness" of the evaluation team

---

[53]. Bamberger and Hewitt,1986:57.

....[54]

An elaboration of this view is provided by Richard P. Nathan:

> [Q]ualitative studies provide the understanding upon which formal modelling can be done. Qualitative research should be used to provide hypotheses, lists of important variables, and tentative conclusions. Quantitative researchers can then subject it to the falsification criteria.[55]

Making choices about evaluation methodology is not, however, a matter of whether quantitative methods 'are better' than qualitative (or vice versa). Even less so is it a discussion of combining "less-than-perfect" variants of the scientific paradigm with innovative research methods - combining quantitative methods with qualitative methods - to make up for the shortcomings of the former.

Fundamental differences exist between a scientific experimental evaluation design and, for example, an ethnographic evaluation design; and rather than seeing them as part of a continuum, they must be perceived as derived from different - and in many ways antagonistic - research paradigms. In the following they shall be referred to as the scientific versus the naturalistic paradigm.[56]

The two paradigms represent different interpretations on

---

[54]. OECD, 1986:33.

[55]. Nathan, 1985:719.

[56]. Michael Patton presents the two paradigms as 'the hypothetico-deductive, natural science paradigm' versus 'the holistic-inductive, anthropological paradigm. Patton, 1980:17.

a number of basic assumptions, as stated by Guba and Lincoln:

> [T]wo paradigms for dealing with "truth" have emerged as major competitors. These are the scientific paradigm, based on logical positivist epistemology, and the naturalistic paradigm, based on a phenomenological epistemology. These two paradigms differ on a number of basic assumptions, of which three have been singled out for attention here: the nature of reality, subject-object dualism, and the nature of truth statements. When an investigator begins the study of some problem...., he should determine as well as he can which set of assumptions best fits the phenomena to be studied.[57]

I shall briefly discuss the differences in the assumptions about 'the nature of the truth statement' represented by the two paradigms. This discussion is important because it can help clarify that the choice to be made in evaluation research is a choice between paradigms and not between quantitative and qualitative methods - nor is it per se a choice between rigor and relevance. The way truth is arrived at from the perspective of the two paradigms determines, in other words, what is termed relevant and how tests of rigor can be met.

In the scientific paradigm truth is verified through experimentation. The hypotheses for establishing truth are deduced from a priori theory, and the experimental testing of the theory will establish generalizations that are widely applicable (non context-related). The scientific inquirer furthermore perceives data as being uninfluenced by interaction between inquirer and phenomenon.

This understanding of truth is contrasted by the

---

[57]. Guba and Lincoln, 1985:62.

understanding contained in the naturalistic paradigm. In this paradigm "truths" - or multiple realities - are discovered through interaction between inquirer and the phenomena. Such interaction will lead to conclusions about what is important, dynamic and pervasive in that field, rather than what has "enduring truth value"[58]

Empirical evidence provided in this paper indicates that the scientific paradigm is unsuited for inquiry in the social-behavioral sciences for a number of reasons. Scientific evaluation approaches depend on the ability to generate rigorous and reliable primarily quantitative data - data that often do not exist, especially in developing countries, and which are hard to recreate, as proved by the impact studies of the First El Salvador Sites and Services Project.

The scientific evaluation design also generally fails to take the context into account - again because of the "data limitations and insuperable statistical problems"[59] experienced when trying to measure complex situations with experimental, quantitative methods. Quantitative methods are primarily 'knowledge-verifying' and ignore information that cannot be quantified, regardless of its importance to project

---

[58]. ibid.:55.

[59]. Aiken et al., 1985:467.

processes, outcomes and impacts.[60]

The shortcomings of scientific research methods cannot, however, simply be "patched up" by methods based on a naturalistic paradigm, as suggested by the methodological discussion by the World Bank and the DAC. Instead choices must be made as to which aspects of project evaluation lend themselves better to which methods within the naturalistic paradigm.

Evaluation Approaches

The previous section was devoted to a general discussion of the different research paradigms applied to evaluation research, namely the scientific and the naturalistic paradigm, and their 'fit' for evaluation of social action programs. These two paradigms are reflected in two general approaches to evaluation: 1) the scientific paradigm tends towards a quantitative, cost-effectiveness approach; and 2) the naturalistic paradigm tends toward a qualitative, illuminative approach.

In this section I discuss the sample of housing evaluations as examples of the two different approaches to evaluation. This discussion will serve not only to explain

---

[60]. Barclay Hudson writes: "[S]cientific interpretation of value choices can only deal with techniques, the superficial expression of goals, just as the scientist looking at a great work of art - a painting or a piece of music - can only describe the technology and mechanics of producing it, not the substantive content." Hudson, 1975:82.

the general rationale behind the evaluations, but also to show how it affects the design of an evaluation and the type of methods used.

Two general evaluation approaches: cost-effectiveness and illuminative evaluation.

Howard Richards (1985) defines two general approaches to evaluation, namely cost-effectiveness evaluation - or 'the systems approach' - and illuminative evaluation. According to Richards a system is cost-effective

> ...when, compared to the available alternatives, it achieves the same objectives at a lesser cost. Or when, compared to the alternatives, for the same amount of money it achieves the objectives to a higher degree.[61]

Following this rationale evaluation has to do with measuring the degree to which the objectives of a system are achieved, and how efficient the system is in achieving the objectives.

As a challenge to this understanding of evaluation Richards uses illuminative evaluation, belonging to a family of evaluation approaches, arising from the naturalistic paradigm, which includes, for example, responsive, ethnographic, transactional and democratic evaluation approaches.[62] Malcolm Partlett defines illuminative

---

[61]. Richards, 1985:8.

[62]. Malcolm Partlett, in his foreword to Richard's book sees these approaches as belonging to the same family, but "a family that, like all families, has different degrees of closeness between its members, a natural tendency for members to regard themselves each as the family's centre, and despite overlapping concerns, is not without its own

evaluation as being 1) holistic, 2) responsive, 3) heuristic, and 4) interpretive.[63] It is holistic because it seeks "to portray the program as a working whole, as an individual organizational construction that needs to be examined simultaneously from many different perspectives."[64] It is responsive, because it "designed to interest, to inform, and to add to [the] understanding"[65] of all concerned with the program. It is heuristic, because "it is not a preordinate design, ... Rather, a study evolves, with 'leads' being followed up and new questions coming to fore."[66] And, it is interpretive, because "through tightly woven descriptions with examples, with significant facts and figures, an overall depiction of a program that does justice to the inherent complexity and which throws light (hence 'illuminative') on little-known or previously taken-for-granted features [the illuminative evaluator acts] as an interpreter - interpreting the program for those outside, or helping to interpret it for those inside it."[67]

Examples of cost-effectiveness and illuminative evaluation in the sample of housing projects.

---

disputes." Richards, 1985:xv.

63. ibid.

64. ibid.

65. ibid.

66. ibid.

67. ibid.

I will argue, that the nine housing evaluations introduced in first chapter can be understood from the perspective of these two evaluation approaches. The distribution between cost-effectiveness evaluations and illuminative evaluations in the sample is very skewed; only the evaluations from Bolivia and Ecuador (which were carried out by the same evaluator, as part of a World Bank experiment with 'innovative' evaluation methods) and the evaluation from Southern Africa use illuminative approaches. The remaining six evaluations use a cost-effectiveness approach.

Examples of illuminative evaluation.

Lawrence Salmen, the principal evaluator of the World Bank projects in Bolivia and Ecuador states explicitly in his book that his participant-observer evaluations have much in common with illuminative evaluation.[68]

The evaluator of the Southern Africa project describes his evaluation - or 'assessment' as he prefers to call it - in terms that resemble illuminative evaluation on a number of counts: The evaluation is meant as a resource for "those locally involved in upgrading"[69] by highlighting situations,

---

[68]. Salmen, 1987:122. Salmen actually uses the term "illustrative evaluation", rather than illuminative evaluation in his book. This may be a typographical error, since a footnote in the text makes reference to "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs", Malcolm Partlett and David Hamilton in Evaluation Studies Review Annual, 1976:140-157.

[69]. Southern Africa Evaluation Report:4.

posing questions and raising issues concerning the project. The evaluation had no preconceived design; the issues included are those "which arose in the discussions [with families and professionals involved in the project] and were pointed out as more critical."[70] And, the role of the researcher is defined "more as a documentor and organizer of the experiences and lessons, than as an evaluator."[71] On the other hand, there is little in the report to suggest that cost-effectiveness was a major concern in evaluating the project.[72]

Examples of cost-effectiveness evaluation.

The remaining evaluations all represent the cost-effectiveness approach. The following quote from the terms of reference for the evaluation of the First El Salvador Sites and Services Project illustrates the cost-effectiveness approach nicely:

> The Bank evaluation will seek to determine whether the objectives mentioned elsewhere in this report are being met over time, and whether project components (both physical and institutional) yield the intended social and economic impacts on project participants and on associated institutions. If they do not, evaluation should attempt to provide explanations for the unanticipated effects. It should also test assumptions concerning the project's goals, and thus provide the basis for recommendations on future project design and

---

70. ibid.

71. ibid.

72. This may be related to the fact that the implementing agency saw the project as a social service to one of the poorest communities - as a gift to the people of one project, as stated in the evaluation report.

policy for both the Government and the Bank.[73]

A look at the Table of Contents of the El Salvador evaluation and the Zambia evaluation - which are quite similar - documents the importance placed on project efficiency, which has been devoted a chapter in each report.[74] The two evaluations follow the World Bank's general definition of the primary purposes of evaluation:

> (a) to assess the extent to which the intended impacts (increases in income, reduced incidence of certain infections, improved housing quality etc.) have been produced and
>
> (b) to compare the cost-effectiveness of a project with possible alternatives.[75]

The UNDP/Habitat evaluation in Zimbabwe states as its purpose to measure, analyze and evaluate the actual direct and indirect outputs in relation to physical, fiscal/economic, social and environmental components; and to assess the replicability of the pilot projects in future national housing programs.[76] Again, there is an emphasis on measuring to which degree the objectives of the project were achieved and on replicability, which in this case is another

---

[73]. Bamberger et al., 1982a:278.

[74]. The similarity of the two evaluations should not cause any surprise. Both projects are part of a five-year evaluation study of four World Bank-financed urban shelter projects. The study was designed to assess major features of project design and to test the evaluation methods used. This explains the similarity of the evaluations. Source: Keare, 1982:i-ii.

[75]. Bamberger and Hewitt, 1985:1.

[76]. Republic of Zimbabwe, 1986:12.

way of saying efficiency. The discussion of replicability of the project is mainly concerned with efficiency, e.g. concerning infrastructure, house design, selection of beneficiaries, and project management. The UNDP/Habitat evaluation is also consistent with overall United Nations guidelines for monitoring and evaluation. In a UN source-book the purpose of evaluation is defined as providing

> Information on programme effectiveness which would provide answers to such questions as whether and to what extent the programme has achieved its objectives, and on what external conditions it depended;
>
> Information on programme efficiency which would permit determination as to whether programme results were produced in the most economical way, i.e., by maximizing outputs or minimizing inputs.[77]

This definition of evaluation is clearly in line with Howard Richards' description of the cost-effectiveness approach presented above.

Similarly, the three USAID evaluations belong to the cost-effectiveness approach, focussing on intended versus actual outputs, replicability and comparability with other housing alternatives.[78] USAID general guidelines defines the purpose of evaluation and monitoring as

> [t]o determine, quantitatively, the factors critical to full or partial achievement of the stated policy goals and project purposes[79]

Furthermore USAID'S system for evaluating housing

---

77. UN, 1978:1.

78. See for example USAID, 1982: Attachment A, pp.4-10.

79. USAID, 1978:1.

programs is meant to record "information that is critical to:

- Appropriate and rational program design;
- Efficient and economical program implementation; and
- Providing quantitative data against which to measure project success."[80]

The meaning of evaluation.

Barclay Hudson recognizes two distinctly different dictionary meanings of the verb "to evaluate", which illustrates the differences between cost-effectiveness and illuminative evaluation:

> What [to evaluate] has usually come to mean in the realm of social policy evaluation is to "appraise", to set a value on, to find the worth of, or the amount of. Essentially this type of evaluation seeks a correspondence between one thing and an externally supplied unit of measure. Qualities are reduced to categories and quantities,...In contrast, a second meaning of the term "to evaluate" is to "appreciate" something, existentially, for what it is, "to recognize gratefully, to esteem, to be fully and sensitively aware of."[81]

Whether by 'to evaluate' we mean 'to appraise' or 'to appreciate' in many ways determines our choice of design and research paradigm for evaluations.

There is also an apparent link between the institutional affiliation of the evaluator and the evaluation approach adopted, as demonstrated above. This issue will be discussed in further detail in the concluding chapter of this paper.

---

[80]. ibid.:2.

[81]. Hudson, 1975:86.

CHAPTER 3

UNDERSTANDING THE CONTEXT: A PREREQUISITE FOR EVALUATION.

I have so far reviewed the design and evaluation methodologies used in nine evaluations of low income housing projects in developing countries. I have also discussed the scientific and naturalistic research paradigms and concluded that the shortcomings of the scientific paradigm cannot be "patched up" by methods based on the naturalistic paradigm. Rather, the choice to be made lies within the naturalistic paradigm and concerns which aspects of a project evaluation lends themselves better to quantitative and qualitative methods, respectively. Finally, I have demonstrated that the two research paradigms are reflected in two general approaches to evaluation: the scientific paradigm tends towards a quantitative, cost-effectiveness approach; and the naturalistic paradigm tends towards a qualitative, illuminative approach. The nine evaluations initially reviewed were used as examples of the two approaches.

Effective evaluation is, however, more than a methodological problem. It also depends on an understanding of the real world context in which the evaluation will be carried out. In this chapter I will discuss why I believe that a proper understanding of the socio-political context of

low income housing projects is of overriding importance in evaluation research.

Origin and Assumptions of Evaluation Research Concerning the Political and Administrative Environment.

It has become clear through my research, that project evaluation most often has depended on evaluation methodologies developed in an American context.[82] The underlying assumptions about the political and administrative environment are, therefore, based on the experience of the United States. Thomas Smith characterizes the key features of the political climate in the United States as it relates to policy evaluation as follows:

> a stability in political and administrative institutions and processes with regular elections for political leadership and supremacy of the political sector over the bureaucratic
>
> democratic traditions and practices of individual and group freedom, lack of suppression, reasonably 'open' government, an active investigative media, freedom to criticize political leaders, administrators and government policies and programs
>
> a diversity of centers of power with divisions and rivalries between executive and legislative branches of government and between federal, state and local levels
>
> a slow, deliberate, incremental policy formulation process which, in some policy arenas, leads to competition between organized interests
>
> a reasonably 'satisfied' well paid, professional, non-corrupt public service which expects evaluation and review as part of the job
>
> an active and powerful political party out of power seeking to gain electoral advantage

---

[82]. Feuerstein, 1985; Smith, 1985; and Staudt, 1985.

> a lack of fundamental ideological cleavages with major political groups more pragmatic than ideological in matters of policy, and with those groups strongly supportive of existing social and political processes
>
> an array of government and non-governmental bodies which support, fund or conduct evaluations and a large number of professional and social scientists from various disciplines who teach about evaluation and conduct evaluations
>
> a large database of social and economic information readily accessible to policy researchers[83]

While these key features may not be an entirely accurate portrayal of how western political systems and bureaucracies actually function, they are assumptions which are de facto built into evaluation methodology.

However accurate these assumptions are as a description of the political and administrative system in the United States, they are wholly unsuited to the developing country context. In fact, arguably, one could prepare a list of the key features of government in most developing countries by simply inverting the features listed above: unstable political and administrative institutions, undemocratic policy making processes, centralization of power, abrupt policy changes, etc. But, rather than providing a general outline of the politico-bureaucratic context of developing countries I shall take a closer look at a specific case: the Dandora Sites and Services Project, in Nairobi, Kenya.

The Political and Bureaucratic Context of Low Income Housing in Nairobi, Kenya.

---

[83]. Smith, 1985:129-130.

During a stay in Kenya, where I studied the Dandora Site and Service Project and other low income housing projects in Nairobi[84], I became aware of the extent to which the socio-political context affects the outcome and impact of low income housing projects.

The Dandora Project began in 1975. The project was designed as a sites and services project of approximately 6,000 plots with related community facilities, including 6 primary schools, 2 health clinics, 2 community centers, 6 markets, 400 market stalls, a workshop cluster, and a small sports complex. The plots were serviced with a core unit, water supply, and sewers; the project has roads, street lighting, and garbage collection. Material loans, and assistance to allottees on technical and community development issues were provided. Finance for the project was provided by the World Bank and the Kenyan Government-contributing approximately $16 million and $14 million respectively.[85]

I will show, using a number of examples, how the political and bureaucratic environment in Kenya had a direct influence on the goals, design, implementation, and outcome of Dandora and other projects . These examples paint a

---

[84]. I spent three months in Kenya doing field work for a paper on institutional performance in low income housing. My research approach was similar to that described as illuminative evaluation in a previous section of this paper. See also Rasmussen, 1987.

[85]. Rasmussen, 1987:21.

picture of a political and administrative reality quite different from that which is implicit in many planning approaches. The failure to understand this has led to unrealistic planning and poor project design.

The examples I have selected illustrate the degree to which the political and administrative context has influenced the Dandora project's development and, therefore, how important it is to include an analysis of that context in any evaluation. I have grouped the examples under 5 main headings: 1) Access to the planning process ; 2) donor influence and conflicts within the power structure; 3) patronage politics and corruption; 4)bureaucratic constraints; and, 5) availability of resources.

Access to the planning process.

When Dandora was designed in the 1970's little effort was made to solicit input from the urban poor or from potential allottees. As a result project participants' only option was to participate in the planning process during the implementation phase in an ad hoc manner.[86] A case in point is what was planned for Dandora's commercial and industrial sector and what actually happened.

As envisioned by the project planners, who hoped to control the growth of the informal economy, commercial and industrial activities were to be licensed businesses located in clusters which were to form a "spinal core" in the

---

[86] Grindle (1981) develops this point nicely.

community. Market stalls were built to house the businesses. As a visit to Dandora makes clear, this arrangement has been a failure. Businesses have not been confined to the "spinal core" where it has been difficult to rent the market stalls (compare the K.Sh.300 monthly rent with the K.Sh.10 fee for a street vendors license.) Instead, a large, informal economy has developed throughout Dandora. Many plots have been turned into all-manner of business: shops, tea kiosks, bars, lodging houses, dry cleaners, etc. Unless the structure itself violates codes, the Housing Development Department[87] seems to have acquiesced in the growth of trade in areas of the project not originally planned as commercial areas.

It is by their deeds, then, that the residents of Dandora have had a voice in the shaping of their community and not because they have been given a voice in the planning process. The result is that an important part of Dandora's evolution as a community has been both unregulated and unplanned.

Donor influence on policy and conflict within the power structure.

In many developing countries, where financial resources are inadequate to meet development needs, large donor institutions can have a major impact on development policy. This is especially true because it is more often than not the donor which approaches the recipient government with plans

---

[87] The Housing Development Department is responsible for the project's implementation and management.

for a project and not the government which approaches the donor.[88] By taking the initiative on project design and by controlling the funds, a donor agency is in a good position to tag conditions on to any agreement it makes with an aid recipient. In practice, however, an agreement on conditions between a donor and a government may not have the desired effect if agencies within the government's administrative bureaucracy either ignore or interpret the conditions to suit their own outlooks and agendas. The history of the Dandora project provides a clear example of this process.

The widespread implementation of minimum sites and services projects, of which Dandora was one of the first, represented a break with past planning practice in Kenya. Previously, the Government- lead by President Jomo Kenyatta- had vigorously opposed implementation of low standard housing projects. In 1971 Kenyatta warned that he would fire any official who approved the construction of what he termed "native-type" houses.[89]

---

[88]. Tendler, 1975.

[89]. The quote is from a report cited by Temple and Temple (1980:236): "President Kenyatta warned at the weekend he will sack any of his Ministers who continue to treat Kenyans as the colonialist did by building them "native-type" houses not suitable for human habitation."
An alternative view to the reason for Kenyatta's opposition to sites and services projects is offered by Stren (1982:93): "The reasoning often heard was that, if building and sanitation standards were allowed to fall in Nairobi, tourism would suffer and international firms would hesitate to locate in Kenya."

By 1975 , when the agreement was signed with the World Bank for 6,000 minimum standard site and service projects at Dandora, the government had had a change of heart. There seems to have been two reasons for the policy shift. First, a recognition of the sheer magnitude of the housing shortage caused by a rapidly increasing population[90]. Second, was the influence of the World Bank. I believe it is safe to say that had it not been for the World Bank Kenya would not have seen large scale sites and services projects until much later - if ever.

But the overall acceptance of a low income housing policy based on the self-help sites and services concept did not mean, however, that all aspects of the policy were accepted by all levels of government. The dislike for lowering the standards persisted in a number of government agencies creating serious roadblocks to the project's implementation.

At the insistence of the World Bank the more relaxed Grade II by-laws of the Building Code were adopted for Phase I of the Dandora project (1,024 plots), instead of the more stringent Grade I by-laws. This was accomplished by re-zoning the area through a resolution of the Works and Town Planning Committee in January 1975. In May the following year - after plot allocation had started and construction was well

---

[90]. The average urban growth rate for Nairobi for 1969-79 was 5.3%, which meant adding more than 300,000 new residents of the city. Source: Rasmussen, 1987:11-14.

underway - law enforcement officers, especially from the Public Health Department of the City Council, refused to approve the plans since they did not conform with the Building Code and the Public Health Act. Lowering the standards, these officials claimed, would result in health hazards and unhygienic, uncontrolled human settlements. As the Dandora Project Manager put it: "This was therefore a problem of the interpretation of the law."[91] .

It took almost 6 months of negotiations among the World Bank, the City Council, and the Government to resolve the issue.

Since it was too late to make amendments to Phase I the new standards were accepted, but phase II experienced eleven major design changes, causing serious delays and increasing cost of construction due to inflation and price escalation. The controversy also resulted in delays in issuing house type plans for allottees and in the construction of demonstration houses. The additional cost incurred from adopting higher standards was estimated to K.Sh. 27.6 million.[92]

The compromise did not resolve the issue of standards for low income housing in Kenya, but was instead a political compromise, among various agencies, to save the Dandora project.[93]

---

[91]. Wanjohi and Chana, 1977:18

[92]. ibid.

[93]. ibid.

Vigorous opposition continued against efforts by the main agencies of the central government and the World Bank to impose the new, lower standards. The opposition, in Nairobi in particular, had powerful supporters, among which were the National Housing Corporation (NHC) - a para-statal agency under the Housing Ministry and the intermediary lending agency for government housing projects - and the Director of City Planning and Architecture in Nairobi.[94] The Director of City Planning and Architecture held particularly strong views on low standard housing, as this excerpt from the minutes of a Housing Committee meeting in January 1986 indicates:

> The Director of City Planning and Architecture reported that although the objective of the site and service projects was to stabilize urban population in the City, it had not done much to solve the problem of low income housing in the City. There was, therefore, a need to review the whole concept of tenant purchase schemes versus rental schemes. The Department's views were that the [Nairobi City] Commission should now concentrate more on rental housing rather than site and services schemes or tenant purchase scheme.[95]

---

[94]. It has been said about the NHC: "The National Housing Corporation which is a front runner in the site and services schemes programmes in Kenya is most allergic to alternative technologies."; the comment was made in a recent paper from the Housing Research and Development Unit (n.d.:17) - a government funded research institute, which devotes large amounts of its research to low standard design and construction techniques particularly for site and services projects.

[95]. Nairobi City Commission, 1986. One can only speculate why the Director would express such sentiments in light of the very poor performance of rental housing in providing affordable housing for low income groups. Furthermore, a widespread belief is held that the problems experienced with sites and services projects are not so much related to the concept of sites and services, but rather to

There are efforts on the national level to resolve the problem of housing standards. Most recently, in 1987, the Ministry of Works, Housing, and Physical Planning submitted a policy paper to Parliament recommending that its policy of lowering housing standards be accepted as a sessional paper, thus making it a legal binding national policy.

The issue of standards provides a clear example of government policies and project agreements being severely altered to the detriment of both the over-all project (because of cost increases to the funders) and the project beneficiaries (because of project delays and more costly construction standards). It also had another effect: with the increased cost of construction because of the raised standard of Phase II , many of the original allottees sold their plots, because construction became too expensive. Thus instead of the intended target group, many relatively well-off people entered the project.[96] The end result of the

---

the insistence on literal self-help (thus providing loans inadequate for the sub-contract form of self-help), serious construction delays (mainly due to corrupt government practices), etc. My guess is that it has very little to do with concern for the urban poor - and a lot to do with the politics of power. Sites and services and upgrading projects are the responsibilities of the Housing Development Department, whereas rental housing falls under the auspices of the City Planning and Architecture Department. The rivalry between the two departments has been known for a long time.

[96]. And, "Furthermore, with the increasing cost of construction due to the various delays and inflation ..., it gave an opportunity to the income groups above the specified target income groups to

inter-agency fight over housing standards had the important consequence of re-defining who would live in Dandora.

Institutional aspects of low income housing in Nairobi.

The World Bank also put a lot of emphasis on the institutional aspects of low income housing in Nairobi. In accordance with the trend in development administration, during the 1970´s, the Bank helped create a new department to implement Dandora, as well as a matching City Council committee. The Dandora committee was later to be expanded to include all low income housing in Nairobi, this occurred in 1978 when it became the Housing Development Department.[97]

The Dandora department was originally vested with plenary powers on matters of staff appointments, bidding procedures, and certain expenditure. The granting of plenary powers was intended to provide the department with relative autonomy and thus safeguard it from many of the problems of a highly centralized and politicized government bureaucracy.[98] Although the World Bank managed to put pressure on the government when negotiating the project, within six months the decision granting the new committee plenary powers was

---

occupy the plots,". Chana, 1984:32.

97. The scope of the Housing Development Department´s activities are much more limited than originally proposed by the plan.

98. Vesting the department with plenary powers does not guarantee that mismanagement and corruption will not be factors. By centralizing more of the decision-making in the department, the Bank´s role as a watchdog is, however, simplified.

revoked by the Nairobi City Council. Internal struggles for power among city councilors and officers led to the dismantling of the new housing department's relative autonomy in certain areas of decision-making.

Land for low income housing.

The Dandora Site and Service Project is not the only example of plans running aground on bureaucratic realities. The history of squatter upgrading projects in Nairobi is another case in point. In fact, Nairobi has no significant upgrading projects; out of a total of 8,750 units constituting the Planned Physical Housing Output 1983/84 - 1987/88 only 200 of those units were upgrading - or a little more than 2%. This compares very unfavorably with other cities: in Mombasa 34% of output was upgrading, in Kisumu 67%, with a national average of 22%.[99] These statistics seem even more puzzling in light of the fact that 30% of Nairobi's population - by a conservative estimate - live in squatter settlements, and since the government strategies (as stated in the Development Plan 1984-88) are to produce low income housing through sites and services programs and settlement upgrading.[100]

Nairobi has not lacked initiatives on squatter upgrading

---

[99]. Republic of Kenya, 1984:168

[100]. ibid.:165.

projects. The Second Urban Project[101] funded by the World Bank originally included several sites for upgrading. During a cut-back in 1981 the upgrading components of the project were, however, omitted. These cut-backs were begun in 1980 when revised estimates, projected to the anticipated date of construction, showed a considerable increase over original projections made at the time of the project's proposal. Since the World Bank funds were limited by the project agreement and the Kenyan government would not commit itself to providing additional funds the whole project was revised.

Significantly, the squatter upgrading projects - all of which were deleted from the Second Urban Project in Nairobi during the revision - suffered major problems with land acquisition. In a report submitted by the Director of Housing Development in 1986, identifying housing projects for the Urban IV study in Nairobi these upgrading projects are brought up again:

> These projects[102] involve upgrading of settled areas by provision of infrastructure and sanitation which is generally lacking in all the above areas. <u>One of the major problems in the above areas has been the land acquisition and provision for this should be made during the financing.</u>[103] (emphasis added)

---

[101]. The Dandora project is also known under the name the First Urban Project. The Second Urban Project included sites and services and squatter upgrading projects in Nairobi and other major cities in Kenya.

[102]. Baba Dogo, Mathare Valley, and Riruta/Kawangware upgrading projects.

[103]. Nairobi City Commission, 1986:1148.

Up until this time all but one of the low income housing projects in Nairobi have been implemented on government owned land[104], and the problem of land acquisition has thus not been an issue. The squatter upgrading projects are, however, a different story.

The owners of land in squatter areas usually find very little incentive to sell their land to the government for purposes of squatter upgrading projects. As areal photos and visits to such sites easily reveal, many of Nairobi´s squatter settlements are not incremental developments with dwellings haphazardly constructed by the individual squatter on an ad-hoc basis. Large tracts of land are covered with the so-called ´company houses´ - houses constructed by private land owners and developers. Although the quality of company housing vary greatly, this housing does not comply with government building codes and the Public Health Act. Furthermore, in a tight housing market very exploitative rents are likely to prevail. The value of land owned privately in the squatter settlements is thus very high, because it generates an income while the landowner is waiting for an opportune moment to either sell or develop the land. In addition, in the long term the potential market value of the land compared to the current price for land offered by

---

[104]. Mathare Valley North Sites and Services Scheme was implemented on privately owned land acquired by the government. Land acquisition proved to be a major issue holding up the project and resulting in delays and consequent cutting back of the project.

the government may indeed provide little incentive for landowners to sell. In Kenya, the government has made little effort to acquire the land through legal processes, and it seems likely that landowners and developers of company housing benefit from contacts within the government.

This is another example of how the planning agenda has been set, more or less, by a donor agency (the World Bank) without a clear understanding of the political context. Whether the Kenyan government was aware of the land acquisition problems prior to signing the project agreement - perhaps hoping the funds could be diverted to other use, should the policy flounder - is hard to say. More likely, it is yet another example of dissonance within government - strongly encouraged by a close political relationship between landowners and politicians, among whom there is often overlap.

Patronage politics and targeting the poor.

Patronage politics has influenced the development of low income housing in Nairobi in several ways. Tampering with plot allocation and bidding procedures has been notorious; rules concerning the collection of arrears, eviction of allottees and demolition of illegal structures have not been enforced; and, the operation of stone quarries near Kayole and Umoja projects[105], which threatens both the physical structures and the subterrainean infrastructure, continued

---

[105]. A World Bank and a USAID low income housing project.

several years after an order from the City Council ordered the quarries closed. I shall focus my remarks about patronage politics to those aspects which had a negative impact on the original objective of targeting the poor.

Instead of a city government based on some notion of impartiality and efficiency, the City of Nairobi has experienced times when patronage politics and corruption have run rampant within the City Council. In 1983 the problem became so acute that the Minister of Local Government in the national government replaced all politically elected councilors with appointed commissioners.[106] In 1987 the Minister summarily fired the Town Clerk and most directors of departments.

The patronage politics and corruption at City Hall had direct effect on plot allocation for Dandora Phase II when, according to Father George MacInnes[107] who has lived in the project a number of years, 700 plots were sold cheaply to "friends" of Nairobi city councilors. By displacing 700 allottees belonging to the original target group, the Council's action had the practical effect of re-drawing the original target group by introducing a more affluent group of people into Dandora.

In fact, there are several reasons why Dandora became

---

[106]. New elections were scheduled for 1988, but have not been held to date.

[107]. Interview February, 1987. Several other people, with whom I spoke, confirmed Father MacInnes' account.

attractive to people with incomes well above the target group. First, national housing policy has not adequately addressed the problem of urban middle income groups, thus leading to a spill-over into places like Dandora in a way similar to the gentrification process in cities of the developed countries. Secondly, middle-income rents have doubled in Nairobi between 1980-85.[108] And thirdly, projects like Dandora represent the quickest return on real-estate development available at the low end of the scale. In middle-income developments, the cost of land represents 25% of development cost - a cost omitted in Dandora.

The introduction of more affluent people into Dandora had other consequences as well. This group had greater resources with which to set up businesses in Dandora such as stores and lodging-houses. This put a competitive squeeze on efforts by the poorer allottees to establish businesses. Since over half of the poor allottees in Phase II came from the Mathare Valley squatter areas, where they had made their income in the informal sector often in the provisioning trade or by subletting rooms, the increased competition directly affected their income and their ability to pay project charges and fees.

Patronage politics have also affected the collection of arrears. It has been common practice that eviction notices served to allottees in rent arrears of six months or over did

---

[108]. International Labour Office,1986.

not lead to evictions. Instead through "various Committee resolutions"[109] fresh eviction notices were served to the allottees. While such practices may be justifiable in many cases, the arrears have continued to accumulate and had reached an average of K.Sh. 2005 per plot in Dandora in June 1986 or a total of more than K.Sh.11 million, thus affecting the viability of the project.[110]

Enforcement of rules on building size is another case where patronage politics has altered original plans. Dandora phase II's 96 multi-story buildings stand as landmarks to the failure of the authorities to enforce the project's ban on multi-story buildings. This type of corruption finds its way down to the building inspectors who fail to report illegal structures.

While this type of patronage politics and the petty corruption which accompany it have had a detrimental effect on the project as a whole, its real victims have been the poor who have neither the money nor the influence to participate in the patron-client game.

Bureaucratic Constraints

The Housing Development Department in Nairobi, the direct descendant of the Dandora Project Committee, has responsibility for most low-income housing development in

---

[109]. Housing Development Department, 1981:39-40.

[110]. Interview with accountant in the Finance Section of the Housing Development Department.

Nairobi. Three things characterize the Housing Development Department's contact with its bureaucratic environment: there is a high degree of control and power centralized in the national government; the bureaucratic procedures, which have roots in the British colonial period, are very cumbersome; and coordination and cooperation between the Housing Development Department and other parts of the bureaucracy seems limited.

The centralization of power in the bureaucracy often affects the Housing Development Department in a very direct fashion. For instance, the bidding procedure prescribes that approval of the Minister of Local Government be given, if the lowest bidder is not awarded the contract.[111] This is taken quite literally, and in one case it delayed the awarding of a contract for several months when the Minister was hospitalized and not available to personally give his approval. The Deputy Minister was not empowered to act in his place.

This high degree of centralization has had its effect on Dandora's development, often in ways that may not have been apparent to decision makers in the central government. A case in point is a hiring freeze for manual workers ordered by the Ministry allegedly to curb excessive hiring of manual

---

[111]. Government of Kenya, 1978:79. The anecdote about the delay in the contract award: interview with Assistant Director, Housing Development Department, February, 1987.

workers for patronage reasons. The hiring freeze was ordered across the board without regard for job categories. This meant that the Housing Development Department was unable to hire watchmen to guard construction sites in Dandora. In Kenya, where there is a high incidence of theft from construction sites, this led to cost increase from lost building materials and machinery, from delays in construction due to lack of materials, and eventually from the hiring of private watchmen at higher costs.

The bureaucratic procedures at the local government level in Kenya can only be described as cumbersome. The decision-making process at the municipal level often involves a Chief Officers Meeting, a meeting of the relevant standing committee (i.e. Housing Development Committee), and a Finance Committee Meeting. It is common practice that matters are deferred from committee to committee, and from meeting to meeting. Even very serious issues get deferred, causing delays in decision-making of up to one month or more, until the next committee meeting. Often, further consultations or reports to be prepared between meetings are not completed. The Chairman of the Housing Development Department called this problem, "a kind of disease."[112]

The committee meeting system as conducted by Nairobi City Commission requires the presence of the director of HDD

---

[112]. Minutes of the meeting of the Housing Development Committee, January 1987.

or his representative at any other committee meeting where issues related to housing development are discussed. While in principle this is desirable, the presence of the director is required in a large number of meetings, taking time away from his principle role as the department manager.

Cooperation and coordination with other municipal departments is necessitated by the design of the department, which makes it dependent, for many services, on other departments such as the City Engineer´s Department (CED), the Water and Sewerage Department (WSD), the Social Services and Housing Department (SSHD) and the City Planning and Architecture Department (CPAD). The HDD´s relationship with the other departments is of varying quality. In general, relations to all departments but the WSD, which has a permanent representative assigned for housing projects, are seen as problematic by the HDD staff, because of interference, lack of flexibility or for slowness in completing work. The relation to the CPAD has been characterized as particularly bad. While the CED has been accused of delaying projects, the CPAD is described as notoriously slow in approving plans, as having approved illegal structures, as failing to provide staff requested for specific purposes (i.e. surveying), and as being generally obstructionist.

Lack of resources.

Physical resources, financial resources and human

resources have all been scarce at different times in the Dandora project - Kenya suffered a temporary crisis in the availability of building materials leading to vast price increases; during a time of general revenue-crisis the Kenyan Government withheld disbursements for the project; and understaffing and the hiring of underqualified personnel has seriously impacted the performance of the implementing agency. As a result of this scarcity of resources there have been delays - measured in years - of the implementation of low income housing in Nairobi.

Financial constraints have also affected cost recovery. The City of Nairobi has, since the abolition of Graduated Personal Tax by central government in 1973, practically been without capacity to generate revenue. Thus, the City depends mostly on grants from central government, since the system for obtaining foreign capital is also strictly controlled by the Treasury.[113] Lack of funds obviously means that the provision of services will be poor. In a study of women in the Dandora project Nimpuno describes how the dissatisfaction with municipal services affected some women´s willingness to pay:

> [A category] of defaulters were those who did not pay because of dissatisfaction with the level of services from the municipality. The failure of infrastructural maintenance by the municipality has had serious implications for the area as a whole. On one hand there is a severe drop in standards due to failures, e.g. drains are blocked by refuse dumping, leaking water

---

[113]. Akivaga et al.,1985:4.

> mains are unreported and there is an absence in refuse collection. On the other hand the problem gets compounded by the negative and indifferent attitude of the residents.[114]

In this way resource-shortages find their causes outside the immediate context of the project: supply-side failures in the construction industry, affecting availability of building materials; central and local government general revenue crises, affecting disbursements and public services; and the mismatch between demand for highly educated technical staff, the educational output, affecting staffing of the implementing agency, and the wage structure of the public sector.

Conclusion on the low income housing experience in Nairobi.

The extended discussion of the low income housing experience in Nairobi in this chapter has underlined a number of the kinds of problems one might find in a Third World context. I have portrayed the politico-bureaucratic environment as partisan and inefficient, affecting both the implementation process, the project outcome, and the impact on the original target-group.

In Kenya there is even additional reason to take the political context into account when doing evaluation research. The centralization of power at the top, along with an almost total intolerance of dissidence, has increased drastically in Kenya in the last few years, dictated by

---

[114]. Nimpuno, 1986:24.

President Moi's government. The increased control and at times severe punishment of dissidence creates an increasingly difficult environment for evaluation and poses serious ethical questions. The government may not want an evaluation , if the evaluation points to problems in the politico-bureaucratic environment. The more useful forms of evaluation for policy formulators and program implementors which provide explanations for why policies fail by considering both policy feasibility and execution, are likely to encounter resistance.[115] Alternatively, evaluation strategies may be requested that ignore the political context, such as cost-benefit analysis and other types of analysis not concerned with explaining why certain outcomes occur. Finally, of particular importance to the evaluation from the perspective of ethics, evaluation findings may be misutilized or abused by an authoritarian government to expose and penalize opposition. Or, if indeed the evaluator is perceived as a representative of an unpopular government and its officials, the data-collection process may be severely impeded.

There may seem little hope for effective evaluation in much of the Third World since the most useful forms of evaluation may be perceived as a threat by authoritarian regimes. But, dissonance at different levels of government and the presence of different actors in development

---

[115]. Smith,1985:142.

assistance may constitute an opening for evaluation.

Gender aspects: Women's participation in housing projects.

In the past very limited knowledge has been generated about gender aspects of mainstream development projects, including most sites and services projects. Gender aspects have generally not been included in project or evaluation designs, and little is thus known about whether and how women participate in such projects, and about the impact a project may have on women in particular. This reflects a general lack of concern for women's special role on the part of government, donor agencies and evaluators - and in the society as a whole.[116] Even where baseline data are disaggregated by sex a remarkable absence of conclusions about a given project's impact on women is often found.

Notably, all the evaluations reviewed in this paper but one[117] fail to make conclusions concerning women's participation in the projects. Lawrence Salmen notes that the exposure, he experienced through his participant-observer evaluation, has resulted in a personal awareness that "both women and youths were sources of energy for development that were largely untapped by [the] projects."[118] He further notes that women spend more time in the neighborhood, have a

---

[116]. This trend does, however, appear to be changing; see for example OECD, 1988:8.

[117]. The UNDP evaluation of Kwekwe and Gutu projects in Zimbabwe excluded.

[118]. Salmen, 1987:82.

greater awareness of the relationship between health and sanitation, are more burdened by the lack of good infrastructure, and are often more responsible than many of the men. Salmen, however, also fails to provide a detailed understanding, backed by factual information, about the gender aspects of the projects he evaluated.

The lack of gender related findings in evaluations of low income housing projects is unfortunate since female headed households, as discussed earlier, are common in many low income areas. And, furthermore, the existing evidence - however limited - reveals important differences between women´s and men´s experience with, for example, self help housing projects.[119]

To form an understanding of women´s role in low income housing it is important to understand the social context, i.e. the role of women in society in general. In Kenyan society, for example, the role of women stands in stark contrast to the perceived reality of the decision makers responsible for the assumptions behind and the design of sites and services projects in Kenya. Again, I shall use data from the Dandora Project in Nairobi to illustrate my point.

Two of the requirements that had to be met to qualify for a plot in Dandora were that the applicant had to be the head of the household and could not own property elsewhere in Nairobi. This last requirement also counted for the

---

[119]. Nimpuno, 1986.

applicant's spouse. In Kenyan society, however, polygamy is still common. This could have led to the exclusion of women in polygamous marriages, even if they had their own household - which is the norm in polygamous marriages in Kenya - and were the only source of income of that household. If their husbands owned property in Nairobi, they would face exclusion from the project.[120] This serves as a reminder of the importance of possible ways the project design could have been biased against women and, thus, the importance of extending the evaluation to those who were discouraged from applying or rejected by the project after application.

Another example is the tradition in Kenya which makes it difficult for women to inherit land.[121] The access to owning a plot in Dandora, i.e. legally getting a title to land along with better housing conditions is a significantly bigger asset to women than to men.

In line with the previous discussion about inquiry into social action programs I shall not attempt to provide an exhaustive list of aspects of women's participation in sites and services projects. Some themes for an evaluation can, however, be provided. Kathleen Staudt argues, based on her evaluations of women in three mainstream development projects in the Caribbean[122], that the minimum information required in

---

120. ibid.:9.

121. ibid.:5.

122. Staudt, 1985.

any women-sensitive evaluation is:

- a description of women's and men's labor, incomes and responsibilities, and how these have changed in the household and the community
- analyses of special needs and constraints on women and men and how these affect participation
- sex-disaggregated data on all outcomes and impacts
- analyses of women's and men's participation in project decision-making
- conclusions about changes, absolute and relative at the household level and at the community level
- interpretations about why changes occur

Paula Nimpuno's study of gender issues in project planning and implementation in Dandora[123], from which the examples in this section come, provides valuable insight into which aspects of a sites and services project may be of particular importance to women. Nimpuno' study compares the assumptions behind the design of the project with the reality of the background of women allottees. Major discrepancies were found between the assumptions of the project design and the reality of women's ability to meet the initial requirements for eligibility to the project; their ability to make regular payments and raise money for the down-payment; and their ability to engage in true self help construction.[124] Based on such findings the project would

---

[123]. Nimpuno, 1986.

[124]. Women as well as men chose the contracting form of self help. For women, however, this had greater financial implications than for men. The contracting form of self help increases the cost of

largely discourage women from applying and also cause a lot of women to drop out. Through her study it became clear, however, that women were actually included in the project due to "gender planning purely at the level of the implementing agency."[125] In this sense, some of the bias against women in the project design was eliminated by project management. Furthermore, as described previously, despite the fact that many allottees fall into arrears evictions have been rare. It appears that project management has accepted that many women are not able to pay on a monthly basis, but that they depend on income from gifts and transfers from family and friends or income from crop harvest.

Nimpuno also lists reasons why women heads of households - despite their initial problems - are the most stable group in the Dandora community:

- the project gives women the opportunity to own a plot in an urban area which was not possible before;
- although both women and men tend to consider the house an economic asset, women also see the house and plot in relation to their families and emphasize therefore the social value of the house and plot;
- the possibility of subletting rooms is very important for many women as it is often their main source of

---

construction, and women - a large percentage being self employed - had much greater difficulties in raising the extra money for construction than many men, many of which had jobs and could obtain loans from employers or cooperatives. Furthermore, women being unfamiliar with the role of supervisor of a construction site were often cheated by the contractors, which again resulted in increased costs.

[125]. ibid.:25.

income.

Conclusion on gender aspects.

Through my discussion of women´s participation in Dandora I have pointed out significant differences in the experience of women and men in low income housing projects. Such differences are undoubtedly common. From this perspective it is a serious omission that gender issues rarely enter into evaluation studies. For evaluation to be effective it must also be gender sensitive.

The general lack of information about gender issues related to low income housing projects has significance for how an evaluation of such issues could be approached. If pre-project data is generally limited, then sex-disaggregated data is, to be sure, even more limited. It is also more difficult to establish initial working hypotheses for a gender-specific inquiry, since little previous knowledge exists on gender issues in relation to low income housing. Clearly, an evaluation of gender issues must lean heavily on ´knowledge discovering´ methods: participant-observation, in-depth interviews, case studies, etc. - of course supported by quantitative interpretations whenever feasible - to generate an understanding of what is significant to the experience of women in this type of projects.

CHAPTER 4

INSTITUTIONAL ASPECTS OF EVALUATION AND CHOICE OF EVALUATOR.

The final aspect of my discussion of evaluation in a Third World context concerns the choice of evaluator and - as discussed in chapter 2 - the apparent link between the institutional affiliation of the evaluator and the evaluation approach adopted.

The discussion of the choice of evaluator can be approached from two perspectives. First, there is the individual skills perspective, which is concerned with the match between the professional and personal skills of the evaluator (or the evaluation team) and the nature of the evaluation, including choice of methodology. Second, there is the institutional perspective, which is concerned with the type of institutional setting conducive to the goals and methods of the evaluation. I shall discuss both perspectives.

The Skills Perspective.

The role of the evaluator - and hence who can fulfill this role - is directly dependent on choice of evaluation method. For instance, participatory evaluation practiced by Oxfam America and other agencies[126], in which the role of the

---

[126]. See Appendix I for a detailed description of participatory evaluation.

evaluator is that of a facilitator[127], may require skills in the areas of participatory learning and teaching, negotiation, qualitative methods and basic quantitative data analysis. On the other hand, evaluation using experimental design and cost benefit analysis, in which the role of the evaluator is that of an expert, requires skills in quantitative analysis. This places some initial conditions on the choice of evaluator or evaluation team. I have in a previous chapter argued for the application of a naturalistic evaluation paradigm, making choices within that paradigm concerning the use of quantitative and qualitative methods for evaluating different aspects of a project. From this argument it follows that both quantitative and qualitative skills will be needed. The evaluation of water and sewerage systems provides an example of the need for both qualitative and quantitative skills, since it involves both project participants' subjective satisfaction and priorities for expenditure, as well as the system's affordability and physical capacity.

I will argue that there is a need to engage more people with an understanding of the socio-political context in evaluation. I shall not, however, embark upon a discussion here about the qualifications and qualities of the different professions engaged in evaluation.

Status of affiliation of the evaluator is one of the

---

[127]. See for example Feuerstein, 1985.

aspects in the debate over the choice of evaluator that has received attention. The debate has often been phrased as dichotomies, i.e. insider versus outsider, or expatriate versus local evaluator. The debate has concerned the evaluator's ability to produce unbiased, insightful findings that also are of practical relevance to the immediate users of the evaluation.

First, the insider versus the outsider. An outsider is usually defined as somebody external to the project team as well as to the agency responsible for the evaluation. Outsiders can bring a breadth of experience to the evaluations and limit the bias of self-evaluation. As the Chairman of ODA's[128] Projects and Evaluation Committee has pointed out:

> It would ... look rather odd to outsiders viewing our evaluation system as a system if it seemed to consist almost entirely of people inside evaluating the organisation from within.[129]

A more practical reason for choosing an outsider lies in the small evaluation units of many agencies, representing limited numbers of staff and different skills.

The outsider's lack of understanding of the information needs of the users of the evaluation is often cited as a

---

[128]. ODA is the British bilateral development assistance organization.

[129]. Browning, 1984:137. Other donors do not share this view. They have found that insiders with no direct association to a project can bring the same objectivity to an evaluation as an outsider. OECD, 1986:46.

problem. The outsider does not understand the internal dynamics of the organization using the evaluation and thus has difficulties in making appropriate policy recommendations, a demand many evaluators are faced with. I believe that adequate briefing of the evaluator about the user's needs can significantly enhance the evaluator's ability to produce useful findings. Whether the evaluator should be charged with the responsibility of making specific policy recommendations is another issue. A solution to this problem, which is practiced by an increasing number of donors[130], is participation by an in-house member of staff on the evaluation team.

The debate over the expatriate versus the local evaluator has generally been over objectivity versus insight in the context of a given project. Again, there are advantages as well as drawbacks to each type of evaluator. The expatriate may indeed be more objective and, due to previous experience from similar projects elsewhere, contribute a broader understanding of both project context and evaluation methodology to the evaluation. The expatriate evaluator, who may be more independent of any local power structures, may be able to establish a more open relationship with the project participants and avoid cooptation. On the other hand, there is a danger that the expatriate may fail to understand the local context, which can lead to a failure to

---

[130]. OECD, 1986:47.

grasp the reasons for certain types of social behavior. As a result evaluation findings may have little relevance. A bigger problem, however, is cooptation of the expatriate evaluator by the employing agency.

A shortage of local evaluators may affect the choice of evaluator, as might the degree of personal autonomy and the practice of patronage politics in the local context. It would be unwise, however, to opt for an expatriate evaluator based solely on the risk of cooptation of the local evaluator. Lawrence Salmen's experience with participant-observer evaluation by locals, as described in chapter two, suggests that local evaluators can indeed be impartial.

As should be clear at this point no one evaluator can easily fulfill the different requirements outlined in this section. A team seems the obvious solution for how to combine different skills, breadth of experience and local understanding, and at the same time avoid serious biases. But, what may be ideal, may not be possible. As with any other activity in development assistance the available funds will determine what is possible. Within this framework the choice of evaluator must be made. The institutional perspective, presented in the following section, offers an additional perspective on the choice of evaluator.

The Institutional Perspective.

The choice of evaluator depends on the role the evaluator is expected to play in the evaluation. This role is

defined by the characteristics of the institutional arrangement for the evaluation. Two such roles have already been described, e.g. the facilitator and the expert. In this section I shall discuss the different roles of the evaluator in relation to the involvement of donor and host country institutions in evaluation of aid projects.

Michael Patton provides a definition of three possible role models for evaluators with different "missions":

> The first style is the "surveillance and compliance" approach of aggressively independent and highly critical auditors committed to getting the goods on a program. In this type of evaluation, the evaluator is the Lone Ranger and program personnel are potential and suspected outlaws. The evaluator is on a mission of law enforcement. The second style of evaluation is epitomized by the aloof and value-free scientist who focuses single mindedly on acquiring technically impeccable data. The objective scientist is on a mission seeking truth. Program personnel are research subjects to be labeled and studied in accordance with the rules of science. It is not so much a matter of the white hats against the black hats as it is of the white laboratory coats against ordinary street clothes. The third style is where the evaluator works in a consultative, consensus-building process to help policy-makers and program personnel cooperatively and openly clarify their information needs and use information to improve their effectiveness. In this approach, evaluators and information users are partners in the search for useful information.[131]

There is evidence that evaluators who carry out evaluations of donor funded development projects in the Third World are often perceived - by project management and involved government agencies - as assuming the "surveillance and compliance" role. Often recipients of aid regard

---

[131]. Patton, 1982:58. This set of definitions is based on Barkdoll, 1980.

evaluation as an auditing or fault-finding process and may fear that evaluations will lead to decreased aid, if the evaluation come up with negative findings. This negative image of evaluation among aid recipients is one of the reasons why the donor agencies have dominated evaluation of development projects.[132] Project participants may also perceive the evaluator as having a similar role. As pointed out in a previous chapter, participants may not want to provide information on sensitive issues to the evaluator out of fear that it may be used punitively against them.

The resentment of the second role, the evaluator in the laboratory coat, probably has less merit at the level of project management and government agencies. Rather, potential users of evaluation findings may respond to this evaluator with indifference. Evidence from a number of studies[133] suggests that knowledge produced by the scientist-evaluator

---

[132]. Other reasons explaining why aid recipients typically have not invested resources in evaluation of donor-funded development projects are 1) the different priorities of donors and recipients for evaluation; 2) the lack of influence of the recipient on design as well as implementation; 3) the poor performance of some evaluation missions undermining the trust between recipient and donor, e.g. drastic changes of aid policies based on superficial evaluations or reports that allocate blame to governments and individuals; 4) the use of evaluations as an instrument to terminate policies or projects; and the lack of human and financial resources. OECD, 1986:49.

[133]. These studies were based on the American experience. I would argue, however, that such observations would be at least as likely in a Third World context less familiar with the scientific approach to decision-making.

may have less potential for utilization in decision-making processes, regardless of its merit:

> [A]s Caplan (1977) indicates, decision-makers appear more likely to make decisions based on "soft" knowledge considerations of the social impact of a particular policy rather than on "hard" knowledge..."hard" knowledge was defined as "research based, quantitative, and couched in a scientific language".[134]

Specifically, representatives from a host of developing countries have expressed the need for a "common-sense approach" and the view that evaluators should assume the role of "activists" rather than "theoreticians".[135] On the other hand, the scientist-evaluator may be welcomed, since he or she tends to disregard potentially sensitive issues, such as political matters.

In regard to the project participants 'the objective scientist' may not establish the feeling of trust necessary for communication of personal information.

Evidence from the literature on project evaluation suggests that the last role of the evaluator, as the consensus-building consultant or the 'facilitator', is a rarity. The rejection of an approach where "evaluators and information users are partners in the search for useful information" may at first seem odd, but some explanations can be provided.

One explanation is the perception of the role of the

---

[134]. Cook et al., 1980:489.

[135]. OECD, 1988:16.

evaluator as seen by the evaluator. Many evaluators - with different educational backgrounds - see themselves in the role of the "expert" rather than the "servant". A sample survey of program evaluators in the United States for example revealed that the respondents tended to see their role as that of an expert rather than a servant to program managers or stakeholders.[136]

A more compelling explanation is the differences that exist between the understanding and purposes of evaluation as seen by recipients and donors. The recipients, to the degree they are interested in evaluation, see the evaluation as a potential means for improving planning and implementation of their own development plans, whereas the donors are mostly concerned about improving the cost-effectiveness of their development assistance, as discussed in chapter 2. In other words there is a conflict of goals for the evaluation. Following Michael Patton's model it is not the partnership between the evaluator and the information-users that is at stake - rather, it is the fundamental differences in purposes between the two groups of information-users that prevents the

---

[136]. The survey used a stratified, random sample of the membership directories of the Evaluation Network and the Evaluation Research Society. Another finding from the survey revealed that, although the surveyed evaluators see themselves as experts, there is a very low level of familiarity with theory among the respondents. The authors conclude: "Such findings suggest there is a danger of scholarly illiteracy in evaluation about its own writings and concepts." Shadish and Epstein, 1987.

evaluator from assuming the role of consensus-building consultant.[137]

The problems standing in the way of donor and recipient partnership were highlighted at a recent seminar of the Development Assistance Committee of OECD to which representatives of 13 developing countries were invited.[138] The donors and recipients each had an opportunity to present their perspective on the 'state of the art' in evaluation of aid projects.[139]

Recipient country participants complained that donors frequently want to carry out evaluations without taking sufficient account of the recipients own needs and criteria and thus give the host country an inadequate role in the evaluation. Donors schedule evaluation at very short notice or at inconvenient times during the budgetary cycle. Furthermore, the absence of coordination between donors in their scheduling makes it a common event that host countries have several evaluation missions in the country at the same

---

137. This may explain why even agencies like SIDA (Sweden), that have placed considerable emphasis on strengthening local evaluation capacity , have experienced scant interest for its support; OECD, 1986:48-9.

138. This provides a rare opportunity to hear the perspective of the recipient countries on evaluation of aid projects.
The members of DAC are Australia, Austria, Belgium, Denmark, Finland, France, the Federal Republic of Germany, Italy, Japan, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, the United States and the EEC.

139. OECD, 1988.

time.

Participants representing the donor interests blamed the lack of involvement of the recipient in evaluation on the lack of sufficiently experienced staff due to competing demands from other planning tasks.[140] When donors try to identify local skills - in consulting firms, universities, management and research institutes - as team members they are exposed to accusations of imposing their own preferences, instead of accepting people nominated by the local authorities. Mixed teams also frequently experience difficulties in reaching forceful conclusions on the evaluation.

Another important factor of the institutional setting for project evaluation in the Third World is the financing of evaluations. The financial constraints of most host countries are such that evaluations generally are funded by the donors. This inequality may easily lead to dominance by the donor.

Both donors and recipients have expressed their desires that more joint efforts will be made in the future.[141] Whether in reality a basis exists for reconciliation of the

---

[140]. The World Bank has on a separate occasion expressed a similar sentiment, arguing that the recipient countries have either failed to see the use of evaluation, because of the long-term focus and the cost involved, or that they are simply not capable of addressing "the 'what happened and why' questions systematically and forthrightly,..." because of the constraints of an over-extended bureaucracy. Weiner, 1984:128.

[141]. OECD, 1986 and 1988; Weiner, 1984.

differences between donor and recipient shall remain unanswered. On the part of the recipient, an inherent resistance to evaluation on grounds of unwillingness to suffer 'exposure' may prevent such reconciliation. On the part of the donor, there may also be a fear of exposure, but more likely, there may in reality be little incentive to involve the recipient - for the sake of the recipient - if donor agencies are generally satisfied with their own evaluations. Some institutional arrangements can, however, be suggested which may lead to a better balance between donor and recipient in evaluations. Let us for a moment again turn to the case of the Dandora project.

The monitoring and evaluation of the Dandora project was an integral part of the project agreement between the Government of Kenya and the World Bank. The evaluation was carried out by a team of local consultants under the coordination of an inter-ministerial/inter-agency steering committee under the Ministry of Finance.[142] While the evaluation can be criticized on other accounts, the institutional arrangement may provide some useful ideas. The funding for the Dandora project was divided almost equally between the Kenyan government and the Bank[143]. If the funds for the evaluation are a part of the total initial funding

---

[142]. Chana, 1984:28-9.

[143]. Rasmussen, 1987:21. This suggestion may prove of little utility in situations where the recipient has a limited economic stake in a project.

for the project the dominance of the donor, often justified by its role as funder, may be lessened.[144] Coordination of the steering committee by a host-country ministry may further increase the participation of the recipient. Likewise, local consultants may have a more profound knowledge of the local decision-making processes, and thus increase the likelihood of utilization. The evaluation of the Dandora project indicates that there is a role for the recipient to play in evaluation of aid projects, provided the funds and the institutional arrangements for the evaluation are part of the project agreement.

A possible institutional approach to evaluation that includes the recipient and the donor as well as others affected by the evaluation is the ´stakeholder approach´. This approach is based on the assumption that "people who have a stake in an evaluation should be actively and meaningfully involved in shaping the evaluation to focus it on meaningful and appropriate issues, thereby increasing the likelihood of utilization."[145] But, what does it mean to ´have a stake´ in an evaluation? Carol Weiss offers a definition of a stakeholder:

---

144. I have previously criticized the inclusion of a priori definitive statements about how an evaluation should be conducted in the terms of reference; this poses a problem for a pre-project decision about how much to set aside for evaluation. My suggestion here would be based on a set percentage-figure.

145. Patton, 1987:116.

> I interpret the term stakeholder to mean either the members of groups that are palpably affected by the program and who therefore will conceivably be affected by evaluative conclusions about the program or the members of groups that make decisions about the future of the program, such as decisions to continue or discontinue funding or to alter modes of program operation.[146]

Some are very enthusiastic about the potential of stakeholder participation. Jennifer Greene provides empirical evidence from the American context that stakeholder participation can have a considerable positive effect on utilization[147], and that for politically controversial evaluation settings "it is particularly important to establish from the outset a forum of policymakers and stakeholders ... that can serve as a microcosm of the political world into which the research will later cast their results."[148] Others are less enthusiastic. Carol H. Weiss provides a "tentative balance sheet" for stakeholder evaluation.[149] She points out that whereas the stakeholder approach can improve the fairness of the evaluation process, counter organizational tilts, democratize access to evaluation information and equalize the potential power

---

[146]. Weiss, 1983:84. It should be noted that the term 'program', used by Carol H. Weiss to define who should be considered a stakeholder, is broader than the term 'project'

[147]. See also Patton, 1982:56-67. Michael Patton establishes a rationale for consultative evaluation practice.

[148]. Kleinfeld and McDiarmid, 1986; quoted in Greene, 1988:110. See also Patton, 1982:55-98; and Patton,1987.

[149]. Weiss, 1983:91-2.

knowledge provides, it cannot bring harmony to contentious program arenas - it may in fact make people even more aware of the conflicts that exist. The stakeholder approach also places new demands on peoples time and attention - an even bigger problem in developing countries than in the American context, for which the article was intended.

The real test of the feasibility of the stakeholder approach is the willingness of those who hold power to share it with those who do not. Are the donor agencies willing to share their influence on the evaluation with the host country agencies? And are the host country agencies willing to share their influence at the community level? The answer to these questions can only be empirically based.

Conclusion on institutional aspects of evaluation.

Although suggestions can be made for future institutional arrangements that could change project evaluation, the current reality is that the evaluations' terms of reference often appear as part of the project agreement and often list the type of evaluation to be conducted, aspects of the project to be evaluated and methodologies to be entertained. There is even evidence that terms of reference for development projects "...are becoming increasingly specific about the design and implementation....,and usually require recommendations as well as major findings."[150] Should this trend continue,

---

[150]. OECD, 1986:11.

considering the predominance of cost-effectiveness evaluation, there may be little hope for different and better evaluations. At the very least, there will be little left for the evaluator to decide in terms of the nature of the inquiry.

CONCLUSION

The nine evaluation reports, I review in the paper, as examples of the state of the art in evaluation of housing projects in developing countries, have adopted four different design strategies for evaluating a housing project. While they differ in this respect they are alike in their, at best, very superficial treatment of the socio-political context of the projects they evaluate and, at worst, total neglect of the influence of the context on the project process, outcome and impact.

In the critique of the four evaluation designs and the methods used for data collection and analysis I single out four evaluations for scrutiny. Each of the designs used has its own practical or theoretical problems, but the scientific experimental design proves most inadequate for evaluating social action programs, such as housing projects. This holds true particularly for the Third World context. Most new housing projects require active involvement of people in new and unfamiliar roles. Evaluating such projects depends to a large degree on understanding changes in attitude and behavior. In this particular context quantitative, primarily ´knowledge verifying´ methods are wholly unsuited. These methods - adopted from scientific research - can in reality

only invalidate a theory about a causal relationship. They do not generate "new" knowledge.

The question of methodological choice is often seen as a choice between quantitative and qualitative methods. The solution is as often presented as a combination of "less-than-perfect" variants of scientific research methods with more innovative methods to make up for the shortcomings of the former. I believe, that this solution is wrong. It results from a failure to recognize the existence of two different research paradigms, both of which use quantitative and qualitative methods. The two paradigms, here termed the scientific and the naturalistic paradigm, are based on different philosophical premises, and must in many ways be seen as antagonistic rather than supplementary.

Behind the choice among evaluation designs and research paradigms lie two distinctly different approaches to evaluation. I use the terms cost-effectiveness evaluation and illuminative evaluation to describe these two approaches. The nine evaluation studies, I review, are used as examples of the two approaches. The two approaches have a great deal in common with two distinctly different dictionary meanings of the verb ´to evaluate´. Cost-effectiveness evaluation fits a definition of evaluation meaning ´to appraise´; illuminative evaluation fits a definition meaning ´to appreciate´. Whether by ´to evaluate´ we mean ´to appraise´ or ´to appreciate´ in many ways determines our choice of design and research

paradigm for evaluations.

Effective evaluation is, however, more than a methodological problem. It also depends on an understanding of the real world context in which the evaluation will be carried out. Project evaluation has most often depended upon evaluation methodologies developed in an American context. Underlying assumptions about the political and administrative environment, that fail to appreciate the typical Third World context, are built into the evaluations. Rather than providing a general outline of the politico-bureaucratic context of developing countries I use my observations from a stay in Kenya, where I studied the Dandora Sites and Services Project and other low income housing projects in Nairobi, to serve as a real-life example of the overriding importance of including the socio-political context in evaluation.

A full appreciation of the context must also include an understanding of the fundamental importance of gender issues in evaluation. Paula Nimpuno´s research on women´s participation in the Dandora sites and services project in Nairobi points out the significant differences in the experience of women and men in low income housing projects. Data from other projects indicate that such differences are common.

No one evaluator can easily fulfill the requirements needed to do effective evaluation as outlined in this paper. A team seems the obvious solution for combining the different

skills, breadth of experience, and local understanding necessary, and at the same time avoiding seriously biased views. But, what may be ideal may not be possible. Scarcity of funds may largely determine the choice of evaluator.

I argue that many of the shortcomings of evaluation are tied to institutional constraints. The different agendas and pressures on donors and recipients often stand in the way of effective evaluation. Indications are that both donors and recipients desire more joint evaluation efforts in the future. Institutional approaches can be suggested, such as joint financing and the stakeholder approach, that may help overcome some institutional constraints. But, whether in reality a basis exists for reconciliation of the differences between donor and recipient is less clear. On the part of the recipient, an inherent resistance to evaluation on grounds of unwillingness to suffer ´exposure´ may prevent such reconciliation. On the part of the donor, there may also be a fear of exposure, but more likely, there may be little incentive to involve the recipient, if donor agencies are generally satisfied with their own evaluations.

Considering the predominance of cost-effectiveness evaluation, the current trend of increasingly specific statements in the terms of reference of development projects, about the type of evaluation to be conducted, may also limit the hope for different and better evaluations.

Despite these constraints I believe that evaluation is

the only effective way in which we can systematically learn from our experience with development projects and not confine this knowledge to a few. The critique presented in this paper, of the state of the art in evaluation of housing projects, is a contribution to an understanding of how projects should be evaluated.

In his book "The Evaluation of Cultural Action" Howard Richards describes the type of evaluation which, I believe, grows out of the critique I have developed in this paper. I can do no better than to end by borrowing from Richards' conclusion:

> Everything seemed clear to me in retrospect. The evaluation of a programme, I now saw, consist to a large extent of describing it in an illuminative way, in a way that helps the reader to see it as if he or she were there. The description should first be expressed in the language of the participants themselves. To see how and why....change takes place, the viewpoints of the actors who are changing, and an understanding of the conflicts they are involved in, are indispensable. A quantitative estimate of how much change is taking place and checking - triangulating - to improve the accuracy of the picture emerging from the description, are important but secondary.[151]

---

[151]. Richards, 1985:242.

BIBLIOGRAPHY

Aiken, Linda H. and Barbara H. Kehrer, eds. Evaluation Studies Review Annual. Beverly Hills, CA: Sage, 1985.

Akivaga, S.K., W. Kulundu-Bitonye and M.W. Opi. Local Authorities in Kenya. Nairobi: Heinemann, 1985.

Andreasen, Jorgen, Niels H. Kryger and Njoki Ndungu. Case Study of Sites and Services Schemes in Kenya: Lessons from Nairobi and Thika. UNCHS-KUL Workshop on Housing in Development, January 1987. Nairobi: United Nations Centre for Human Settlements (Habitat), n.d.

Anyanwu, C.N. "The Technique of Participating Research in Community Development," Community Development Journal, Vol.23 No.1 (January 1988): 11-15.

Bamberger, Michael, Bishwapura Sanyal and Nelson Valverde. Evaluation of Sites and Services Projects: The Experience from Lusaka, Zambia. World Bank Staff Working Papers, Number 548. Washington, D.C.: The World Bank, 1982.

Bamberger, Michael and Eleanor Hewitt. Monitoring and Evaluating Urban Development Programs: A Handbook for Program Managers and Researchers. World Bank Technical Paper Number 53. Washington, D.C.: The World Bank, 1986.

Bamberger, Michael, Edgardo Gonzalez-Polio and Umnuaye Sae-Hau. Evaluation of Sites and Services Projects: The Evidence from El Salvador. World Bank Staff Working Papers, Number 549. Washington, D.C.: The World Bank, 1982a.

Browning, R. "Evaluation in the ODA. A view from the inside," Public Administration and Development, Vol.4 (1984): 133-139.

Chana, T.S. "Nairobi: Dandora and Other Projects", in Low Income Housing in the Developing World, 17-36, ed. Geoffrey K. Payne. Chichester: John Wiley & Sons, 1984.

Community Aid Abroad. Guidelines for Self-Evaluation. n.d.

Cook, Thomas D., Judith Levinson-Rose and William E. Pollard. "The Misutilization of Evaluation Research: Some Pitfalls of Definition," Knowledge: Creation, Diffusion, Utilization, Vol.1 No.4 (June 1980): 477-498.

Dandora Community Development Project Department. Annual

Report. Nairobi, 1975.

Dandora Community Development Project Department. Annual Report. Nairobi, 1976.

Dynes, Patrick S. and Mary K. Marvel. Program Evaluation: An Annotated Bibliography. New York: Garland Publishing, 1987.

Feuerstein, Marie-Therese. "Finding the Methods to Fit the People: Training for Participatory Evaluation," Community Development Journal, Vol.23 No.1 (January 1988): 16-25.

Government of Kenya. The Local Government Act, rev.ed. Nairobi: Government Printer, 1978.

Government of Kenya. Organization of the Government of the Republic of Kenya. Presidential Circular No. 1/86. Nairobi: Office of the President, 5th August 1986.

Greene, Jennifer G. "Stakeholder Participation and Utilization in Program Evaluation," Evaluation Review, Vol.12 No.2 (April 1988): 91-116.

Grindle, Merilee S. "Anticipating Failure: The Implementation of Rural Development Programs," Public Policy, Vol.29 No.1 (Winter 1981): 51-74.

Guba, Egon and Yvonna S. Lincoln. Effective Evaluation: Improving the Usefullness of Evaluation Results Through Responsive and Naturalistic Approaches. San Fransisco, CA: Jossey-Bass, 1981.

Guba, Egon and Yvonna S. Lincoln. Naturalistic Inquiry. Newbury Park, CA: Sage, 1985.

Hoole, Francis W. Evaluation Research and Development Activities. Beverly Hills, CA: Sage, 1978.

Housing Development Department. Annual Report. Nairobi, 1978.

Housing Development Department. Annual Report. Nairobi, 1979.

Housing Development Department. Annual Report. Nairobi, 1980.

Housing Development Department. Annual Report. Nairobi, 1981.

Housing Development Department. Annual Report. Nairobi, 1982.

Housing Development Department. Annual Report. Nairobi, 1983.

Housing Development Department. Annual Report. Nairobi, 1984.

Housing Policy and Systems in Kenya: Nairobi's Experience. Nairobi, n.d.

Hudson, Barclay M. Domains of Evaluation. School of Architecture and Urban Planning, UCLA, 1975.

Information Please Publishers. Information Please Almanac 1983. New York: A & W Publishers Inc., 1983.

International Labour Office. Yearbook of Labour Statistics 1986, Geneva: International Labour Office, 1986.

Keare, Douglas and Scott Parris. Evaluation of Shelter Programs for the Urban Poor: Principal Findings. World Bank Staff Working Papers, Number 547. Washington, D.C.: The World Bank, 1982.

Kiamba, Makau. Private Developers on the Housing Scene in Kenya. UNCHS-KUL International Workshop on Housing, Nairobi, 1987. Nairobi: UNCHS/Habitat and HRDU, University of Nairobi.

Maina, G.G. Actors and Housing Systems in Kenya. Nairobi: Ministry of Works, Housing and Physical Planning and National Housing Corporation, 1987.

Morrison, Minion K.C. and Peter C.W. Gutkind, eds. Housing the Urban Poor in Africa. Syracuse, NY: Maxwell School of Citizenship and Public Affairs, Syracuse University, 1982.

Nairobi City Commission. Committee Meeting Minutes. Nairobi: City Hall, 1986.

Nathan, Richard P. "The Missing Link in Applied Social Science," Evaluation Studies Review Annual, Vol.10 (1985): 716-722.

Nientied, Peter and Jan van der Linden. "The 'New' Policy Approach to Housing: A Review of the Literature," Public Administration and Development, Vol.8 (1988): 233-240.

Nimpuno, Paula. Gender Issues in Project Planning and Implementation: The Case of Dandora Site and Service Project, Kenya. DPU Gender and Planning Working Paper No. 13, Development Planning Unit, Bartlett School of Architecture and Planning, University College London, 1986.

Obudho, R.A. and Constance C. Mblanga, eds. Slum and Squatter

Settlements in Sub-Saharan Africa: Toward a Planning Strategy. New York: Prager Publishers, 1988.

OECD. Evaluation in Developing Countries: A Step in the Dialogue. Paris: OECD, 1988.

OECD. Methods and Procedures in Aid Evaluation. Paris: OECD, 1986.

Palumbo, Dennis J., ed. The Politics of Program Evaluation. Newbury Park, CA: Sage, 1987.

Patton, Michael Quinn. Creative Evaluation. Beverly Hills, CA: Sage, 1981.

Patton, Michael Quinn. "Evaluation's Political Inherency: Practical Implications", in The Politics of Program Evaluation, 100-145, ed. Dennis J. Palumbo. Newbury Park, CA: Sage, 1980.

Patton, Michael Quinn. Practical Evaluation. Beverly Hills, CA: Sage, 1982.

Patton, Michael Quinn. Qualitative Evaluation Methods. Beverly Hills, CA: Sage, 1980.

Rasmussen, Susanne. Implementation and Institutions: A Case Study of the Housing Development Department in Nairobi. Master's thesis, Aalborg Universitetscenter, Danmark, June 1987.

Republic of Kenya. Development Plan 1984-88. Nairobi: Government Printer, 1983.

Republic of Zimbabwe. Low-Income Housing Projects: Kwekwe and Gutu - Project Evaluation Report. Zimbabwe, 1986.

Richards, Howard. The Evaluation of Cultural Action. London: The MacMillan Press Ltd., 1985.

Room, Graham. Cross-National Innovation in Social Policy: European Perspectives on the Evaluation of Action-Research. Houndsmills, UK: the MacMillan Press Ltd., 1986.

Salmen, Lawrence F.. Listen to the People: Participant-Observer Evaluation of Development Projects. New York: Oxford University Press, 1987.

Sanyal, Bishwapriya. "Problems of Cost-Recovery in Development Projects: Experience of the Lusaka Squatter Upgrading and Site/Service Project," Urban Studies,

Vol.24 (1987): 285-295.

Shadish, Jr., William R. and Roberta Epstein. "Patterns of Program Evaluation Practice Among Members of the Evaluation Research Society and Evaluation Network," Evaluation Review, Vol.11 No.5 (October 1987): 555-590.

Skaburskis, Andrejs. "Cost-Benefit Analysis: Ethics and Problem Boundaries," Evaluation Review, Vol.11 No.5 (1987): 591-611.

Smith, Thomas B. "Evaluating Development Policies and Programmes in the Third World," Public Administration and Development, Vol.5 No.2 (1985): 129-144.

Soni, Praful Naran. On Self-Help in a Site and Services Project in Kenya. Master's thesis, Massachusetts Institute of Technology, 1980.

Staub, William J. and Bruce Koppel. "Monitoring and Evaluation of Benefits in Agriculture and Rural Development," Asian Development Review, Vol.4 No.1 (1986): 100-110.

Staudt, Kathleen. "A Planning-Centred Approach to Project Evaluation: Women in Mainstream Development Projects," Public Administration and Development, Vol.5 No.1 (1985):
25-37.

Stren, Richard E. "Housing Policy and the State in East Africa" in Housing the Poor in Africa, 81-104, eds. Minion K.C. Morrison and Peter C.W. Gutkind. Syracuse, NY: Maxwell School of Citizenship and Public Affairs, Syracuse University, 1982.

Stren, Richard E. Housing the Urban Poor in Africa: Policy, Politics, and Bureaucracy in Mombasa. Berkeley: Institute of International Studies, University of California, 1978.

Temple, Frederick T. and Nelle W. Temple. "The Politics of Public Housing in Nairobi", in Politics and Policy Implementation in the Third World, 224-249, ed. Merilee S. Grindle. New Jersey: Princeton University, 1980.

Tendler, Judith. Inside Foreign Aid. Baltimore: John Hopkins University Press, 1975.

Tendler, Judith. Turning Private Voluntary Organizations Into Development Agencies: Questions for Evaluation. A.I.D. Program Evaluation Discussion Paper No. 12. U.S. Agency

for International Development, April 1982.

Thorner, Daniel and Alice Thorner. Land and Labour in India. New York: Asia Publishing House, 1962.

Tolchin, Susan J. "The Political Uses of Evaluation Research: Cost-Benefit Analysis and the Cotton Dust Standard", in The Politics of Program Evaluation, 249-269, ed. Dennis J. Palumbo. Newbury Park: Sage Publications, 1987.

United Nations. Systematic Monitoring and Evaluation of Integrated Development Programmes: A Source-Book. New York: United Nations, 1978.

United States Agency for International Development. Evaluation and Monitoring Guidelines. Office of Housing, A.I.D., November 1978.

United States Agency for International Development. Evaluation of Housing Programs of the Government of Sri Lanka. Washington, D.C.: A.I.D., October 1982.

United States Agency for International Development. Housing and Community Upgrading for Low-Income Egyptians, Final Evaluation Report. Cairo, Egypt: Bureau for Near East, A.I.D., February 1982.

United States Agency for International Development. Project Evaluation Mission Botswana. Washington, D.C.: Office of Housing and Urban Programs, A.I.D., December 1988.

Wanjohi, I.G. and T.S. Chana. Strategies for Housing the Lower Income Groups in the Dandora Community Development Project, Nairobi: A Case Study. Nairobi: Housing Research and Development Unit, University of Nairobi, 1977.

Weiner, Mervyn L. "Making Evaluation Work, A Case Study: The World Bank," Optimum, Vol.15 No.4 (1984): 124-133.

Weiss, Carol H. "Toward the Future of Stakeholder Approaches in Evaluation," in Stakeholder Based Education, ed. A.S. Bryk, 83-96. San Francisco: Jossey Bass, 1983.

Weiss, Carol H. and Theodore D. Fuller. "On Evaluating Development Assistance Projects: Some Innovations," Evaluation Review, Vol.7 No.2 (April 1983): 175-190.

Weiss, Carol H. Evaluation Research: Methods for Assessing Program Effectiveness. Englewood Cliffs, NJ: Prentice-Hall, 1972.

APPENDIX I

PARTICIPATORY EVALUATION: THE EXAMPLE OF OXFAM AMERICA.

This appendix provides a brief introduction to the concept of participatory evaluation, its advantages and problems exemplified by the experience of Oxfam America.[1]

Oxfam's approach to evaluation is closely linked to the ideology behind its development assistance strategy in general.[2] Self- and participatory evaluation plays an important role - virtually all projects have some element of self-evaluation and around half the projects choose a participatory approach to evaluation. If, however, Oxfam works with government agencies local Oxfam staff is usually responsible for the evaluation. The typical isolation of many - i.e. rural - Oxfam projects explains this arrangement, but Oxfam's perception of government agency performance also plays a role.[3] Projects are usually evaluated at the end of a

---

1. Oxfam America is an international agency that funds self-help development projects and disaster relief in Africa, Asia and Latin America, and also prepares and distributes educational material for people in the United States on the issues of development and hunger. Oxfam America is one of seven autonomous Oxfams around the world.

2. The description of Oxfam's approach to evaluation is based on an interview with one of Oxfam's Regional Directors, at Oxfam America Headquarters in Boston (October 28, 1988).

3. Government agencies are described as typically having the responsibility of too many other tasks and facing too many other problems to be able to devote time to Oxfam evaluation activities. One may also suspect that

project-cycle - a 3-5 year commitment - if it is decided funding should be continued. The overall use of ´ex-post evaluation´ is, however, rather limited - only around 10% of all projects are evaluated.[4]

Ideally, participatory evaluation seeks to involve the funding agency, any intermediate agency and the local grass-roots organizations representing the project beneficiaries. Participation is sought in identification of evaluation objectives, definition of criteria and indicators for assessing if objectives have been reached and - in the case of self-evaluation - also in the actual collection and analysis of data.[5]

Oxfam´s evaluations tend to be biased towards gender, ethnicity, race, class, and other disadvantaged groups, and towards partnership and equality values. Furthermore, the evaluations tend to be heavily biased towards **process**. The feasibility of the policy implemented is considered of lesser importance, e.g. economic viability and replicability of a

---

Oxfam does not want government biases to be reflected in self-evaluations.

4. Other donors typically evaluate between 15-20% of their projects every year, although the differences are considerable. OECD, 1986:17.

5. According to "A Field Director´s Handbook" (1985) - published by Oxfam UK - in which guidelines are provided for monitoring and evaluation, as well as a host of other aspects of disaster relief and development assistance. Oxfam America does not necessarily follow these guidelines, rather they represent the ´ideal´.

project usually receives the benefit of the doubt and do not become determining factors, when further funding is considered. As expressed by one of Oxfam´s regional directors the evaluations tend to become matters of whether the project beneficiaries have made a genuine effort to make the project work.

A primary goal of Oxfam´s evaluations is to help the project, i.e. the evaluation serves as a planning exercise for project management and beneficiaries. A secondary purpose is to control that the money assigned to the project has been spent well. But, evaluation findings are also meant to aid Oxfam´ board of directors and staff in the formulation of new policies and in making choices about which projects to fund. Oxfam staff does not follow specific guidelines for evaluation design, since only a tentative policy exists for how evaluations should be carried out.

Using self- and participatory strategies has, according to one of Oxfam´s regional directors , had the effect of biasing Oxfam´s evaluations.[6] People tend to judge themselves more positively ("less objectively"), wishing to retain project benefits, even if the project is not cost-effective. And, self-evaluations often only pronounce this problem. Such bias is, however, not inherent to participatory evaluation, and - as pointed out by the Regional Director - can be

---

[6]. Little empirical data exists, since Oxfam America´s evaluation studies have not been made available to the public.

remedied by assistance from outside evaluators "[..] who at the same time are sensitive to the 'environment' and Oxfam's policies."[7]

The problem of bias is also sometimes aggravated by the fact that Oxfam cannot guarantee that the local 'grass-roots' organizations, Oxfam has chosen to work with, actually represent the intended target-group. As documented by Tendler (1982), in a study of 75 evaluations of PVO projects, contradictory to the self-image of many PVO's, participation often means top-down or 'outside-in' approaches through local elites.[8] For obvious reasons - in particular if the evaluation serves as input in decision making processes about further funding - local elites will not point to such issues as, or rate negatively, the project's ability to reach the poorest of the poor, the project's truly participatory nature or its ability to empower the broader base of the community.

Oxfam's bias towards process also has implications. As Tendler points out, being biased towards process can place

---

[7]. Interview, October 28, 1988.

[8]. Tendler, by the way, does not see this as necessarily negative. Successful projects include 1)cases where the poor were reached and their lives improved although strong local-elite leadership or the PVO itself masterminded the activity, with little participation involved; 2) cases where local elites or the middle poor were reached, but substantial indirect benefits or trickle-down affected the poor; or 3) cases where better-off members of a community were provided a service previously not available - even though no attempts were made to reach the poorest of the poor, and even though there were no spread effects. (Tendler, 1982:9-10).

unnecessary constraints on an organization's ability to pose evaluation questions:

> The PVO, for example, may be less interested in whether a community decided to initiate an agricultural extension or a health program than it is in the way the decision was made. Yet if one type of project works better with participatory action than another--or if the PVO is better at assisting in one area than another--then the type of activity <u>is</u> relevant to the successes of the undertaking. Strong commitment to open-ended processes of decisionmaking, in sum, may make it difficult to think about evaluation questions of this nature.
>
> Tendler, 1982:22-3.

In this manner process bias actually defeats one very important purpose of the evaluation activity, namely to increase the knowledge about which projects not only materially benefit the poor, but also which projects - through participation - facilitates organization and empowerment of the poor within their communities.

Not having organizational guidelines or standards has had the effect that Oxfam's evaluation studies vary greatly in scope, depth and general quality.[9] One must suspect that the incentive to carry out truly participatory evaluations - including initiating training - will vary greatly in accordance with the <u>actual</u> participatory nature of the project, the time and financial constraints[10] facing the local field representative or project manager (depending on

---

[9]. Again, according to one of Oxfam' Regional Directors.

[10]. Financial constraints of the project may inhibit the hiring of an outside facilitator/evaluator.

who has been given the ultimate responsibility for initiating the evaluation), the potential problems experienced during implementation with the participation component[11], and - in particular - serious disagreements among stakeholders about the project.

Probably the most difficult aspect of participatory evaluation is making sure that even findings that are embarrassing or unwelcome to some stakeholders are included in the evaluation.[12] Anyanwu writes:

> [..] because the participatory research approach is generally problem-centred, the technique cannot be that of mere data gathering, but must accommodate an active endeavour to understand the conditions underlying community problems, with a view to finding lasting solutions to such problems.
>
> Anyanwu, 1988:15.

In other words, participatory evaluation demands a lot of attention on the part of the field and project staff. Exactly because of its potential ability to bring even unwanted or embarrassing facts out in the open a lot of effort must be put into the training of involved parties to understand the purpose and the processes of participatory evaluation. Successful participatory evaluation depends on

---

[11]. As pointed out by the Regional Director, launching a participatory evaluation - if the project itself could not be characterized as participatory - is most likely a wasted effort.

[12]. Feuerstein points out that the phenomena of embarrassing or unwelcome findings may be the reason why some evaluation reports suffer the fate of 'file and forget' and the result are not shared. Feuerstein, 1988:22-3.

reaching consensus on evaluation objectives and methodology, on teaching evaluation methodologies to participants and on gaining the acceptance of decentralization of decision-making and power from all participants - including representatives from intermediary agencies such as government agencies.[13] The potential of participatory evaluation also depends - as Feuerstein notes - to a large degree on the role of the participatory facilitator:

> There can be areas which local people either forget to look at, or do not want to look at. An outsider can play an important role by asking the right kind of questions and providing useful insights into dealing with dilemmas and uncertainties. (emphasis added).
>
> Feuerstein, 1988:23.

One of the most important aspects of participatory evaluation is thus the facilitator's role as a negotiator between the participants in the evaluation, e.g. the government, agencies, project staff and participants.

In light of the difficulties that participatory evaluation may entail it is easy to imagine - where no guidelines exist - that a field representative will choose not to opt for participatory evaluation.

Dissemination of findings.

According to the Regional Director the institutional

---

[13]. This last feature of participatory evaluation is the most controversial; as Feuerstein (1988:16) notes, while many people in development activities may be ready to share responsibility, there are few who are genuinely ready to share power.

network of the organization is, however, very week and the dissemination of findings is only informal and ad hoc. Ironically, there are examples - for example from Oxfam projects in India - that institutions elsewhere are stronger, and findings have been utilized by local PVO´s and government agencies. In this respect evaluation findings may have a much greater impact for knowledge dissemination in local areas than for Oxfam itself. Oxfam hopes within the next 3 years to establish a more systematic approach to evaluation and dissemination internally, and also for usage by the PVO community in general. An ´embryonic´ approach is under way, developing models for the response to disaster and the evaluation of such disaster relief programs.[14] Currently, however, the dissemination of findings is very poor, which affects Oxfam´s ability to improve its policies consistently with findings from the field, to the effect that policies remain based mainly upon ideological considerations and ´anecdotal´ evidence. It also affects Oxfam´s ability to participate in the debate with other development assistance organizations about future aid policies, limiting the circulation of experience with participatory evaluation to larger donors such as the World Bank and USAID whose evaluation strategies have not included participatory evaluation.

---

[14]. Three programs in the Philippines, India and Ecuador provides empirical data for the study.