



UNIVERSITY
OF OSLO

Working Paper No. 2/02
***Political Evaluation:
Premises, Approaches
and Methods***

William M. Lafferty

Working Paper



Program for Research and
Documentation for a Sustainable Society

Centre for Development and
the Environment

ProSus 2002
Program for Research and Documentation
for a Sustainable Society (ProSus)
Centre for Development and the Environment
University of Oslo
P.O.Box 1116 Blindern
N-0317 Oslo, Norway
Tel: + 47 22 85 89 00
Fax: + 47 22 85 87 90
informasjon@prosus.uio.no
www.prosus.uio.no

CONTENTS

- 1 INTRODUCTION..... 5
- 2 NORMATIVE-EMPIRICAL ANALYSIS OF DEMOCRATIC PERFORMANCE..... 7
- 3 STRATEGIC RESEARCH FOR SUSTAINABLE DEVELOPMENT: INSTITUTIONAL PREREQUISITES 11
- 4 “GROUNDING” EVALUATION: WHAT IS “SUSTAINABLE DEVELOPMENT”? 13
- 5 OPERATIONAL GUIDELINES: THE NEED FOR “RULES OF THUMB” 15
 - 5.1 DIFFERENTIATION CRITERIA 15
 - 5.2 OPERATIONAL PARADIGMS 17
 - 5.3 PERFORMANCE BENCHMARKS 19
- 6 CONCLUDING PERSPECTIVES: STRENGTHS AND WEAKNESSES? 21
- REFERENCES 23

1 INTRODUCTION

The use of evaluation as approach and method in political science has very little prominence. A review of seven major “handbooks” of political science published between 1975 and 1999 reveals no significant reference to the approach as an independent method.¹ The only area where evaluation has played a noticeable methodological role is in the sub-field of “policy analysis”. The tradition here is, however, very strong, connected mainly to the study of policy implementation in the United States, Scandinavia and the Netherlands. A key methodological text in this area is Vedung (1997), and there have recently appeared two major collections of the most significant contributions to the field across a broad spectrum of topics and approaches (Rist 1995 and Nagel 2002).² The approach is also strongly promulgated by The Policy Studies Organization in the United States (subtitled “The International Association for Decision Makers”), and is often featured in major journals on policy analysis.

The common logic underlying nearly all of these studies is that policy implementation is a central feature of public administration, and that it is the responsibility of policy analysts to improve our understanding of both how implementation works, and the degree to which programmes and policies actually achieve their declared goals. But whereas there is considerable disagreement and competition on the first point (among different “schools” of explanatory policy analysis), discussions of different evaluative approaches and methods are relatively seldom. In general one can say that modern political science has been very self-conscious about distancing itself from its normative roots in classic theory, and this has contributed to a certain “nervousness” around issues of values and objectivity. For many positivists, behaviourists, logical-empiricists and rational-choice analysts in political science, “evaluation” sounds too much like “values” and applied science; directions which the mainstream of “academic” political science would prefer to keep at arm’s length. It is no accident, I believe, that the vast majority of evaluative studies in the area of environment and development are carried out by disciplines *other than* political science.

Be that as it may, as the references cited in the present paper indicate, there is more than enough of relevant “discourse” within the sub-field of policy analysis to nourish the development of a systematic and well-disciplined “school” of political evaluation. The key to such a development is not to avoid either values or normative political theory, but to combine the two with established standards of empirical enquiry. This is particularly important in the area of sustainable development, where conflicts on values and principles are rife, but is also a vital aspect of democratic performance in general. The present author has, for the past several decades, been working on research issues in this area, and can illustrate the approach by first looking briefly at what has come to be known as the “normative-empirical

¹ The works consulted were *The Handbook of Political Science* (Greenstein and Polsby, 1975 and subsequent years); *The International Handbook of Political Science* (Andrews, 1982); *The Dictionary of Political Analysis* (Plano, Riggs and Robin, 1982); *Bridges to Knowledge in Political Science: A Handbook for Research* (Kalvelage, Meland and Segal, 1984); *Political Science Research: A Handbook* (1996); *A New Handbook of Political Science* (Goodin and Klingemann, 1996); and *The Blackwell Dictionary of Political Science* (Baley, 1999).

² Additional recent, and highly relevant, contributions are: Mayne and Zapico-Goñi (1997), Riper and Toulemonde (1997), and Bemelmans-Vidéc et al. (1998).

approach” to democratic performance, and then turning more specifically to the evaluation of strategies for sustainable development.

2 NORMATIVE-EMPIRICAL ANALYSIS OF DEMOCRATIC PERFORMANCE

It is central concern of democratic governance that democracies should be *both* legitimate and effective. Most arguments for democracy are addressed to the first characteristic, claiming that democracies are *per se* the most legitimate form of authority. Many (hardly all) are also willing to argue that, in the long run at least, democracies are also the most efficient form of governance. These arguments usually build on one or another version of the “mobilization-of-competence” argument, namely that complex societies require highly integrated systems of personal and institutional interaction, and that it is only by including as many actors as possible in the decision-making process that one can maximize the effectiveness of the overall performance. Combinations of both arguments are currently prevalent in the debate over the legitimacy and effectiveness of the European Union.

Evaluation is an approach that rather uniquely cuts across both types of argument. On the one hand, evaluation is necessary for supporting democratic legitimacy since it is vital that the “people” see that programmes related to electoral majorities are being carried through. Democracy must be *perceived* to be implementing the goals, programmes and policies that politicians have promised. Such perception can, however, be either highly symbolic or highly instrumental. Democracy *may look like* it is working to follow up political commitments – but the *actual degree of instrumental output and change* may be another story all together. Evaluation is thus also necessary to verify the instrumental effectiveness of the implementation process.

Further, both tasks – guaranteeing legitimacy and documenting instrumental effectiveness – require *external objective* evaluation. “Form follows function”, and it goes without saying that *political neutrality* is necessary to monitor the follow-up of political commitments; and *methodological objectivity and discipline* are necessary to document effectiveness. Some journalists *may* be neutral enough to perform the first task (most are not), but even the best and most neutral journalist will lack the professional ethos and methodological schooling of a scientifically trained evaluator.

While the second of the two tasks indicated (monitoring effectiveness) is a common declared goal among professional evaluators, the first task (guaranteeing legitimacy) is not a goal – or at least not a *declared* goal. Indeed the essence of conflict associated with most programmatic or institutional evaluations is the *degree of association or compliance with the contracting party*. Anyone who has either conducted or been subjected to a “conditional evaluation” (an evaluation where the result will directly effect the continuation of the activity under evaluation), knows that the degree of “sympathy” between the contractor and contractee is a key issue affecting the results of the evaluation. The notion that an agent would contract an evaluation which seriously undermines the legitimacy of the agent itself, is simply not part of the normal evaluation “game”. It *is* a part, however – and *must be* a part – of the evaluation of democratic societies and their political programmes. Hence the very basic initial premise that, if evaluations of democratic performance in general, and of national strategies for sustainable development in particular, are to be effectively carried out, the

evaluating unit must have a mandate guaranteeing political neutrality. Whether or not the unit manages to live up to the mandate and manifest it in practice, is, of course, another question.

Given that such a mandate has been both issued and maintained, how can it be applied in practice to democratic performance? As an approach within the Department of Political Science at the University of Oslo, “academic freedom” enabled the development of a programme of “normative-empirical analysis”. During the period 1981-92 (roughly), numerous projects related to different aspects of democratic performance were carried out, all operating on a dual normative-empirical track. The basic steps of the approach were quite simple – though up to that time there were virtually no “paradigms” from the discipline to build on. Over the years the approach led to numerous publications in both English and Norwegian, including several graduate theses and doctoral degrees. The elemental steps of the approach are listed in Box 1

Box 1: Basic steps of normative-empirical research

1. Identification of a specific practical discourse where questions of democratic norms were at issue: these could be questions of equality, freedom, participation, rights, or whatever, and the discourse could be related to any aspect of political performance: national or local democracy, institutions, workplace democracy, women’s rights, etc.
2. Connection of the specific issue in question to one or more academic discourses related to the problematic.
3. “Translation” of the practical-discourse problem into a normative-theoretical discourse problem: clarifying the implications of the problem within a normative-theoretical context
4. Formulation of empirical criteria, drawn from the field of the practical discourse, by which the normative problem could be addressed and clarified
5. Determination of the relevant empirical methodology necessary to an objective analysis of the normative problematic
6. Execution of the empirical analysis, with conclusions for *both* the practical discourse and the normative-theoretical discourse.

Several different types of study were carried out during this period:³

Parliament as a discursive forum: An analysis of the practice of the Norwegian parliament with respect to norms of democratic discourse. Focusing on the debate over the 4-year long-term plan in Norway – where expectations as to meaningful dialogue and constructive discourse should be stronger than for normal parliamentary inter-party exchanges – the study first clarified the issue vis à vis discourse theory, and then applied the norms to an empirical analysis of the debates. The results revealed a clear bias inherent in the rules for debate to the advantage of the largest parties, indicating that normal procedures for allocating parliamentary seats according to electoral pluralities can detrimentally influence the quality of discourse on long-term, overarching issues.

³ All of the studies carried out are in Norwegian. An overview of the relevant theses and dissertations is available at the website of the Department of Political Science: <http://www.statsvitenskap.uio.no/fag/hoved/sammen/>. Some of the more relevant works are Hovde (1982), Langeland (1985), Christensen (1987), Mageli (1987), Tørres (1987), Auråker (1988), Raaum (1988), Skogerbø (1988), Hov (1989), Strandhagen (1990), and Kjelland-Mødre (1991).

Norway as a “Distant Democracy”: Taking its point of departure in a major empirical study of political participation in Norway (Martinussen 1973), the goal of this analysis was to demonstrate that the negative conclusions of the study (portraying Norway as a “Distant Democracy” characterized by low levels of political involvement and alienation), were not warranted on *either* empirical or normative grounds. By initially explicating and reworking the vague normative premises of the study, and then applying many of the same indicators and analytical methods to similar data sets, the study clearly showed that the situation for participation and democracy was much less negative than originally indicated (Lafferty 1981). The debate raised by the study led to numerous other studies and debates focused on the normative criteria for participation in a democracy, with significant consequences for both normative theory and everyday expectations as to what *kind* of participation should/can lead to what kind of results.⁴

Democracy and the media: The general theme here was a major trend during the 1980’s in Norway to “democratize the media”. Taking its point of departure in high-profile political proposals and debates as to the “democratic nature” of the mass media – primarily radio and the press at this juncture – the study did a thorough analysis of the public documents in relation to relative norms for democratic media. Criteria were developed for different aspects of the democratization problem, and numerous analyses of different types of media were then carried out. Once again a major finding was that norms had been vaguely expressed and often incorrectly related to different proposals for change. The empirical analyses also showed (again) that different, and often competing, democratic norms could lead to contradictory performance and conflicts among valued norms (Skogerbø 1988 and 1996).

Democracy and learning at the workplace: Finally, we can mention a whole series of studies where the theme was workplace democratization, and the problem was to untangle a whole spectrum of broadly and vaguely expressed values and expectations. A major challenge for this subset of studies was to differentiate between the specification of democratic norms for a political community *per se*, and the specification of democratic norms for both workplaces in general, and specific types of workplace in particular. It was here that the notion of “form follows function” emerged as a key guideline, indicating that specific types of functional activities – activities which by functional definition require a division of labour and responsibility – require specific modes of democratic procedures and values. In short, the “democratization” of an activity is an adaptive process whereby the essential nature of the activity in question must be given due and balanced consideration in relation to democratic norms.⁵

A major result of all these studies was to develop a specific awareness and identity as a separate research approach to practical problems where democratic norms were directly at issue. We were able to address political issues which were often very contentious; to place those issues within relevant systematic discourses of normative political theory; and to assess existing states of affairs against the normative criteria derived. All this was done, however, within a normal university setting, where the task was principally conducting basic research and schooling political scientists. In the next phase of the work, the profile of the

⁴ The issues have been debated primarily in Norwegian. Interested readers of a Scandinavian language can consult Berg 1983 and Hagtvet 1980. There is also an interesting graduate thesis devoted to the debate between Martinussen and Lafferty on democracy in Norway (Strømsnes 1993).

⁵ See the list of graduate theses in note 2. More general treatments are presented in Lafferty 1984 and 1989, and Lafferty and Raaum 1992.

approach was specifically recognized in a programmatic context, and connected directly to the evaluation of national democratic performance on achieving sustainable development.

3 STRATEGIC RESEARCH FOR SUSTAINABLE DEVELOPMENT: INSTITUTIONAL PREREQUISITES

In 1992 the current author was asked to head an existing programme devoted to the development of alternative macro scenarios for the socioeconomic development of the Norwegian society. The “Project for an Alternative Future” (PAF) had been initiated in 1982 by a small group of NGO-leaders and concerned academics, with the intention of developing alternative paths of development which could have political influence. The project was given direct support by the Norwegian parliament, and for a number of years functioned as a more-or-less ad-hoc think tank for the Committee on Education and Research within the parliament. The project was evaluated by an independent academic committee in 1988, and it was decided to transfer responsibility for the activity to the then Norwegian Research Council for Science and the Humanities (NAVF). A new five-year period was inaugurated in 1990, but the constellation of diverse actors and interests had difficulty finding an appropriate form of applied research. When new leadership was introduced in 1992, it was gradually decided that the major thrust of the activity should be altered from one of developing alternative holistic scenarios, to one of assessing Norway’s progress on issues of sustainable development. The change was inspired largely by the broad promulgation of the findings and recommendations of the World Commission on Environment and Development in its concluding report *Our Common Future* (WCED 1987). The change of leadership took place, moreover, at the exact time of the follow-up conference to the WCED, the United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro in June of 1992.

In 1993 the research community in Norway underwent a major transformation when five separate councils were merged in one Norwegian Research Council (NRC). In addition – and most significantly for the present “story” – a totally new division was created within the Council, the Division on Environment and Development (DED). A major task of the new division was to integrate research on environment and development, both within the division itself but also, as far as possible, within the Research Council as a whole. It was also indicative of the times that one of the principle new objectives adopted for the NRC was to promote research that contributed to “the creation of value-added capital within a framework of sustainable development” (“*verdiskapning innenfor en bærekraftig utvikling*”).⁶

It was within this context that the Project for an Alternative Future was terminated in 1995, and replaced by a new “Programme for Research and Documentation for a Sustainable Society “ (ProSus). The leadership and most of the staff were the same, but the remit from the Research Council was significantly altered. ProSus was defined as an “applied strategic research program” designed “to produce new knowledge and conduct informational activities in support of a better realization of national goals for sustainable development” (ProSus 2000: 1). The activity of the programme was to be conducted along three lines (Box 2).

⁶ The interesting conjunction of “value-added capital” and “sustainable development” was also typical of the times, and particularly in line with similar signals from the European Union, particularly in the Maastricht Treaty.

Box 2: Principle activities of ProSus as defined by the Research Council of Norway

- Documentation and evaluation of Norway's follow-up of the Rio accords and the guidelines of the United Nations Commission for Sustainable Development. The programme is to focus on the political, social and economic goals of the UNCED process, and should submit regular reports on Norway's progress with respect to the declared goals and values.
- Goal-oriented strategic research on the barriers which stand in the way for a more rational and effective realisation of sustainable development. The activity here should be conducted in cooperation with other research institutions, both nationally and internationally, and in close dialogue with voluntary organisations and representatives for business and labour.
- Information and dissemination of alternative strategies of governance, steering instruments and normative future perspectives for more sustainable societies, locally, nationally and globally. The activity here to be coordinated through networks with other research and information efforts in this area.

Of most direct relevance for the issue of political evaluation, is the fact that ProSus receives the major part of its funding from the Ministry of Education and Research, but the allocation is channelled through the Research Council of Norway, guaranteeing that there is no political connection between the Ministry and the Programme. Furthermore, ProSus was, as of 2000, designated a "strategic university programme", placing its evaluation and research activities within the administrative domain of the Centre for Development and the Environment (SUM) at the University of Oslo. SUM is an "all-university" research and teaching unit, established directly under the University Senate. Its remit also prescribes research and information on sustainable development, but is both more basic-research oriented and more focused on development and environment issues in developing countries. The two unites are, therefore, highly complementary, with ProSus focusing more on the pursuit of sustainable development within Norway, and SUM concentrating on sustainable development in an international, North-South context.

These administrative details are necessary to understand the nature of the evaluative strategic research being conducted at ProSus. By monitoring and evaluation of Norway's initiatives to achieve sustainable development, ProSus fulfils an important task for democratic legitimation, at the same time that an empirical foundation is laid for focused research on the barriers to, and new possibilities for, more effective implementation. Such a research operation is a vital adjunct of the strategic intentions of both UNCED and, more recently, the European Union. The political bodies of the international and European communities have made sustainable development an overarching goal for economic and developmental activities in general, and it is absolutely essential that scientific bodies be accorded a role in evaluating and improving implementation efforts.

4 “GROUNDING” EVALUATION: WHAT IS “SUSTAINABLE DEVELOPMENT”?

The first issue that must be resolved in any evaluation is a common understanding of the nature of what is to be evaluated. A high degree of consensus on this issue is a necessary, but not sufficient, condition for a “successful” evaluation. This is true for professional programme and policy evaluation; and it is *particularly* true for independent political evaluation. Clarification of the goals and mandate for change under evaluation is a crucial condition for converting analytic results into better practice.

In developing evaluative criteria for ProSus, we have relied on a relatively simple three-point programme for “grounding” analyses of initiatives for sustainable development.

Firstly, we consciously opt out of the business of trying to tell anyone what sustainable development “really is”. ProSus does not have a substantive programmatic position on sustainable development. Our strategy is to declare that sustainable development is what the governments of the world have committed their countries to within the UNCED process. All of the accords adopted at Rio, particularly the *Rio Principles* and *Agenda 21*, have sustainable development as their goal, and virtually all members of the United Nations have committed themselves to pursuing the goal. Our solution to the definitional problem is, therefore, conventionally political. We believe that this solution provides by far the closest approximation to a “consensus” on what sustainable development is. It is an *instrumental* and *political* solution, which, in our opinion, means that it provides a much better point of departure for change than any of the myriad of alternative definitions, positions and programmes that one can find in the voluminous literature on sustainable development.⁷

Secondly, we point out that the Rio documents make no extensive effort to define sustainable development. This is because the Rio Summit was designed and constructed to specifically follow up, specify and facilitate the understanding of sustainable development as put forth in the WCED report, *Our Common Future*. Insofar as UNCED commits the nations of the world to work for sustainable development, it is the sustainable development of *Our Common Future*. Does this not mean then that we are simply “passing the buck” on the definitional problem to *Our Common Future*, the meaning of which is notoriously contended? We think not – for two reasons: (1) we believe that a textual analysis of *Our Common Future* reveals that the type of development envisaged as “sustainable” is distinct enough on its own terms to contrast sharply with major characteristics of currently competing paths of *non-sustainable* development⁸; (2) we believe that the Rio accords (particularly *Agenda 21*), and the follow-up work of UNCSD, provide more than enough specifics for change (objectives, guidelines, actions, indicators) to move the agenda on sustainable development a *considerable* distance towards change.

Thirdly, we have developed three distinct types of “evaluative criteria” to provide meaningful, and politically relevant, nuances to our evaluations: (1) *External criteria*: the documents and guidelines provided by the WCED, UNCED, UNCSD – and more recently

⁷ See Lafferty 1996 and 1999 for the complete argument here.

⁸ The argument for this position is presented in a collection of studies devoted solely to *Our Common Future* (Lafferty and Langhelle 1999).

UNEP, the OECD, and the European Environmental Agency. These sources provide the most general level of standards for assessing Norway's follow-up to the UNCED-agreements; (2) *Internal criteria*: the strategies, goals, programmes, action plans, policies etc., adopted by government as implementation of the UNCED programme; (3) *Comparative criteria*: the activities, initiatives, programmes, policies, etc., undertaken by other governments in their pursuit of the Rio goals.

This position on sustainable development is explicitly "operative". It is designed to build as much objectivity into the evaluation process as possible; an objectivity anchored in the commitments of democratic governments, and applicable within a very intuitive understanding of the relationship between research and politics. Given that governments mean what they say when they sign international or regional agreements to pursue sustainable development, social research (in this case) has both the right and the obligation to monitor the implementation process. The position is, therefore, not designed to either defend or "preach" a specific understanding of sustainable development, but to accept, clarify and evaluate the semantic and operational implications of the political discourse in question.

5 OPERATIONAL GUIDELINES: THE NEED FOR “RULES OF THUMB”

While the “grounded definition” and standards outlined in the previous section serve as basic orientations for political evaluation, it is also the case that one encounters numerous situations in the evaluation process where the criteria of either *Agenda 21*, the UNCSO guidelines, or the proscriptions of the EU on sustainable development, are open to interpretation, weighting, ranking, etc. This is a most normal situation when pursuing highly normative goals; goals moreover which clearly must be realized in a multitude of different types of society, with markedly different resource bases, economies, social structures, levels of development, etc. Even the most specific sections of *Agenda 21*, or the most refined of UNCSO indicators for sustainable development, will have to be *relativized* to the individual community. Form must be adapted to function, but form must also operate within contextual constraints. Mediating between the definition and standards outlined above and the specifics of any given society requires “rules of thumb” for effective communication and consistent research results. Demands for “cumulative knowledge” are no less relevant for evaluative research than for more “basic” modes of nomothetic science.

“Rules of thumb” are simple prescriptions for the basic application of any craft or discipline. They are what carpenters, plumbers and electricians learn as apprentices, and what social scientists learn (ideally) as graduate students. In applying evaluative standards for sustainable development, we can identify three types of “rules of thumb” which aid in the standardization of the research effort: *differentiation criteria*, *operational paradigms*, and *performance benchmarks*.

5.1 Differentiation criteria

One of the key operational tasks in the evaluation of sustainable development is to know it when you see it. There are a great number of phenomena going on in any society, at any level of development, which may, or may not, be judged as “sustainable development”. Shakespeare made the basic point by pointing out that “Beauty is in the eye of the beholder”; but beauty is not a strategic programme for change, and Shakespeare was not a social scientist. We need help.

Two logical “devices” can provide assistance. Though neither is so intuitively clear as to guarantee consensus, they are at least consistent with the grounded position here outlined. The first has to do with differentiating sustainable development phenomena from other similar phenomena, and the second has to do with differentiating among the so-called “environmental, economic and social pillars” of sustainable development. Both aspects are illustrated by the simple list provided in Box 3. The list provides an overview of “basic components” of sustainable development, with the term “gradients” also used in parentheses to highlight an underlying logic of stages. The list thus reflects a *selection of identifiable characteristics*, formulated as a *sequence of prioritized issues and standards*.

Box 3: Basic components (“gradients”) of sustainable development

The environmental component:

Consisting of three major aspects (phases) of normative environmental politics:

- nature conservation
- environmental protection
- ecological balance

The economic component:

Consisting of the key elements of the Brundtland/UNCED goal of a “qualitatively” different mode of (“sustainable”) economic development (“de-coupling”):

- eco-efficiency (reducing the impact on natural resources for producing goods and services)
- sustainable consumption (consumer responsibility for eco-efficiency)

The social (equity) component:

Consisting of four dimensions of equity for adjusting the satisfaction of “basic needs” to the sustainable functioning of natural life-support systems:

- national equity for current generations
- national equity between current and future generations
- global equity for current generations
- global equity between current and future generations

As a “rule of thumb” for evaluation, the list says: “These are the essential elements of sustainable development (within the UNCED context), and they are to be given prominence in the order listed”. The key implications of this in an operational context are as follows:

(1) Sustainable development incorporates, but is different from all three of the “environmental components”. When “looking for” sustainable development to evaluate, it is important to differentiate it from “pre-sustainable-development” environmental concerns. The conservation of nature, protection of the environment from pollution, and maintenance of ecological balance are all vital aspects of sustainable development, but they were in place and being pursued prior to the political commitment to sustainable development, and their achievement is not sufficient to qualify for high marks on SD implementation.⁹

(2) Having “said” this, it is nonetheless the case that the most fundamental dimension of sustainable development is its environmental or “ecological” component. This is what anchors the concept logically, making it different from other normative concepts,

ideologies and programmes devoted to socio-economic welfare, justice, political reform, etc. In short, the ecological component is a necessary, and “lexicologically” prioritized, dimension of sustainable development – but it is not sufficient to qualify as an independent standard for implementation.

(3) The “economic” component is the second-most important aspect of the concept. It reflects the vital, and highly differentiating, message of *Our Common Future*: namely that any level of economic growth must be fundamentally assessed as to its impact on the sustainable functioning of life-support systems. This involves, if one will, the “materialist bias” of the WCED position. It is the basic relationship between natural life-support systems and the

⁹ The classic statement on this aspect from *Our Common Future* is as follows: “The environment does not exist as a sphere separate from human actions, ambitions and needs, and attempts to defend it in isolation from human concerns have given the very word ‘environment’ a connotation of naivety in some political circles. The word ‘development’ has also been narrowed by some into a very limited focus, along the lines of ‘what poor nations should do to become richer’, ... But the ‘environment’ is where we all live; and ‘development’ is what we all do in attempting to improve our lot within that abode. The two are inseparable.” (WCED 1987: XI).

dispositions of the “means of production” which warrants primary attention in a sustainable development context. Moreover, this relationship must be conceptually and analytically “decoupled”. Where it can be demonstrated that modes of production driven by the under-satisfaction of basic needs are causing harm to sustainable life-support systems, these must be rectified. Likewise, where it can be demonstrated that modes of production driven by excessive satisfaction of basic needs, these too must be rectified. In general, conditions of poverty characterize the first type of imbalance, and conditions of over-consumption characterize the second type.

(3) The “social dimension” of sustainable development (once again, in the UNCED context), is a generalized reference to the ethical stipulation of *Our Common Future*, to the effect that issues of decoupling should be assessed and adjusted according to four dimensions of equity applied to the interdependency between modes of production, the level of satisfaction of basic needs, and impacts on natural life-support systems. These are: (a) national equity for current generations, (b) national equity between current and future generations, (c) global equity for current generations, and (d) global equity between current and future generations. The “social dimension” is viewed in this light as a normative perspective for the assessment and regulation of the relationship between economic activity and natural life-support systems according to fair standards for the satisfaction of basic needs.

(4) Though it is clearly difficult to reach a consensus on benchmarks for both “sustainable life-support systems” and “basic needs”, the conceptual and operational difficulties encountered here are considerably over-dimensioned by both public and academic discourse. Reasonable indicators for critical thresholds of life-support systems – air, water, nutrients, sinks, and non-renewable vital resources – are available, as are reasonable conceptual standards and indicators for basic needs. In an evaluation context, the emphasis must be placed on “reasonable” within the context in question, and not on “reasonable beyond doubt”. The community of evaluation practitioners should, however, be willing to devote considerable more attention and resources to the question of consensual standards.

Once again, it is important to emphasize that these criteria are not meant to impose *general* standards on the understanding and evaluation of sustainable development (probably an impossible task), but to assist in the evaluation of sustainable development as conceptualized and adopted as a political programme within the UNCED process.

5.2 Operational paradigms

Another device which has proved particularly useful in the evaluative work done at ProSus, is the notion of an “operational paradigm”. This is particularly important when the task is one of comparative assessment across different political constituencies, where there is an acute need for common points of reference. By way of illustration we can look at two types of criteria which have been employed with considerable success in the SUSCOM project, an EU-funded “concerted action” to determine the status and implications of the implementation of “Local Agenda 21” (LA21: Chapter 28 of *Agenda 21*) in Europe.¹⁰ Before the project could proceed to an empirical monitoring of progress on LA21, we had to

¹⁰ The SUSCOM project was conducted between 1997 and 1999, covering LA21 activities in 12 European countries. The project is comprehensively reported in Lafferty and Eckerberg 1998, Lafferty 1999 and Lafferty 2001a.

establish common criteria across the 12 research partners as to what would be considered “a Local Agenda 21”. This was accomplished in two steps.

The initial step was on a normative-descriptive level: Once again the question had to be answered – “How would we know ‘an LA21’ if we saw one”? The problem was particularly acute for an assessment of Chapter 28 of the *Agenda*, since the guidelines for objectives and activities were (of necessity) extremely open. Given the enormous diversity of communities throughout the world that would have to work with the Chapter, it was clear that any attempt to provide substance benchmarks for achievement would be counter-productive. We developed, therefore, a relatively simple logic: (A) Chapter 28 was an explicit appeal to “local authorities” to engage in a “dialogue” with their citizens and stakeholders, so as (implicitly) (B) to promote sustainable development. After considerable internal “dialogue” on our own, the SUSCOM team arrived at the set of criteria shown in Box 4.¹¹

Box 4: Basic criteria for identifying Local Agenda 21

- A more conscious attempt to relate environmental *effects* to underlying economic and political *pressures* (which in turn derive from political decisions, non-decisions and markets)
- A more active effort to relate local issues, decisions and dispositions to *global impacts*, both environmentally and with respect to global solidarity and justice.
- A more focused policy for achieving *cross-sectoral integration* of environment-and-development concerns, values and goals in planning, decision-making and policy implementation.
- Greater efforts to increase *community involvement*, i.e. to bring both average citizens and major stakeholder groups, particularly business and labour unions, into the planning and implementation process with respect to environment-and-development issues.
- A commitment to define and work with local problems within: (a) a broader ecological and regional framework; and (b) a greatly expanded time frame (i.e. over three or more generations)
- A specific identification with Agenda 21 and the UNCED process

SUSCOM Project (Lafferty 2001: 4-5)

Once these criteria were agreed – and brought to bear in our joint assessments of the original monitoring efforts – the question arose as to whether it was possible to further clarify the reporting effort by outlining a more dynamic and systematic “model” – an “ideal-type” of “best case” which would reflect the most ambitious aspirations of the Rio action plan. While the criteria were highly useful in established a set of individual characteristics, we soon found that the criteria could be fulfilled individually and piecemeally, without other criteria being manifest at all.

On this second task, it was highly fortunate that one of the most significant documents related to the entire LA21 effort – the Ålborg Charter – had in fact outlined just such a process. Though hardly promulgated at all across Europe (it was included in one of the Charter’s subsections), the list of integrated criteria here provided exactly what was needed for “evaluative imaging”. As shown here in Box 5, the paradigm from the Ålborg Charter

¹¹ These criteria were later followed up in greater detail and submitted in a report to the Ministry of the Environment in Norway (MoE 2000), from where they have gradually been circulated as standard benchmarks for assessing LA21 in Norway. The English version of these expanded criteria is available in Lafferty 2001b.

outlines in much greater detail the logical steps which both could and *should* accompany a commitment to LA21. Since there were numerous signatories to the Charter in all 12 of the SUSCOM countries, this meant that we also here could bolster the validity and legitimacy of our assessments in and through the open democratic commitments to the LA21 goals. The devices in question thus provided key operational criteria on the basis of widely sanctioned and internally consistent programme standards.

5.3 Performance benchmarks

Finally – and very briefly – there is a clear need in all forms of political evaluation for developing some kind of performance benchmarks. Differentiation criteria are necessary to make the first “cut” among competing relevant phenomena; and operational paradigms can significantly improve comparative monitoring; but neither device can help us to determine whether the glass is “half full” or “half empty”. Even accepting that we are operating on a very general and complex level of evaluation (systemic political performance), we still need some form of basic benchmark for strengthening the practical implications of the evaluative results.

Box 5: The Ålborg Charter “paradigm” for Local Agenda 21

- Information-gathering and consciousness-raising
- Interpretation and relativization of Agenda 21 to local conditions and problems
- Development of priorities and local action plans with both general and sector-specific targets
- Determination of appropriate steering instruments, including the procurement of voluntary agreements among sector-relevant social actors (“stakeholders”, “major groups”, “target groups”)
- Goal-specific procedures for implementation
- Monitoring and evaluation (of both the enactment process and its effects)
- Revision of goals, plans and initiatives

The Ålborg Charter, Part III

Turning once again to the SUSCOM project, we can conclude the overview of devices by considering the categories presented in Table 1. In contrast to the first two devices, which were largely developed prior to the actual monitoring and reporting of the implementation efforts, this final device summarizes some of the more significant findings of the comparative evaluation. The table gives expression, on the one hand, to an acknowledgement by the project of a need for differentiating between different levels or types of LA21 performance (the “Four P’s”): which in turn led to a broad categorization of the results in terms of 5 different “modes” of LA21 implementation. Taken together the results constitute a new set of “normative-empirical” benchmarks for LA21 achievement – with clear indications as to what will be the strengths and weaknesses of each mode. By employing the term “mode”, instead of “model”, we clearly indicate how and why the four dominant extant “modes” deviate from the Ålborg paradigm. The benchmarks in question thus reflect specific types of “outcome” from LA21 initiatives, expressed as deviations from the ideal type. The evaluative approach to LA21 is thus provided with a new knowledge base and point of departure for further empirical assessments, as well as new insights for revisions the action plan for sustainable development in this area.

Table 1: Stylized modes of LA21 implementation in Europe

Mode of implementation	<u>The Four P's</u> Emphasis is place on:			
	<u>Process:</u> New methods of mobilisation and cooperative governance for achieving SD	<u>Plan:</u> Adopting a strategic plan for SD with targets and indicators	<u>Policy:</u> Adoption of single-issue SD policies and programs	<u>Product:</u> Achievement of confirmed SD-goals and targets
Ålborg	Yes/strong	Yes/strong	Yes/strong	Probable/strong
Paternal	No	Possible	Possible	Possible/Weak
Integrated/Single-issue	Possible/partial	Partial/strong	Partial/strong	Probable/narrow
External/Forum	Partial/strong	Possible/weak	No	Unlikely
External/Fragmented	Possible/partial	No	No	Possible/narrow

Source: Lafferty 2001a: 292

6 CONCLUDING PERSPECTIVES: STRENGTHS AND WEAKNESSES?

The major strength of the type of political evaluation here described is that it is not political. Indeed, insofar as it *is* political, its quality and impact are weakened. The evaluation of political processes requires, in short, a studiously *apolitical* institutionalization and execution if it is to have broad-ranged and long-lasting impact. This is particularly true with respect to at least four crucial “mediators” of social change: the mass media, science, public administration and business. No meaningful progress towards sustainable development will be achieved without the positive engagement of each – and all four are particularly wary of politicized research. Insofar as evaluation of political implementation is perceived as politically biased in any recognizable party-political sense, one can guarantee that the effects of the results will be proportionate to the influence of perceived party beneficiaries. Not only that, but one will also contribute to a direct politicization of the programme, policy or issue in question, thereby enhancing the conflictual nature of the issue, resulting in yet higher thresholds for change. The negative prospects of such a development are particularly acute with sustainable development, since the issue has thus far enjoyed a very high moral status. Such an admonition does *not* mean, however, that evaluation of initiatives for sustainable development must be non-controversial. To the contrary, the more controversial the better, since change is nourished by controversy – but it must be controversy related to deviations from clearly documented political commitments, and clearly delineated efforts to achieve change.

To the degree that this sounds like an overly naive and relatively hopeless prescription for evaluative research in this area, we stand face to face with the potential limits of evaluation as a medium for change. The slightest concession towards the institutionalization of “deep structure” politics in and through evaluation practice, however “balanced” in relation to other similarly “disguised” practices, the more quickly professional evaluation will transcend, and ultimately sever, its contacts with academic policy analysis. The ability to maintain a proper disciplinary balance – and even more crucially, the ability to finance such a balance over the long run – is the outstanding challenges of the persuasion. They are, however, only marginally different from similar challenges in the natural, technical and economic sciences. It’s all part of developing a truly effective science for change: a science particularly suited to the democratic pursuit of sustainable development.

REFERENCES

- Berg, T. (1983): *Deltagerdemokratiet* ("Participatory Democracy"). Oslo: The University Press.
- Bemelmans-Videc, M.-L.; R. C. Rist and E. Vedung (eds) (1998): *Carrots, Sticks and Sermons: Policy Instruments and Their Evaluation*. New Brunswick: Transaction Publishers.
- Hagtvet, B. (1980): *Den Vanskelige Lydigheten* ("The Difficult Obedience"). Oslo: Pax Forlag.
- Hagtvet, B. and W. M. Lafferty (eds) (1984): *Demokrati og Demokratisering* ("Democracy and Democratization"). Oslo: Aschehoug Forlag.
- Lafferty, W. M. (1981): *Participation and Democracy in Norway: The "Distant Democracy" Revisited*. Oslo: University Press.
- Lafferty, W. M. (1984): "Workplace Democratization in Norway: Current Status and Future Prospects with Special Emphasis on the Role of the Public Sector". *Acta Sociologica* 27: 123-138.
- Lafferty, W. M. (1989): "Work as a Source of Political Learning Among Wage-Laborers and Lower-Level Employees", In Roberta Sigel (ed.) *Political Learning in Adulthood: A Sourcebook of Theory and Research*. Chicago: University of Chicago Press.
- Lafferty, W. M. (1996): "The Politics of Sustainable Development: Problems and Prospects after Rio". *Environmental Politics* 5: 185-208.
- Lafferty, W. M. (ed) (1999): *Implementing LA21 in Europe: New Initiatives for Sustainable Communities*. Oslo: ProSus and the European Commission.
- Lafferty, W. M. (ed) (2001a): *Sustainable Communities in Europe*. London: Earthscan.
- Lafferty, W. M. (2001b): "Local Agenda 21: The Pursuit of Sustainable Development in Sub-National Domains", In *Sustainability Assessment at the Local Level*, Dimitri Devuyst (ed), New York: Columbia University Press.
- Lafferty, W. M. and K. Eckerberg (eds) (1998): *From the Earth Summit to Local Agenda 21: Working Towards Sustainable Development*. London: Earthscan.
- Lafferty, W. M. and O. Langhelle (1999): *Towards Sustainable Development* London: Macmillan Publishers.
- Lafferty, W. M. and N. Raaum (1992): *Makt og Politikk med en Annen Røst* ("Power and Politics in Another Voice"). Research Report No. 02/92, Department of Political Science, University of Oslo.
- Martinussen, W. (1973): *Fjerndemokratiet* ("The Distant Democracy"). Oslo: Gyldendal.
- Mayne, J. and E. Zapico-Goñi (eds) (1997): *Monitoring Performance in the Public Sector*. New Brunswick: Transaction Publishers.
- MoE (2000): *Innføring av Lokal Agenda 21 i Norge* ("The Implementation of Local Agenda 21 in Norway"): Report prepared by Trygve Bjørnæs and William M. Lafferty, Program for Research and Documentation (ProSus), for the Norwegian Ministry of the Environment. Oslo: Ministry of the Environment.
- Nagel, S. S. (2002): *Handbook of Public Policy Evaluation*. London: Sage Publications.
- ProSus (2000): *Strategi- og virksomhetsplan for ProSus, 2000-2005* ("Strategy and Action Plan for ProSus, 2000-2005"): Oslo: Centre for Development and the Environment, University of Oslo.
- Rieper, O. and J. Toulemonde (eds) (1997): *Politics and Practices of Intergovernmental Evaluation*. New Brunswick: Transaction Publishers.
- Rist, R. C. (1995): *Policy Evaluation*. The International Library of Comparative Public Policy Vol. 3. Cambridge: University Press.

- Skogerbø, E. (1988): *Demokratiperspektiv på lokalfjernsyn: En studie av Jevnaker lokal-TV* ("Democratic perspectives on local television: A Study of Jevnaker Municipality"). Report No. 2. Oslo: Department of Media and Communication, University of Oslo. (Master's thesis in Political Science)
- Skogerbø, E. (1996): *Privatising the Public Interest: Conflicts and Compromises in Norwegian Media Politics 1980-1993*. Report no. 20. Department of Media and Communication, University of Oslo. (Doctoral dissertation in Political Science)
- Strømsnes, K. (1993): *Politisk Deltakelse I Norge: Variasjoner over et Tema* ("Political Participation in Norway: Variations on a Theme"). Master's thesis in Political Science, Institute for Comparative Politics, University of Bergen.
- Vedung, E. (1997): *Public Policy and Program Evaluation*. New Brunswick: Transaction Publishers.