

# Why Government and Non-Governmental Policies and Projects Fail Despite ‘Evaluations’: An Indicator to Measure whether Evaluation Systems Incorporate the Rules of Good Governance

David H. Lempert  
*Independent Consultant*

**Background:** While much has been written in the evaluation literature on the theory of evaluations and on specific cases, there is still no comprehensive and easy to use indicator that can be used to hold organizations to the principles of effective evaluation, to score their quality in several areas, and to offer an immediate diagnostic for improvements.

**Purpose:** The article offers an easy-to-use indicator for measuring whether organizational evaluation policies for government and non-governmental organization spending actually protect the public interest in line with good governance and management principles or whether they serve, instead, to protect bureaucracies and hide wrongdoing under the cover of an “outside” evaluation.

**Setting:** The primary focus of the piece is on international development evaluations but the author shows how the same indicator and model can be used for other government agencies as well as businesses, with modifications.

**Intervention:** The article examines failures of evaluation systems in light of the principles for quality and shows how an indicator can be used to measure and prevent those failures.

**Research Design:** The piece defines the principles of evaluation systems and accountability using both the frameworks of international agencies, themselves, and professional texts, places them in a model framework for effective evaluations, and then turns this framework into a set of questions to derive an indicator.

**Data Collection and Analysis:** The article offers a sample detailed test of the indicator using the United Nations Development Programme (UNDP) as a case study. Several other organizations of different types are also tested briefly to yield a variety of results on the quality of their evaluation systems.

**Findings:** Use of this indicator on several organizations including those of the UN system, the EC, U.S.A.I.D., reveals that a number of governmental agencies and contractors (and particularly in the field of international development) are actually failing to protect the public interest and are using evaluation processes as tools to cover up abuses and mistakes and to advocate or advertise for more funding.

**Keywords:** *evaluation, international development, transparency, accountability, monitoring, governance*

The lesson that I learned about the lengths government agencies would go to avoid public oversight was made clear in my first job with the executive branch of the U.S. government some 25 years ago. Even before I began, my interview for the position centered on my earlier experiences working as an intern for a United States Senator who exercised some budgetary oversight over the agency, rather than on my skills or goals for the work. The reasoning for the intense questioning about my relationship with that Senator became clear when I began work. “We wanted to make sure that you wouldn’t report back to Congress on how we use public money,” the department deputy told me. Shortly thereafter, I witnessed lawbreaking by the agency and flagged it in my first report to my supervisors in the belief that they would be eager to assure their compliance with laws and improve use of their funds. The reality was just the opposite. The agency then classified the report as a “secret” document so that I couldn’t even read my own work and to ensure that the public would never see it. Then, they told me that if I ever reported what happened that they would act to assure an early end to my career in their agency. From hiring to routine work activities, I learned the lesson that the priority of government employees was not serving the public but on evading it by circumventing the democratic controls of the U.S. Congress and public laws.

Now, with more than 25 years experience working with several U.S. government agencies as well as international organizations and foreign governments, ironically on projects that are designed to bring about “good governance” and to oppose “corruption” or to educate the public on their “rights,” I have become familiar with a wide range of

devices that government agencies, implementing organizations, and businesses routinely use to thwart oversight regulations and requirements for transparency and accountability in order to assure continued or expanded funding for their projects of dubious benefit.

I have learned from colleagues that my experiences are not unique (Wenar, 2006; Dakin, 2003). The majority of evaluations in “development” projects in many international agencies follow a single formulaic result, guaranteed by an apparently manipulated (or subconsciously homogeneous) process to yield a typical conclusion: “The transfer of money or skills to ‘poor’ beneficiaries made them ‘richer’ or “better off” than they were without the transfer. The project should be continued and expanded.” Questions about comparative benefit, about long-term consequences, about dependency, interference in cultural systems, violations of international law, inconsistency with organizational mission, or about specific financial abuses, mismanagement or waste are routinely eliminated from their Terms of Reference or from their final reports as “outside of the scope” of the evaluation. More likely, these are replaced with requests to explain ‘delays’ in the transfer and the administrative structure for the transfer of resources alongside “lessons learned” and ideas for the “next phase” with the very same implementing agents.

With evaluations like these, bureaucrats always win because the answer about whether transfer of money makes a recipient richer on receipt is always, “Yes.” With this logic, the result of an “evaluation” is pre-determined and no one is held to any measures of performance, with the shared understanding that asking for accountability would be “too difficult” or

would be “politically sensitive.” The real underlying goal is to turn evaluation into an advertisement for more money and power for the administrators, no matter how potentially harmful, how weak or unprofessional, how corrupt, how strongly in violation of international or local laws or development principles, or how self-interested the transfer. In evaluations like these, where spending does not focus on the root causes of a problem to be treated, where evaluators are not empowered to stop or replace projects or redesign failing systems, or to hold anyone accountable for performance, the public that funds the projects and the intended “beneficiaries” who need solutions to problems rather than subsidies to treat symptoms, are always the losers.

The above critique is simply firsthand evidence of several observers and it is easy for most agencies to wave off such criticisms as “subjective” or “uninformed.” To determine the reality would require a more objective tool that could both measure whether monitoring and evaluation systems fulfill their principles. Both critics and defenders currently lack such a tool that can test the critiques and constructively reveal exactly what is missing and what needs to be fixed to make an evaluation system work professionally.

This article offers exactly that remedy and device; a measurement tool that takes the textbook professional standard of what an effective management and evaluation system is supposed to do and translates it into an easy to use indicator to hold organizations accountable. It does that by providing measures for assuring adherence to three simple categories of management control, accountability and transparency, along with professional protection of the evaluator in his/her evaluation role. This standardized critique

scores evaluation systems using a total of 25 questions. Each question reveals a part of the system that may be missing or weak and that needs to be fixed to assure credibility and effectiveness.

The indicator offered in this article is a tool that government reformers, taxpayer advocates in the developed world, community activists in the developing world, public interest lawyers, and journalists can use to examine government systems that seem complex and difficult to understand, in order to measure their performance and turn those measures into clearly defined remedies. For professionals in development and in the evaluation field, who are more likely to defend the status quo than to call for actual measures of their work, this piece offers a reality check on public and professional expectations of their performance that can serve as a wake-up call. For students and scholars in management, it offers an additional framework for looking at management accountability and oversight systems.

The importance of this indicator is reflected by a bit of irony. The author of this article has begun to establish the architecture for effective evaluation systems in the field of international development including invention of several new peer reviewed and published indicators for evaluators to use to hold public agencies accountable for results that were claimed to be hard to measure, but that detail inconsistency between project spending and international law and development missions. Yet, few of these indicators are being used because indicators have little or no impact without an effective system to use them. Moreover, there is a likelihood that evaluators who use them could be blacklisted and any use of indicators at all could be suppressed by agencies whose

performance is being measured. In offering change to developing nations, the developed nations have a responsibility to look at their own systems first to assure that they are actually modeling and support what they claim to be providing in their interventions overseas; a phenomenon that this author previously referred to as “Building Democracy...with Mirrors” (Lempert, 1996).

Recent articles by this author have called for new organizations to monitor government agencies (Lempert, 2008a) and have offered indicators and benchmarks through which the public and organizations can hold international development actors accountable to international law and to their mission statements for their interventions. These “litmus-test” type indicators include an indicator for “sustainable development,” (Lempert and Nguyen, 2008) and other impact areas such as “democratization” (Lempert, 2008b) and “self-reliance” (freedom from donor “dependency” (Lempert, 2008c) among others (Lempert, 2009a, 2009b). The author has also begun to publish the ethics codes for professionals working in the field that protect public beneficiaries and legal requirements (Lempert, 1997).

This piece goes one step further to provide a measure of the quality of the key control systems that must be in place, in monitoring and evaluation, and offers a scoring system to rate them based on the real world failures that currently thwart accountability.

This article fits within the literature of evaluations, as well as of government structuring of good governance systems and how to measure and achieve them. It also provides a model framework of administration and control systems in international development and in general. For readers who work directly in applied

settings in international development, the indicator offered and tested in this article can be used to hold donor countries and non-governmental organizations to the very standards they claim to be promoting overseas for good governance but for which they have yet to clearly define measures.

The piece begins by defining the principles of evaluation systems and accountability as defined by international agencies, themselves, and by professional texts, so that these can be placed in an indicator that holds these organizations to their own standards. Following the description of principles is a brief summary of previous calls and attempts at designing an indicator, a section on the kinds of failures that now occur in the absence of any oversight indicator holding organizations to these principles, followed by a presentation of a new indicator that meets that goal. At the end of the article is an extremely detailed examination (a step-by-step introduction) to using the indicator on an organization like the United Nations Development Programme (UNDP), as well as on several other governmental and non-governmental organizations.

Note that some readers may find the theoretical portion of this article a bit difficult since it combines work in various disciplines of management (public and private) and law. That, in itself, serves as a wake-up call to those who are currently involved in the field of evaluation, that they need professionals with full knowledge of these areas to build an effective evaluation system. Those readers who have difficulty with the theoretical material are advised to turn right to the indicator itself and to read the 25 questions. These questions are relatively easy to apply even for non-experts since they apply the principles in practical ways.

## Principles of Evaluations and Accountability

The basic principles for effective evaluation systems are clear and a recent study by the World Bank's Independent Evaluation Group defines these key areas as (World Bank & Mackay, 2007):

- strong incentives for conducting M&E (monitoring and evaluation) and using the information and
- structural arrangements to ensure objectivity and quality

It seems to go without saying that a third prong in addition to incentives and objectivity is the existence of the standard management controls and reporting to provide quality information that evaluators can review (for which they need the “strong incentives” and “objectivity and quality” mentioned above). The functional skills and tasks that are essential can be found in any textbook of organizational management, though they are usually designed for private managers rather than adapted directly to public organizations.

These three principles can be described more clearly as follows, with the skills and tasks placed before the World Bank's two principles:

1. Establishing a functioning management control system: Utilizing proper evaluation roles to provide information that the “owners” of the project need to review their spending. These are standard management control functions and reporting that provide quality information for review.

2. Assuring incentives for transparency and accountability of the evaluation so that information reaching the “owners” is objective and appropriate: The approach to ensuring strong incentives for conducting evaluations generally requires some kind of independent, enforceable, watchdog function over the evaluations, to protect the “owners” of the project; and
3. Protection of professionalism and political insulation of evaluations: These are the structural arrangements to ensure objectivity and quality of professionals in their evaluation roles.

The basic principles and measures of quality in each of these three categories are also relatively well known, though they may never have been collected and presented together for measuring the quality of evaluation systems. They can be elaborated further, below.

### *Principles of a Functioning Management Control System: Measurement for Managerial Control*

It is interesting that while there are standard evaluation frameworks in fields like international development with a recurring set of concerns—efficiency, efficacy, relevance, quality of design, overall value, sustainability, etc.—these are not directly linked with management literature to make it clear what specific *systems* need to be in place to collect the information and perform the functions that are to be evaluated. (The standard frameworks of the UNDP and the European Commission can be found easily at: UNDP Evaluations Office, 2002,

Guidelines for Outcome Evaluators, p. 10; Europeaid project evaluations, 2007). If the professionals who are supposed to be making judgments about efficiency (usually accountants), design (technical professionals), relevance (legal experts and business strategists), and so on, are not the appropriate personnel for the specific measures required at each step, the determinations they make will be entirely inconsistent with the professional needs of the organization. While businesses are not likely to make this mistake given their incentives, there is often no link at all between the types of measures and the types of professionals needed to make those judgments in development organizations, and one evaluator with only one profession may be making decisions in several fields where he or she has no training.

Standard management textbooks to train business and public managers for effective measurement and control of their organizations are clear in defining the kinds of systems that need to be in place for an organization to run efficiently and for it to be subject to controls that assure its objectives are best achieved. (See, for example, the standard textbooks in "Managerial Accounting" (Garrison, Noreen & Brewer, 2005) and in "Accounting for Management Control" (Emmanuel, Merchant, & Otley, 1990).

Table 1, Principle I, takes the set of evaluation concerns required by development organizations like the UNDP and the EC (referenced above) and places them next to the standard management functions and departments that routinely measure these concerns as a way to distinguish the specific professional functions that are required in an effective management control system. This table may seem foreboding at first, since it takes a massive amount of information

and tries to summarize and short-cut it, but it is an easy way of presenting and seeing a large volume of information to see the parallel needs of different types of organizations to hire specific professionals for particular kinds of evaluation measures.

The table compares three different kinds of organizations: international development organizations (the first two rows), businesses (the third and fourth rows) and government and civil society organizations (the fifth and sixth rows). In fact, development organizations are just one type of government and civil society organization and the material in the fifth and sixth rows expands in detail on the functions and specific information needs for effective monitoring in international development organizations as well as for any government or civic organization spending. Those who are most familiar with business or government organizations can start in those sections and look at the comparisons. Overall, the final conclusion of the table as to the specific professionals needed for different parts of the flow chart of activities (the required oversight and measurement functions in organizations) is presented in the final row at the bottom.

To walk through the table, those who are familiar with the evaluation measures for development organizations should start with the second row as a way of understanding how the first and second rows work together. This row (Evaluation Categories/Questions of UN/EC/USAID Systems) presents the series of questions and information sought by development organizations in evaluations.

Look next at the top row. The top row does is takes the standard questions for evaluations in development organizations from the second row and shows that these questions actually follow the flow of

activities that a development organization is supposed to follow in designing and managing its projects. Each evaluation question set should be and is linked to a specific activity or function that is presented in the first row. The arrows show that these functions are linked in a process or flow and that the entire process is really part of a feedback loop. Ideally, information from an evaluation leads to a check on the organization's mission and to new designs and project innovations.

Note that no matter what kind of organization to be monitored, evaluations regularly correspond to seven different and independent lower management systems in the cycle of operations of an organization. These range from strategic planning and mission setting and protecting to design and innovation, purchasing of inputs, production of outputs, outcome results, financial and business planning for sustainability and growth, and research and development. Depending on the type of organization, they have different names, but the concept of each stage is the same. Looking down the columns at the different organizations is useful as a check, particularly on public organizations, on the professional role of each of these stages.

Overall, the rows describe the kind of information that is essential to be routinely collected and then made available to evaluators for the evaluation to have any real meaning or purpose.

The final row of the table makes it clear that each of these sub-systems is not only distinct but that successful evaluation of organizational performance requires specific professional skills to match the needs of each of these sub-systems. Asking evaluators to review systems for which they lack the wrong professional expertise (a common problem in evaluations that assume one

evaluator can do everything), providing too little time or resources for effective oversight, or simply asking evaluators to review systems that do not even exist as distinct in the organization are clear signs of evaluation failure. Reading across the rows it is easy to see that the necessary expertise includes business planning and law, managerial accounting (different from cost efficiency accounting though most development organizations don't recognize it), auditing and personnel auditing, cost accounting (quite different from auditing, though most development organizations don't recognize it), public accountability/ law, financial investment, and social science, in addition to sustainable development training and other project technical skills.

Understanding this alone is the key to understanding what constitutes quality and why most measurement systems fail at even this most basic step of professionalism. It is easy to start with this set of principles of what an evaluation needs to do and what measures and systems need to be in place, as a very quick means of determining whether the evaluation system is meaningful or not.

**Table 1**  
**Principle I: Design of a Proper Management Control System**

<i>Types of Organizations (Rows) and Evidence Required for that Type of Organization (Alternate Rows) to Fulfill the Function</i>	<i>Development Organization Functions, (Generalized to be Applicable to Other Organizations as Well)</i>	<i>Flow Chart of Functions Following the Cycle of Activities of an Organization</i>						
		1. Mission and Strategy; Defining the scope and advantage of types of interventions	2. Design and Development of the intervention and the implementation/ management system	3. Inputs (Purchase (Procurement))	4. Intermediate Inputs/ Outputs Produced	5. Outcomes Result	6. The Outcomes Generate Resources to Sustain and Maintain the System and/or Expand it	7. Changing Competitive Environment/ Needs: Measures on Value to Influence Design and Strategy to Next Phase
<i>Measures Needed -- Evaluation Categories/ Questions of UN, EC, USAID systems</i>		<i>Relevance of the Intervention: Organizational contribution and role; Relation to Organizational Mission; Relationship with partners</i>	<i>Project Design; Efficacy of the Systems and of Management</i>	<i>Efficiency (of Inputs)</i>	<i>Efficiency (of Outputs)</i>	<i>Overall Value in Sustainable Human Development; Relevance; Public Participation of Beneficiaries</i>	<i>Sustainability/ Institutionalization and Self-Replication of Models</i>	<i>Best Practices and Lessons Learned</i>
<i>Business Functions and "Evaluation/ Valuation" (by Analogy)</i>		<i>Strategic Plan including the right management system for the organization as a whole</i>	<i>Product Research and Development including the organizational management system for the product/service</i>	<i>Book-keeping; Cost Control; Bidding; Personnel management</i>	<i>Intermediate and Final Product Produced with Quality and Cost Control</i>	<i>Balance sheet shows asset growth and investor returns</i>	<i>Long Term Business Plan that shows how growth or maintenance is funded (re-investment)</i>	<i>Market Research and Development</i>
<i>Preliminary Measures Needed</i>		<i>Market Segments and Branding</i>	<i>Product Technology and value to consumers</i>	<i>Accounts: Cost Control Measures for Purchase/ Comparative Costs of labor and inputs</i>	<i>Cost Savings through Efficiency/ Scale/ Investment/ Process</i>	<i>Sales and Profits</i>	<i>Financial measures of assets, returns, reinvestments and depreciation</i>	<i>Focus groups; Surveys; Product Testing</i>
<i>Government or civil society Functions in development (recognizing the different missions of both but similar systems)</i>		<i>Upper Level Organizational Management: Mission Statement of the Function (Promotion and</i>	<i>Professional Expert Project Design: Intervention logic solving a problem of value to</i>	<i>Book-keeping; Cost Control; Procurement systems; Personnel management</i>	<i>Professional quality review of the output or service</i>	<i>Increased/Stable Per Capita Assets, Sustainable Development and additional sub-goals that</i>	<i>Tax Policy (Government) or Beneficiary funding (NGO) and Systems for social marketing of models fit into a</i>	<i>Social science research and development</i>



	Protection of an Asset for Government; Change in Government Action for NGOs) and of the Specific role and comparative advantage of the unit in achieving it and in allocation of funds	beneficiaries including the right management system			promote these for Beneficiaries and Taxpayers with an appropriate Sustainability Plan for the social/cultural unit	long-term financial plan	
Preliminary Measures Needed in Design Phase to Make Evaluation Useful and to Hold Bureaucrats/ Implementers Accountable	Ability of the governmental system to function and of the unit to achieve the mission, with detail explanation showing how the intervention directly contributes to achieving the mission in ways that justify the spending on this intervention compared to others (strategic value) for the organization	Expert problem analysis with root causes; Sustainable development framework; Logframe relating intervention to specific factors in the solution; Justification for why the specific intervention is the most strategic use of resources to change specific behaviors and incentives that will change outcomes	Comparative Costs of labor and inputs	Benchmarks: Cost Benefit Analysis and Return on Investment per specific unit of professional quality in the relevant categories using specific professional indicators	Sustainable development plans, systems, indicators and per capita asset measures for the cultures and comparative to other interventions; Measures relevant to citizen beneficiaries that they can measure; Impact screening and analysis models	Measures of the value of the benefit and how beneficiaries are taxed to promote it in an independent way and within their resource and asset base; and Measures of the information system and reinvestment in new technology to avoid dependency	Control groups, hypothesis testing using the scientific method
Type of Evaluation/Evaluator Skills to be combined with Technical Expertise in the area of the intervention [Where most evaluation systems fail to recognize expertise]	High Level Strategic Planning Consultant; Organizational Management consulting; Governance and Legal Specialist for government functions	Managerial Accounting/ Results based management	Independent Cost Control Audit with random spot checking (different from a due-diligence book-keeping audit)/ Purchase Accounting and Personnel Cost management	Cost Accounting	Watchdog Role with accountability to taxpayers, to beneficiary groups, and professional measures relevant to development (sustainable cultures, increased per capita assets)	Systems Analyst, Financial systems analyst of public spending, Social marketing systems analyst, Innovation investment analyst	Social Science and Indicator Research

### *Principles of Transparency and Accountability of the Evaluation Role, Itself*

There is plenty of literature on the legal principles of transparency and accountability, alongside the literature of management control. In legal doctrine, transparency and accountability are parts of (management) control and “agency,” described by relations of “principal and agent” or “master and servant.”

The basic principle is simply that the decision-maker (in the case of public organizations, the “taxpayer” public; in the case of NGOs, the citizens who donate) must be assured that it receives all of the information it needs in the form that makes it useful. In the case of international development, where there are recipients of aid, laws also define the rights of “third party beneficiaries” (the aid recipients who are members of the public in the developing country) as that of being owed “duties of care,” who also have rights to information and oversight. The duty of care is the duty to provide information to these citizens who receive the services or could be harmed by other actions taken by the project. The project management is the agent (the government agency or the non-governmental implementing agency) and their agents are the people they hire for technical work (professionals and partner implementing agents). Each of these four types of actors has different roles and responsibilities in an accountable and transparent management and evaluation system.

An accountable management control system in international development add two groups of actors, the “taxpayer” public and the aid recipients in the developing country, to the system of oversight in a private business activity (the “taxpayer”

public and the aid recipients in the developing country), and one group to the system of oversight ordinarily find in domestic public administration (the aid recipients in the developing country). In a shareholder owned corporation where the owners and managers are different, the shareholders/owners take the place of the “taxpayer” public, and government takes the place of the community in protecting against any harms of business activity, with the other principles the same. These two groups (the “taxpayer” public and the aid recipients) are the real decision-makers, expected to have the power of oversight in international development projects. Their role is that of watchdogs and they make the decisions on spending and on programs.

Within the project management, in the implementing agency and among the professionals and partner implementing agents they hire (including for evaluations) the principles for assuring accountable and effective oversight are simple and clear.

- Management (i.e., the control agency of the evaluation system on behalf of the taxpayer public and aid recipients in international development; on behalf of shareholders/owners in a business) must be completely separate from the project/agency offices under review without conflicts of interest (i.e., not belonging to same union, responding to another common manager above them, and having no shared interest in project funding or continuation or favorable review) and they must answer to higher authority with real power, including budgetary control or the power of removal/firing rather than simply

one in name that has actually delegated power back to the managers.

- Evaluations must feed into a system with performance based incentives from the design to the review stage of operations, such that results of the evaluation lead to rewarding and punishing specific employees and offices on the basis of performance. These measures of performance need to be clearly defined so that punishments and rewards have a rational basis linked with organizational objectives.

Table 2, Principle II, helps to generate a list of the measures and safeguards that an evaluation system needs to ensure accountability and transparency for different actors in different roles (the third column of the table). This list can then be used to test the quality of an evaluation system.

Though the table may seem packed with information, it is also simple to follow. The left column lists the four types of actors, identified above, who are the different stakeholder-managers of an evaluation and who have legal obligations on public spending for international development projects. The second column describes their general role in assuring effective, accountable, transparent oversight. Note that for a “Government implementing Agency” the roles are exactly that list of functions that were listed in the previous table, Principle I. The key detail information is in the third column, defining exactly who (what specific professionals in what kinds of organizations) can assure performance of the role and how they do it. This column

describes the specific duties of care as well as the specific mechanisms that need to be functioning in order to assure information transparency and accountability related to the function of evaluations. Finally, the fourth column describes what happens when a specific part of the oversight and control system fails, as a way to help diagnose the origin of problems in a co-opted or corrupted evaluation system.

For those who want a summary of the information, rather than follow all of the detail, the key general ideas that can be summarized from the table and applied directly for international development projects (with analogies to other public projects and to business for those who wish to make other uses of the table) are that:

- The “Taxpayer” Public and Aid Beneficiaries must be able to use the information that evaluations provide as a real control mechanism and not for just an information network or “communities of practice” among the implementers;
- Some controlling authority directed by the “Taxpayer” Public and Aid Beneficiaries must set proper timing and funding for the evaluations, based on principles of cost/benefit (matching the costs of the evaluation to the benefits to the Beneficiaries), so that the tasks of review are regular and appropriately funded; and
- The Information system, itself, must meet certain quality standards so that useful and clear information reaches *all* of the oversight groups and is protected from interference.

**Table 2**  
**Principle II: Transparency and Accountability of the Evaluation**

<i>Stakeholder Interest group in the evaluation with responsibility for managerial control and oversight (with Legal Term to describe them)</i>	<i>Role (Building on Principle I Table)</i>	<i>Who Assures Accountability to the Beneficiaries and Public Contributors and How</i>	<i>Implications of failure of this system</i>
Public beneficiaries of the aid ("Third Party Beneficiaries" owed a "Duty of Care")	Watchdog role/ Direct participation and ability to assure the project is transparent and accountable	<u>Who:</u> Requires a protected civil society (NGO and Media) with investigatory powers and public accountability of civil society through public participatory court and judicial processes. <u>How:</u> Direct and transparent oversight of the government donor, of the implementing partners and of the government body acting on behalf of the public.	If this system is absent or weak, none of the M&E by the stakeholder implementing agents will improve or be held to standards and there is a likelihood of abuse, corruption, and subversion of the stated agenda.
Public contributors to the donor (taxpayers or contributors) – (The "Master" controlling the "Servant" "Agency")	Watchdog role/ Direct participation and ability to assure the project is transparent and accountable	<u>Who:</u> Requires a Parliamentary our outside watchdog agency with investigatory powers and public accountability. Requires a protected civil society (NGO and Media) with investigatory powers and public accountability of civil society through public participatory court and judicial processes. <u>How:</u> Independent evaluation findings communicated directly to the public through standards and benchmarks, with potential to change funding.	Same as above
Government implementing agency with the power to hire specialists (following the functions of the previous table) – (The Agency)	1) Protecting Organizational Mission and Strategic Planning to fulfill it; a) Interventions screened to assure most strategic promotion of the mission: Fulfillment of a specific accountable public function to protect and promote development of a resource in a sustainable way on behalf of the public beneficiaries and contributors. b) Advocacy for more projects on the basis of innovations and measures	<u>Who:</u> - High Level Management Trained in Strategic Planning with regular exercises; <u>How:</u> - Organizational Management reviews; <u>Who:</u> - Governance and Legal Review to assure accountability and implementation of government functions <u>How:</u> Review and screening of all interventions to assure compliance with the mission and comparative/ strategic advantage	The organization becomes a tool of the donors and its own bureaucrats, setting an agenda and choosing projects that do not advance its mission on behalf of the public and substituting hidden agendas or choosing interventions randomly and inefficiently with weak results at best.
	2) Professional Project Design and of Management Systems for Implementation:	<u>Who:</u> Professional Experts in a Unit for Research and Development of interventions or coming from the outside. (Measures include those listed in the table on the Activity cycle of organizations.)	Project design and standard setting offer no real measures for the evaluation that have a professional basis and advocacy is on the basis of conservatism and hidden

		<u>How:</u> Support should be 5% of spending on the project and in addition to any project manager/ administrator	agendas rather than results, resulting in failures and waste.
	3) Cost Control a) Cost Accounting / Bookkeeping in Procurement Systems and b) Independent Watchdog Audit for Efficiency of Inputs	<u>Who:</u> - Internal procurement systems and personnel systems; <u>Who:</u> an independent comptroller/ auditor doing spot checks for waste <u>How:</u> - External, independent Auditing and accounting of inputs (should be 2% of spending) by	Theft and waste
	4) Managerial Accounting for Efficiency of Outputs for Measurable Levels of Technical Outputs (Quality per unit cost)	<u>Who:</u> Key Function of Independent Evaluators in a Technical Area, but relies on professional project reporting and data collection from the design phase. <u>Who:</u> Upper level management must understand skills of Managerial Accounting <u>How:</u> assure that project designers understand quality measures and have offered real benchmarks of output value that are quantitative and scored if qualitative, so that there are measures of technical value per unit of spending	Waste and degradation of interventions and a result where advocacy is on the basis of connections, lobbying, politics, and sales/marketing rather than results
	5) Measuring Outcomes in Professional fields and in additional categories of impacts (social, cultural, resources) – a Managerial accounting role linked to Step 2)	<u>Who:</u> Public Watchdog role in coordination with objective evaluators offering information and exerting incentives on management to achieve high standards through: <u>How:</u> Professional measure of changes in behaviors, relative powers, and shifts of assets to promote sustainable human development balances, including the measures of unintended effects scaled for an overall measure of benefit/ performance (and losses) to individuals/ the system per unit of spending	If unintended outcomes are hidden or not measured (particularly long-term impacts/ externalities), if the wrong measures are substituted for real development goals, or if measures are poor or false, projects may be doing more harm than good
	6) Strategic Planning for sustainability/ maintenance/ growth: - - Systems Analysis, - Financial systems analysis of public spending, - Social marketing systems analysis, - - Innovation investment analyst	<u>Who:</u> Upper Management <u>How:</u> - Benefit evaluation and tax/ donation policy linking contributions to benefits - Social marketing to promote replication	Projects are just quick fixes that are treating symptoms and may be creating dependency.
	7) Research and Development	<u>Who:</u> Social science professionals accountable to professional standards and the public through: <u>How:</u> Analysis to lead to design of improved indicators and better interventions through scientific methodology like control groups (should be 5% of project spending)	Project design and standard setting offer no real measures for the evaluation that have a professional basis and advocacy is on the basis of conservatism and hidden agendas rather than results
Professionals as implementing partners (including evaluations specialists) -- (Sub-contractors or Agents of the Agents)	Professional Expertise: Outside professionals in specific categories can fill the roles in the government organization in the 7 categories above with the role of meeting professional standards in these roles and to advocate for more work on the basis of professional quality.	<u>Who:</u> Internal systems of professionals similar to those in government organizations (7 categories above) as well as through professional organizations and the public. <u>How:</u> - Maintenance and protection of professional standards and objectivity of work through ethics codes, unionization, legal protection, certification, professional peer review, and contracting procedures that assure these in several different ways such as direct contact and accountability to beneficiaries and independent oversight bodies	Collusion occurs between the agency being evaluated and the evaluator, leading to a situation where results are politicized/pressured/falsified causing the implementation or evaluation process to serve only as a rubber stamp or advertisement for previous decisions and hidden agendas

### *Principle for Protecting the Professionalism and Objectivity of the Evaluators*

The final principle for assuring effective managerial oversight and evaluations—that of protecting the professionals who perform the evaluation, measurement and reporting functions—is the simplest one to state, though possibly the hardest one to implement. Protecting the professionalism and objectivity of the work of evaluators requires a similar set of safeguards to those for protecting the quality of work of any professional. Selection procedures must meet professional standards that adequately test the range of skills needed, the professional should meet certain standards of certification, and the conditions of work must protect the freedom and objectivity of the professional by contract, contract enforcement, and conditions.

Professionalism requires looking at the cycle of the evaluator's work and assuring that there are protections at every step, from hiring to certification, to setting the standards and conditions of the work, to protecting against pressures of firing without cause, to assuring appropriate measurement of performance and rehiring without politicization. Though the different steps and stages that require protection are well recognized, the standards for achieving this and the measure of quality are continually revised based on understandings of legal mechanisms for protecting professionalism. Some of these can be listed in an indicator to measure at least basic compliance with accepted procedures.

*Previous Attempts to Design an Indicator for Measuring the Quality of Evaluation Systems.* Though there seems to be overall agreement among those who write about evaluation policy that, "It is a good idea to regularly monitor and evaluate the way the M&E system has developed and how well it is performing" (World Bank & MacKay, 2007, p. 2), most attempts to design a monitoring tool to test the quality of evaluation systems fall short. Many authors define the process and categories to examine, but fail to focus on the specific, substantive needs of a working system. None offer a clear overall indicator, though some do focus on how to measure performance of individual sub-systems with some guidelines that fit the three principles above.

*General Attempts to Design an Indicator for Evaluation Systems.* Some analysts of evaluation systems have described the range of "types" of monitoring and evaluation systems in different categories of approaches, with strengths and weaknesses as ways of classifying them and moving towards a diagnostic, but without offering a useful indicator.

House (1978), for example, has identified a number of types of evaluations, including

- politically controlled evaluations to promote an agenda
- public relations evaluations to advertise for funds
- research evaluations
- objective based evaluations
- decision oriented evaluations (for managers); and
- accountability for performance, between consumers (public) and implementers

Yet, identification does not help to determine whether the common requirements that apply to all types of evaluations are being met, and labeling evaluations may not be that useful. In fact, no governmental evaluation system (and probably no business or civil society organization evaluation, either) should ever be a public relations evaluation or a politically controlled evaluation. These descriptions are really descriptions of failures of the evaluation system and its subject to capture for different reasons. A good indicator should actually be able to distinguish and label evaluations based on weaknesses. House does not provide one.

Many different articles and books describe the “steps” (considerations and political processes) involved in building an evaluation system, while never defining what exactly needs to be built and how its quality is measured. Among the 7 or 8 or 10 steps, experts usually suggest actions like “doing a preliminary study” and “developing indicators” after “agreeing on what needs to be measured” but not describing how to protect the system of measurement or actually offering such indicators. (cf., Kusak & Rist, 2001).

Mackay promotes the idea of a diagnostic and offers a list of 8 areas to consider, but does not define the key principles of success that could be used for scoring in such an indicator. He suggests reviewing the history, the departments, and the process and even whether or not there is “donor support” (which seems to reverse the principle of accountability rather than implement it). Rather than offer a way to evaluate an evaluation system, with a list of key requirements of that system, he offers “success factors.” Among the 14 elements he notes are “substantive government demand” and whether there is a

“diagnosis of existing M&E as a first step” as well as whether an evaluation is “centrally driven by a capable Ministry.” This is a process oriented approach that potentially needs to an endless loop without defining the principles or substance. It also appears to be a self-serving guarantee for the need to hire consultants because there is no way that citizens or consultants can use the process, themselves (World Bank & MacKay, 2007).

The typical book on “building capacity” of evaluation systems similarly ends with suggestions for the “need to educate” for routine and quality, but without defining specific standards and without offering measures (Boyle & Lemaire, 1999).

While there are several organizations that are seeking to coordinate standards in specific fields for measurement of different kinds of government services, they have also failed to define systems of accountable evaluations or to provide standards for measures. Among these are: the International Development Evaluation Association: <http://www.ideas-int.org/>, and the International Organization for Cooperation in Evaluation (IOCE): <http://internationalevaluation.com/index.shtml>, including its partner the American Evaluation Association/American Journal of Evaluation, and other national associations.

*Indicators of Evaluation Quality for a Functioning Management Control System: to Achieve Measurement for Managerial Control.* There seems to be the only indicator currently available to measure whether evaluation systems fit overall managerial control needs and it is not particularly streamlined to identify the 7 different management control functions that are outlined in the table,

above for Principle I. A team of evaluators has developed a sophisticated “evaluability” indicator that rates the performance of an organization in professionalizing systems of measurement that evaluators can use (Poate et al., 2000). They ask whether an organization has sufficiently clarified its objectives and its measurements such that clear measurements are even possible. This is an important part of examining the evaluation process. However, it is only one part. Even if a project is “evaluable,” the evaluation may be irrelevant if its results do not feed back into management, if the evaluation is not transparent, or if the evaluator is pressured and the process is corrupted.

*Indicators of Evaluation Quality for Accountability and Transparency.*

Various governance indicators already exist in the field of “democratization” though they are not particularly applicable to evaluations. Among them are those that include shadow measures of accountability of organizations or that seek to measure attributes like public voice or government “effectiveness.” None of the specific indicators offer enough detail to measure the effectiveness of different watchdog and beneficiary roles for holding public and private bureaucracies accountable. The author has discussed some of these measures in an article on democratization indicators that can be summarized briefly as follows (Lempert, 2008b).

Freedom House’s “national democratic governance” category, among its measures of “freedom,” for example, implies that it measures accountability of government bureaucracies along with overall separation of powers. However, Freedom House’s scoring system does not list attributes of a system that characterize

accountability. Instead, scoring is based on subjective indicators in categories of performance (Freedom House, 2008, p. 3). The indicator scores are based on broad, subjective questions such as, “Does a competent and professional civil service function according to democratic standards and practices?” without listing those standards.

The World Bank’s governance standards are also subjective, relying on surveys of concepts like “voice and accountability,” asking questions like: “How much do you trust the Parliament?” and “How satisfied are you with the way democracy works in your country” and with “Free and fair elections” (Kaufman, Kraay, & Mastruzzi, 2007, p. 44). This is an indicator of perceptions, not of the quality of mechanics of an evaluation system.

A third organization, Citizen Report Card, offers a citizen watchdog tool as an example of the citizen oversight role and how it can be made effective, but the Report Card cannot be used as an indicator. Citizen Report Card bases its concept of accountability on whether an organization relies on user surveys but not on whether it shares information with the public and is accountable to it (Citizen Report Card, <http://www.citizenreportcard.com/>).

*Indicators of Systems for Protecting the Professionalism and Objectivity of the Evaluators.* There do not appear to be any current indicators in this area.

*The Wrong Focus.* The problem of failed evaluation systems, particularly in governance, appears widespread and many of the harms are easy to spot. Unfortunately, while the three principles for quality evaluation systems are simple and not so difficult to measure or



estimate, many evaluation systems avoid even the basics.

Failure occurs at all three levels: lack of overall management control systems in the organization to ensure real evaluations; the lack of an effective public watchdog role to oversee or improve the process on behalf of beneficiaries and taxpayer contributors, and the lack of professional safeguards of the evaluator's independence and professionalism.

The information presented thus far in this article can be summarized as follows, with clear implications for where failures occur. For organizational evaluations to be effective there need to be roughly 8 to 11 levels or "systems" that are working, depending on how one counts (seven different management systems within the organization that were listed in the first table, and up to four other systems of accountability to the four different groups of stakeholders listed in the second table, above). Each of these systems requires a specific kind of expertise. Despite this need for multiple professional skills, international donor assistance project evaluations often rely only on one evaluator to serve all of these multiple interests and trying to assure their effectiveness, with a minimum of funding! Moreover, the basic systems on which evaluators need to rely for information, may not even exist. In most organizations, audit and accounting systems usually exist in a basic form, but the lack of other management systems for other necessary functions means that evaluations are already severely handicapped before they begin.

Where Principle I is undermined and internal systems to routinely collect and process information that supports their functions are missing, evaluations are often transformed into exercises in fiction writing with the goal of creating measures

after the fact, where none exist, under pressure to justify a project. Lack of standardization—making every project an independent case study without any measures under the assumption that every project is "unique" or incomparable or that measures are "too difficult to create" (assertions that are rarely made subject to professional challenges)—undermines the concept of management control and evaluation. With no basic data or systems to review, evaluators are often pressured to just keep their eyes closed to the absence of information and systems and to just assume good faith and appropriate functioning even though the failures are evidence that the opposite is true.

Where Principle II is overridden and there is no effective watchdog role and procedures for accountability, evaluations often become advertisements for bureaucrats and the people they hire. This reverses the relationship of master (taxpayer public, aid beneficiaries, and their designated managers) and servants (implementing agents); with the servants setting the rules and evaluating themselves; declaring that they have done an excellent job and deserve higher pay.

Where Principle III is overridden and professionalism lacks safeguards, evaluators are pushed towards a role of propagandists and fundraisers.

While some evaluation systems are unavoidably weak as a virtue of organizations being small and unable to hire a full range of qualified professionals for an entire array of measures and managerial functions, the failure in larger organizations, and particularly in government systems, is hard to excuse. In larger organizations where failures occur, they are often the result of the weakening the evaluation process by design; increasingly through sophisticated

measures to corrupt, distort, and falsify any attempts at oversight and accountability. Corrupted evaluation systems may have any one or more of these characteristics:

- Hiring process and structuring of the evaluations, that deliberately under-fund or under-represent the necessary expertise (the several different professional skills) that are necessary to ask appropriate questions, to ferret out or reconstruct key data, and to appropriately apply the different types of professional analyses that are required.
- Transference to the private sector of public sector functions of evaluators in order to enable bureaucrats to directly (and intentionally) strip evaluators of public protections (free speech; whistle blowing; public and legal reporting) that they would have if they were working directly for government;
- Monitoring and pressuring of evaluators in ways that impede their ability to freely exercise their professional responsibilities, sometimes under the rationalization that asking the right questions is “an insult” to those whose failures (or corruption) are being exposed.
- Direct suppression, falsification or replacement of honest and objective evaluations;
- Firing, blacklisting, or harassment of professionals
- Anti-Fraud or Ombudsman agencies that lack real power to hold anyone accountable or are tasked only to target outsiders or

embezzlement, not protection of projects and the bureaucracy; and

- Anti-fraud or hotline offices that are within the same agency, part of the same union, in the same buildings, etc. as staff being evaluated (an example of the absence of independent systems that are legally enforceable or backed by any public or professional bodies to protect the process).

In fact, it is possible to use a relatively simple indicator to separate failures from successful projects simply by looking at whether the fundamental principles are appropriately measured and incorporated into a system of objectives.

*An Indicator to Measure Effective Evaluation Systems.* The indicator presented below can be used easily, even by non-experts, as a litmus test of the quality of an evaluation system, differentiating quickly between effective and ineffective (or hidden agenda, “propagandistic”) evaluation approaches. By asking a set of 25 “Yes or No” questions in three different performance areas, tallying up the results in each of the three areas, then reading the final determination on a special grid (see “Scoring Table”), it is easy to score the relative quality of any evaluation system and reveal its weaknesses.

The three categories of performance, that can each be scored, follow the three Principles described in the first part of this article.

- I. Fulfillment of the functions of a management control system: Utilizing proper evaluation roles (divided into two scoring categories) (11 questions);

II. Quality of transparency and accountability of the evaluation: Whether the system effectively serves as a management watchdog (for non-governmental organizations) and government watchdog and public protector (for both governmental and non-governmental organizations) (6 questions); and

III. Ability to safeguard professionalism and politically insulate evaluations from outside pressures and conflicts of interest (8 questions).

While this indicator is designed for use in the most complex type of organizations—international donor assistance—it can also be used on domestic government agencies, private businesses and civil society organizations (with some modifications). Those who wish to use the indicator on business organizations should note that the approach should be slightly different on the second category of questions, relating to transparency and accountability, where there are different stakeholder groups (corporate shareholders as well as the overall public). To measure the quality of a business or civil society in responding to its shareholders or members, simply insert this category in the category of “beneficiaries” and ask the same set of questions in Category II, a second time. To see whether a business or civil society organization is accountable internally (to its shareholders and members) and externally, to the public that charters business and non-governmental organizations to operate and that has the power to regulate them, the same questions must be asked again from the perspective of citizens as beneficiaries. In other words, there is a need to score

accountability to shareholders, first, then to score accountability to the public, second, to measure whether the evaluation systems are working effectively to hold the organizations accountable to the public for issues of concern to the public, as well.

Before beginning the scoring, note that the indicator is not an absolute scale to measure the quality of each internal system, but offers a checklist of whether procedures exist at all. The indicator is not offered as a social science research tool but, instead, as a project evaluation and selection tool. It is best used to show the relative value of different projects, with some leeway offered in judgments. Note that the indicator is “relative” rather than “absolute” and each person doing scoring might score higher or lower, but should produce similar overall relative measures. Like most indicators, answers to each question would need to be “calibrated” if the goal were to assure that different observers make the exact same determinations. To do so would require a longer manual for standardized, precise answers across observers. To try to help calibrate the indicator as a guide for scoring, a detailed example of scoring is provided at the end of this article and several organizations are also scored.

After coming up with individual scores (“Yes” or “No”) on each question, the way to interpret the results is to take scores from each category and to put them into the Scoring Table (see Table 3) to see the final result. To use the Scoring Table, the scores for the different categories should be placed directly on the table. Then, the final assessment is made by taking the score that is the lowest (furthest left) of the three different measures, and reading the label at the top row of the table. Since the first category is broken up into two parts, there is also an explanation of how

to take the two scores and come up with an overall score for the first category. The overall score of the three categories is determined by the lowest of the three categories since the quality of the evaluation is determined by its weakest link. Even if everything seems to be in place in two of the areas, the degradation of one of the systems makes the overall evaluation ineffective.

For a quick understanding of different results and possibilities, a second table (see Table 4), "Classification of Evaluation Systems based on Indicator Measurements" shows all of the possible combinations of scores and what each set of scores would imply about the evaluation and management system. This table just presents the different "types" of systems that might be found. Since there are three evaluation categories and they can be classified as "High" or "Low," there are eight potential "types," each corresponding to a certain kind of organizational model. Actual examples of specific systems with these scorings are shown towards the end of this article, following presentation of the indicator. A number of organizations whose scores fit these "types" are presented. At the end of the article is a through, detailed analysis of the UNDP's evaluation system that walks the reader through the organization, culling information that can be used to score its evaluation system and then walking the reader, step-by-step, through the scoring process. This is a case study of how to examine an evaluation system, break it down into its component parts, identify its weaknesses, and then score its effectiveness.

Overall, the power of this indicator is that it highlights how a weakness in one category of performance points to a very specific, systemic type of problem that

needs to be corrected. For example, an organization:

- Can have a professionally protected evaluation that has absolutely no impact because the results are not linked to any real changes or management oversight.
- Can have a strong management system that is open to the public, but can use evaluators to promote a hidden or political agenda if there is no professional protection of the evaluator; or
- Can follow "professionalism" in every way but destroy the public interest and public purpose because of no accountability to the public, such as businesses that have no effective public regulation.

**Table 3**  
**Scoring Table for the Evaluation Indicator and its Components**

	<i>SCORING</i>			
<i>Measures</i>	<i>Corrupted Evaluation Likely Used as a Propaganda and Fundraising Tool</i>	<i>Vulnerable Evaluation</i>	<i>Partly Effective Monitoring but Potentially Vulnerable Evaluation Subject to Influence</i>	<i>Professional, Accountable Evaluation Protecting public spending</i>
I. Fulfillment of the Functions of a Management Control System: Utilizing Proper Evaluation Roles, Expertise, and Data Collection (Weighting of scores of A. and B.)	At most 0 – 1 point in category A or 0 – 2 points in category B The evaluation is just for show (window dressing; a formality only) and has no impact.	At most 1.5 - 2 points in category A or 2.5– 4 points in category B The evaluation system is probably driven by an outside agenda or is co-opted internally.	At most 2.5 – 3.5 points in category A or 4.5 – 6.5 points in category B A strong system needing only some changes	All 4 points in A and All 7 points in B A model professional management system
A. Managerial Controls that Apply Evaluation Systems to Rewards and Punishments: (0 to 4 points)	0 to 1 point An evaluation system that is symbolic but that has no impact.	1.5 - 2 points An ineffective or vulnerable evaluation system	2.5 – 3.5 points A working evaluation system that could be more professionalized	All 4 points— A model professional management system
B. Quality of Seven Individual Oversight Systems that Evaluations Monitor (0 to 7 points)	0 – 2 points No measurements or systems. Evaluation is for show.	2.5 - 4 points Weak organization probably driven by an outside agenda or run for bureaucrats and their friends	4.5 – 6.5 points – Strong management system in most areas	All 7 points – A model professional management system
II. Transparency and Accountability of the Evaluation, Itself: Watchdog Role (0 to 6 points)	0 – 2 points The system is designed for bureaucrats to protect their interest against the public	2.5 – 4 points There is lip service or recognition of public protections that sometimes work but can be thwarted. The public may be fooled by what it is able to see.	4.5 – 5.5 points The system is open. Nevertheless, quality is vulnerable and some opinions may still be suppressed	All 6 points – The public receives everything and can be assured of quality
III. Protection of Professionalism and Political Insulation of Evaluations (0 to 8 points)	0 – 2 points No real protections at all	2.5 – 4 points Professionals will have to battle against pressures and take considerable risks	4.5 – 7.5 points Vulnerable to financial pressures or risk of violations of professional standards.	All 8 points – A model system of protection and professionalism
Total	Scores cannot be totaled. The overall measure is determined by the LOWEST column score among the three measures.			

**Table 4**  
**Classification of Evaluation Systems (8 Potential Types) Based on Indicator Measurements**

<i>Type of Evaluation System</i>	<i>I. Functioning Management Control System</i>	<i>II. Transparency and Accountability of the Evaluation (Government Watchdog and Public Protection)</i>	<i>III. Professionalism and Political Insulation of Evaluations</i>
<i>1) Model Evaluation System</i>	High	High	High
<i>2) Efficient BUT Totalitarian System (Government) or Strong Business Management with weak public oversight</i>	High	Low	High
<i>3) Politicized Evaluation to Promote a Hidden Agenda</i>	High	High	Low
<i>4) Corrupted Evaluation System Protecting Bureaucrats</i>	High	Low	Low
<i>5) Evaluation System that Exists only for Show as a Formality</i>	Low	High	High
<i>6) Democratic oversight exists but without professionalism (common to NGOs)</i>	Low	High	Low
<i>7) New or small organization without institutional procedures</i>	Low	Low	High
<i>8) Useless or Non-existent Evaluation System</i>	Low	Low	Low

*Measures/Sub-Factors.* Below are the 25 questions of the indicator in the categories described above. Most of the questions are clear cut “Yes” (1 point) or “No” (0 points or negative points for harms) in scoring, but in cases where there is a judgment call, it is possible to opt for a “Debatable” (0.5 points for benefits and 0 points for harm). (Note that the questions are numbered by their category, rather than from one to 25.)

*I. Fulfillment of the Functions of a Management Control System: Utilizing Proper Evaluation Roles, Expertise and Data Collection.* These questions go to the heart of what a management control system is designed to do to make an organization run effectively. The questions reveal whether or not the necessary feedback and monitoring systems and evaluation functions actually exist to make evaluations meaningful exercises for measuring organizational and individual performance, or whether they exist just for show. The first sub-category looks at whether there is a management feedback system using results monitoring to reward or punish performance and the second sub-category lists the required performance measurements that every organization needs. (11 questions with a total of 11 points divided into two sub-categories)

*I.A. Managerial Controls Exist that Apply Evaluation Systems to Rewards and Punishments: There is Professionalism in the Various Management Functions in the Cycle of Activities of the Organization so that the Evaluation is Meaningful.* These questions reveal whether evaluations are actually used as a form of managerial control that makes a difference in organizational performance and change, or whether they are just “due

diligence” paperwork for show. (4 questions with a total of 4 points)

Question I.A.1. *The Evaluation Acts as a Control Linked to Personnel Systems and Audits, for which Employees are Rewarded or Punished.* There are performance-based incentives in the system, itself, from the design phase to the review phase of every project/product. The evaluation feeds into those incentives with the results of the evaluation used not only to change organizational spending but to reward and punish employees on the basis of performance. The evaluations are not just perfunctory, providing required information or seeking to convince those in the organization to change or follow “best practices” and share in a “community of practice” but are managerial control tools with serious implications. The evaluation works to measure achievement on specific “functions” in an organization that have specific relevance to beneficiaries – to promote or protect a particular resource, for example – with each function designed to have measured results rather than just serving as a title or “area” of control. No points if the evaluation does not go directly to the funding/hiring authority but goes instead to the implementing agency to review itself, since this is immediate evidence that results are not used for punishment of poor performance.

Scoring: Yes = 1

Debatable = 0.5

No = 0

Question I.A.2. *The Evaluation Function is Appropriately Measured, Itself, for its Value Added: Evaluation Spending/Investment is Related to Calculated Organizational Income and Savings.* The evaluation serves a vital role

in organizational performance and is funded in ways that directly impact on organizational performance. Evaluation results feed into key decisions on spending and strategy. Specific evaluations and funding for them are determined through studies that link the benefits of information with the costs of that information. As a general rule, organizations should commit 5 to 10% of funding to research and development that includes the evaluation function, with auditing separate. No points if the evaluation is simply “required” by a donor and serves no purpose other than as a legal obligation to justify funding, without any rational relation to other organizational incentives for quality. In such cases, the system is short-circuited, with pressures to merge or delay or defund evaluations. “Final evaluations” to merely close out projects and offer “lessons learned” suggest that the evaluation system is just for show.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.A.3. *There Exists a Backstop and Protection of the Integrity of the Evaluation System, Itself, Assuring Direct Accountability to the Actual Owners/Managers.* The evaluation is conducted and overseen by an independent unit that reports to the highest management level of the organization and is independent of implementing managers but directly accountable to owners (stockholders for a business; taxpayers, a legislature with spending authority, and an independent and public representative judiciary for public spending) without any intervening bodies that can work to fortify opposition to objective evaluations. No points if

internal “Ombudsman” or “hotlines” and fraud units are not publicly representative and directly accountable, or if they report to management rather than the public, or if they are called independent but have shared interests with implementing management (such as membership in the same union). In such cases, the evaluation unit can potentially thwart evaluation integrity by creating additional hurdles to the public (slowing down the process, wearing down complainants, and serving to target complainants). The backstop process must be backed by strong laws and offer sufficient interest/ penalties and support as incentives to complainants to act freely to protect the system. The costs to the complainants must be minimal (not requiring process in another country, or before an unenforceable body, for example).

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.A.4. *Specific and Distinct Measurement Systems Exist that can be Evaluated and that Create Individual Accountability.* There are distinct sub-systems into which evaluations can feed. The organization recognizes that evaluations need to include the professional skills that correspond to these 7 key systems/ roles for the evaluation to be meaningful. These 7 key systems/ roles include (and are tested individually in the next section)—strategic planning and assuring adherence to the organizational mission, fulfillment of requirements of project design; auditing of input spending; managerial accounting of efficiency of outputs; managerial accounting of efficacy of outputs/relevance; research and development functions for



benchmarks/standards/improved projects; and assurance of the organizational sustainability and/or growth and replication of projects. The evaluation system needs to separately establish schedules and funding for measurement of these different roles and assure that evaluators that are hired have the skills to fit these different roles. No points if evaluators are given all of these multiple roles at once or if there is no rational plan for evaluators to allocate time and resources in ways that does justice to each of these areas in a strategic way.

Scoring: Yes = 1

Debatable = 0.5

No = 0

*I. B. Quality of Seven Individual Oversight Systems that Evaluations Monitor.* This section measures the quality of management systems in the seven key areas of operations (seven system areas) of an organization. It determines whether evaluations will have the necessary information to come to useful conclusions about performance. The questions closely follow information presented in Table II in the first part of the article. If such measures are weak, evaluators are likely being commissioned to perform a political or advertising role under the form of an evaluation, in order to justify more funding and cover for officials. (7 questions with a total of 7 points)

Question I.B.1. *System 1. Protection of Organizational Mission and Strategic Planning to Fill It Efficiently and Competitively.* There is a regular strategic planning function. Every intervention is rigorously screened to ensure that it meets

two tests. First, it promotes the specific mission of the organization in a measurable way (solving a specific type of problem with a particular tool through a comparative advantage of the organization: to protect and promote development of a resource in a sustainable way on behalf of public beneficiaries and contributors for a government function and in a way that changes government action or provides a non-governmental function for a non-governmental organization). Second, is a comparatively better (more strategic) investment than other interventions in doing so. The measure of this function is whether there is a list of comparative benefits of projects and a screening device to reject projects that are weak or inappropriate after carefully measuring their fit. One type of failed system only offers categories and slogans of fit (“poverty” or “development” or “equity”) and actually substitutes other concerns such as opportunistic ones (donor funding) and attempts to do everything. Another failed type defines interventions on the basis of a specific tool that the organization uses (“capacity building”) rather than on a comparative advantage of the organization in solving the root causes of a specific problem. The measure of whether a government function and mission are accountable is the quality of legislation, public oversight, and public legal challenge to assure compliance and implementation. As a test of quality of this system, there should be regular strategic planning exercises (usually on an annual basis) and organizational management reviews, with expert managers and high level management should be trained in strategic planning. A simple spending plan or action plan or list of symptoms to be targeted, or a plan developed with partners or on the basis of slogans does

not qualify as a “strategic” plan. The evaluation role is to review the data and measures used for screening and for setting the organization mission. The evaluation should yield review of the mission and lists of projects and new approaches that would further the mission of the organization in a measurable way. No points if there is no measurable data to demonstrate the organization’s comparative advantage versus other organizations or if there is no one with comparable professional skills to a strategic planner, who is conducting the evaluation of the work of strategic planning.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.B.2. *System 2. Professional Project Design and of Management Systems for Implementation.* The project design applies professional expertise and detailed measures that link activities to root causes of the problem to be solved, not to symptoms, and that show the specific behavior changes/ changes in incentives and target actors. There are measured impacts and benchmarks that offer specific measurable goals per unit of spending on particular links in the causal chain of the problem. There is a clear distinction between “outputs” (processed inputs that result in a product like a book or policy or law) and “outcomes” that are measurable benefits to target groups and society. There is a clear logical framework that shows how specific inputs lead to specific outputs with measurable changes in root causes and measurable changes that follow to solve the problem. There is an explanation of how the specific intervention is the most strategic use of resources to change specific behaviors and

incentives that will change specific outcomes and how the structure of management relates to the most efficient way to achieve this. For development projects, this framework is placed directly in the context of a sustainable development framework. A typical benchmark for this function in an organization is 5% of the project spending. The evaluation role is to review the data and logic of the project intervention with specialized skill in the area of the problem and root causes and in interventions that provide measurable changes in behaviors and measurable impacts on the problem. No points if the expert and the system do not offer specific quantifiable measures of the stages of problem solution rather than the symptoms.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.B.3. *System 3. Cost Accounting and Cost Control Systems for Procurement and for Personnel.* There is a professional cost control system that fulfills a watchdog function through independent comptroller/auditor spot-check audits (different from due-diligence book-keeping audits) relying on a system of clear measurements of potential suppliers/ product costs/ quality levels of each purchase, as well as competitive bidding procedures. The cost control system works to examine ways of saving funds to achieve the same objectives, rather than simply tracks spending to purchases (book-keeping). The rule of thumb is generally that an external, independent auditing and accounting of inputs should be 2% of spending and that other cost control oversight measures pay their way by finding cost savings less than

the cost of the oversight to identify over-spending. The evaluation role assures that spot checks have been independent and effective and that there are specific cost measurements and systems for continually generating savings and reducing waste. No points if the evaluation does not have the expertise to spot check and re-analyze each item of spending and to show where savings could have occurred and the amount of savings, as a check on whether the overall audit functioning is performing effectively.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.B.4. *System 4. Managerial Accounting for Efficiency of Outputs for Measurable Technical Levels (Quality per unit cost of output benefit)*. A managerial accounting system exists that combines technical measurements of quantifiable results with units of spending to yield efficiency measures (technical value for unit of spending). There are also measures and screening devices for whether the appropriate financial mechanisms were used to fit behavioral incentives and to assure the appropriate functioning of the market, of government, and of NGOs without improper subsidization of either. Managers understand their role in managerial accounting of outputs and they appropriately use technical experts in the design phase, making strategic choices of interventions on the basis of output per unit cost. The evaluation role is to use technical expertise to oversee the measures and choices of interventions on the basis of quality, to see that the appropriate technical and quality measurements are included and that they are linked with spending measures, and that the benchmarks in use reflect

standards of the industry. No points if the evaluation does not have the expertise to analyze each intervention on the basis of its spending and to indicate specific parts of the budget that were inefficient or improper based on different kinds of financial alternatives and on professional benchmarks of impact on root causes of problems per unit of spending. No points if the organization does not understand the difference between cost of inputs (regular accounting) and cost per output and output benefit (managerial accounting) or claims that it “can’t” be measured.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.B.5. *System 5. Clear Benefits Measures and Externalities Measures as Linked to the Organization’s Mission and to Solutions that Address Root Causes (for Public Policy)*. A system of screening interventions for harms (social, cultural, environmental) and of measuring benefits already exists. The system records clear and quantifiable measures that are linked with an overall outcome related to the organization’s mission, including such categories as changes in behaviors, in relative power, and shifts of assets to promote sustainable human development for government projects. The system also offers relative comparisons to other interventions. For public projects, the outcome measures are fit directly into schemes for sustainable human development balances (population, consumption, and resources) and for total assets per capita for each ethnic-cultural group. The measures are also relevant and useful to citizens so that there can be clear oversight of the value of the organization. The evaluation role assures that all the

systems are in place, verifies the quality and accuracy of the measures, and offers comparative benchmarks. No points if unintended outcomes are hidden or not measured (particularly long-term impacts/ externalities), if measures that reflect a hidden or biased agenda are substituted (such as “income” for real development goals in development projects), or if measures are not linked to solutions to the root causes of the problem (rather than to amelioration of symptoms).

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.B.6. *System 6. Planning for Organizational Sustainability/ Maintenance/ Growth and Avoiding Co-Dependency on the Problem.* There is an organizational business plan and a strategy for promoting the recipient or institutionalizing the benefit in ways that are not dependent on continuing symptoms that need to be treated but that offer solutions to eliminate the problem. Systems produce such benefits and replicate/expand the models and advocate for them in ways that promote replication. For public projects, the business plan for organizational sustainability avoids dependency of the recipient on outside funds of a donor. The professional skills that are part of the planning for organizational sustainability include some version of financial systems analysis (public finance for government activity), financial analysis of linkages between the measured benefit and the beneficiaries in ways that links contributions with benefits, (social) marketing systems to promote replication, and systems of reinvestment and innovation (research and development in competitive

technological advances that are commercially viable) that include calculations of the benefit of these investments in organizational sustainability. The evaluation role is largely that of financial expertise, to review the funding mechanisms to see if the measures have been correctly calculated and if the beneficiaries recognize these measures and purchase the product (for business) or contribute the service (government or civil society). No points if the measure of sustainability is based on “support from foreign donors” and if the beneficiaries of the project outcome are not directly linked to the purchasers in some way and the focus is only on whether the benefits of the project continue into the future.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question I.B.7. *System 7. Scientific Research and Development System into which Evaluations Feed.* A scientific measurement system of results exists that uses at minimum the methodologies of control groups and hypothesis testing and that employs individuals who are bound to professional codes and certification in the fields of outcomes and externalities that require measurement. This subsystem feeds into the design and organizational mission systems. As a benchmark, spending on research and development should be some 5% of project spending to assure adequacy and quality of data measures, interpretation, and use of the data for development of new approaches. The evaluation role does not seek to do the measures for the first time but takes existing measures and offers scientific hypotheses, interpretations, and improvements,

building on the professionalism and systems of research that are institutionalized in the organization and that build on each other sequentially. No points if there is no evidence of following the scientific method.

Scoring: Yes = 1

Debatable = 0.5

No = 0

*II. Quality of Transparency and Accountability of the Evaluation, Itself (Watchdog role).* This scoring category can be used both for government processes as well as for private organizations (businesses and civil society organizations that conduct evaluations) to measure of the integrity of processes to meet an organization's needs by providing information directly to the decisionmakers. If the organization being measured is a government agency, the "beneficiaries" are the citizens paying for the government activities and the citizen beneficiaries, and the questions are asked once. If the organization being measured is one that the public charters, as it does for private businesses and civil society organizations, the questions can be asked twice: once where the decision-makers are the shareholders/owners/members of the organization (to yield a stakeholder accountability score) and once to consider transparency and accountability to citizens, as for government organizations, to test whether government oversight/regulation of business and civil society organizations is accountable and transparent. (6 questions with a total of 6 points)

Question II.1. *Beneficiary (including Taxpayers for Public Projects) Has Control over the Process and There is Direct Evaluation Reporting to the*

*Budget Authority.* Evaluations respond directly to the publicly accountable body that prepares the budget and not to the implementing stakeholder agencies. The evaluation is also insulated from any marketing unit or function of the organization that is seeking or can seek additional funding or work. For public spending, the commissioning body of the evaluation is one directly accountable to taxpayers and or beneficiaries of the spending through either direct and competitive election/removal and/or legal challenge and this commissioning body is separate from the implementing agent of the program being evaluated. For example, the evaluator reports to a legislature with separate powers and interests from the executing agency. For a business or NGO, the reporting is to top management with freedom to report directly to shareholders and also to government. (This question is similar to I.A.2. but goes further, not only protecting the evaluation quality but also assuring reporting.)

Scoring: Yes = 1

Debatable = 0.5

No = 0

Question II.2. *No Intervening Agent or Screening Authority, Between the Individual Evaluators and the Budget Authority, Acts to Prevent Full Reporting and Openness on Public Matters to the Public, (Public or Public Bodies for Public Spending).* Each evaluator has a direct relationship with the budget authority (for public spending, the public/public body) without any intervening private company management or other screening or vetting body to the evaluation. The evaluator's work directly incorporates all public laws and protections of public whistleblowers and government actors, without placing

the evaluation in the private sector (under a private contractor who has the power to hire or fire subcontractors at will in ways that eliminate their public protections). The evaluator is not placed in any position where the standards of work override public interest and the public right to information that is not commercial, proprietary information or that of national security in BOTH the funding and recipient countries. Contracting provisions on public evaluations bind the evaluator to public law of the highest standard of EITHER the taxpayer or recipient country or funding organization (rather than the lowest standard of one of these). These laws cannot be overridden by any other kind of subcontracting or intermediate contracting with private contracting firms running the evaluation. Nor can any evaluation contract seek to create any agreements for secrecy or withholding information from the public that is different from the laws and protections of business secrets, national security secrets or protection of individual informants, and that interferes with the public right to information.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question II.3. *Drafts Reach the Public/Budget Authority for Discussion.* The public/ budget authority is assured of seeing the raw reports of evaluators and has an opportunity to present information directly to, receive full information from, and to interact directly with evaluators in ways that assure no information is suppressed. Any pressures or changes or replacement of a report or of an evaluator will be reported and will be visible. Draft reports and any changes in draft reports

as well as the reasons for them are placed on public record.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question II.4. *Information Access is Enforceable and Dissemination Regulations Protect the Beneficiary and Evaluator.* The organizational regulations and public law are sufficient to assure protection of the rights of the beneficiary to information (e.g., Freedom of Information laws and real enforcement through easy-to-use court procedures before publicly representative judges, for public information) and the rights of the evaluator to disseminate information freely to the beneficiaries. In case of public information, the evaluator is able to make that information available without restrictions of copyright or other barriers, controls of ownership or threats of costs to safety or career, and with sufficient Whistleblower protections.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question II.5. *Quality of Evaluation Reporting is Sufficient for Beneficiaries to Use to Make Judgments on Comparative Value and Spending.* Evaluation reporting is sufficiently standardized so that the evaluations contain appropriate and adequate data, processed with benchmarks and information sufficient for a cost benefit or other comparative return/benefit measure on relevant categories of evaluation. Decision-makers, including the public, can read the reports and use them to make comparative funding decisions that are rational and defensible. Reports that

do not refer to common and comparative standards but that describe project spending as in its own category earn no points.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question II.6. *Real and Measurable Information Delivery and Access.* Project design and information on the choices available at the design stage, interim evaluation reports, and final evaluations all reach the beneficiaries in a regular and routine manner such that the information is readily available in ways that the beneficiaries (not the implementing stakeholders) can influence the funding and professional quality of the interventions at each stage, as well as hold individuals accountable.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

*III. Ability to Safeguard Professionalism and Politically Insulate Evaluations: A Test of Quality and Objectivity.* The questions here cover the most common and key areas where pressures are currently exerted to compromise evaluations. There may be other areas in the steps of the process from hiring to conduct of the evaluation to presentation of results to final payment that can compromise the objectivity of the evaluation and put pressures on evaluators. In fact, one weak link in this chain of protecting professionalism can circumvent it. It may be important to continue to revise this list to add tests of other practices. (8 questions with a total of 8 points).

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.1. *Consultant Hiring Focuses on Professional Skills.* The hiring process of consultants is conducted with appropriate measures of the skills that evaluators and decision-makers should use in determining specific value per unit spending. The hiring system understands the professional standards and separates the standard requirements from other considerations. There is a whistleblower law, transparency of selection, and a working legal procedure in which candidates can challenge the hiring procedure before a neutral public body, to assure against politicization and retaliation/blacklisting in hiring. No points if the hiring process is conducted by non-experts who focus on factors such as likeability or experience with the organization, ability to please project partners or attract funding, or factors like gender or nationality or cost or ease of administration of the evaluator (e.g., hiring of a company that can replace or backstop/pressure evaluators).

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.2. *Professional Standards of Work are Protected.* A professional body—either a union or a professional certification organization—exists to enforce an ethics code AGAINST organizations that hire/commission evaluators and on behalf of professionals acting as evaluators. These codes are required in an evaluator's contract and are fully protected by law, with both the professional organization and the evaluator having legal standing to enforce

judgments in a court. These codes act not to protect salaries but to protect the public interest and rights of those impacted by the projects and outside of the implementing agencies (i.e., the code is not simply an “ethics code” that commissioning organizations use to shield themselves and improve their reputations, but is a code designed for and enforceable by the public).

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.3. *Independent Judgment is Protected through Evaluator Control of the Process.* The independent judgment of each consultant is protected by contract for the exercise of the evaluation in a manner that protects the independence and objectivity of the evaluator in choice of methodology, questions and forms, and in the presentation of findings. The contract protects the consultant’s ability to be free of reporting back to the implementing agents about questions or methods, to be free of mistranslations or termination of questioning and answers, and any other micro-managing (to “protect relationships” and respond to “sensitivities”) that could interfere with the quality of information and could compromise the evaluation.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.4. *Protected Independent Judgment in Reporting of Minority and Deviant Views.* The evaluation contract and practice protect the independent judgment of each profession important to the evaluation and each evaluator. Each professional is guaranteed the space and

opportunity to record a full dissenting and independent judgment, without fear that these will be overridden by a “team” or a “consensus” view or with any other limitations.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.5. *Protected Continuation of Evaluations without being Subject to Termination without Professional Review.* Evaluators are protected contractually from being replaced in the course of the evaluation or from early termination of the evaluation at any time before their signed final report is complete and made available directly to the beneficiaries, for any reason other than failure to fulfill their professional responsibilities, subject to review by independent professionals.

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.6. *Insulation of Evaluator Payment from Pressures.* Pay is contractually calculated and protected on the basis of professional time, not on client “satisfaction” with the product. Evaluators need not report on interim findings as a justification for continuation of work. The contract is protected in an independent court that is easily accessible without costs to the evaluator (e.g., in the evaluator’s home country).

Scoring: Yes = 1  
Debatable = 0.5  
No = 0

Question III.7. *Pre-Judicial Enforcement of Objectivity and Standards is*



*Independent of the Implementing Agents Being Evaluated.* The control agency of the evaluation, including any Ombudsman, Hotline, or Fraud protection or Comptrollers, are completely separate from the agency under review without conflicts of interest (same union, same boss, shared interest in project funding or continuation or favorable review) and their work is overseen directly by the beneficiaries (including taxpayers on public projects).

Scoring: Yes = 1  
 Debatable = 0.5  
 No = 0

Question III.8. *Whistleblower Protections Protect Standards of Performance and Rehiring of Evaluators.* Performance measures used by the hiring authority to judge the quality of work of the evaluators are made openly available for the evaluator to see. Such measures meet professional standards and are reviewable by a professional body. Whistleblower protections include a system for evaluators/short term workers that allows for challenging performance measures and for review by a transparent body.

Scoring: Yes = 1  
 Debatable = 0.5  
 No = 0

How Some Organizations Do: After understanding how the test works, it is easy to apply to an organization in just a few minutes and with close agreement among anyone using it. Below are several examples showing how different organizations and projects score, from best to worst.

Both generalized and actual examples are presented below; actual examples in the cases of public organizations like the

United Nations system and the European Commission, and generalizations about types of businesses and civil society organizations, without naming the specific organizations.

Note that even though not every question applies to every kind of project, the scoring is still designed to yield a scoring spread that leads to categorization and comparison and that also shows how some organizations can do better or worse depending on their attention to specific features of a management system or of the evaluations process that are highlighted in the scoring system.

*Effective (but Partly Vulnerable) Evaluation Systems.* The most effective evaluation systems are those in private businesses in competitive markets, where the incentives are those of cost savings, competition, and efficiency. At the same time, where public/government regulations are weak, as in most countries in the world today, there are still shortcomings in full public oversight on activities of business that are of public concern. Of the two “idealized” kinds of systems: the “model” evaluation system and the government totalitarian or business/statist system, neither really exists in this ideal form. In fact, totalitarian systems are not particularly efficient for long because the lack of oversight degrades professionalism. Businesses in competitive markets do roughly approximate these ranges.

Table 5

<i>Type of Evaluation System</i>	<i>I. Functioning Management Control System</i>	<i>II. Transparency and Accountability (Government Watchdog)</i>	<i>III. Professionalism and Political Insulation of Evaluations</i>

		<i>og and Public Protection)</i>	
1) <i>Model Evaluation System</i>	High	High	High
2) <i>Efficient Totalitarian or War-Economy System (Government) or Strong Businesses Management with weak public oversight</i>	High	Low	High

Major Business in a Competitive Environment – Most corporate evaluation systems score as effective though partly vulnerable in cases of weak management or less competitive industries where departments or managers seek to hide their failures or avoid major changes. Competitive private companies will generally score all 4 points, reflecting professionalism, in the first category and 5 of 7 points in the second category of management control systems, reflecting a partial effectiveness for public protection (but possibly professionalism for their own concerns of profits) losing only 2 points in measures of whether they actually measure the external harms that business can cause to public interests and measure whether they are promoting products that people don't need (like more drugs rather than preventative health). In other words, if they are judged only by their accountability to their stockholders, they will likely earn all 7 points, but if the public is considered and government regulatory systems are weak,

the score is only five points. Transparency and accountability to management generally scores about 5 points, as “effective,” given that there is a tendency for employees to hide weaknesses from management, and professionalism and protection of evaluations scores about 6 – 8 points, as “effective,” (with loss of points for lack of whistleblower protections for outside consultants who usually work irregularly for a business) in meeting needs of management and shareholders for professionalism, though the score is slightly less if the scoring is done from the perspective of public regulation and protection of business. Note that where upper level managers really want to politicize an evaluation in order to promote a particular agenda that has already been determined, they will remove the professionalism protections in the third category and the evaluation will be of the type often found in domestic policy oversight, where bureaucrats have a strong control system but seek to promote a politicized agenda (below).

*Systems that Can be Partly Effective or Vulnerable.* Domestic government programs can be effectively evaluated in strong democratic systems but are vulnerable in weak democracies to being captured by political interests controlling the government agenda or by government bureaucrats.

Table 6

<i>Type of Evaluation System</i>	<i>I. Functioning Management Control System</i>	<i>II. Transparency and Accountability (Government Watchdog and Public Protection)</i>	<i>III. Professionalism and Political Insulation of Evaluations</i>
----------------------------------	---	--	---

3) <i>Politicized Evaluation to Promote a Hidden Agenda</i>	High	High	Low
4) <i>Corrupted Evaluation System Protecting Bureaucrats</i>	High	Low	Low

*Failures: Systems that are Corrupted Internally or Externally and that Exist Only for Show:* These organizations claim to be open to public inspection and to follow public processes, but the test exposes them quickly. Here is where international organizations that do development are found. They are almost entirely unaccountable to anyone and evaluators are mostly used to help insulate the system from public scrutiny and to keep money flowing. Typically, they score low on all categories, though some may be partly accountable (to donor agendas) or have some professionalism (where there are specific, measurable outputs to be delivered, such as health benefits) but lack the combination of all of these to be effective.

U.S. Domestic Policy Departments: Domestic policy evaluation systems in countries where there is political competition and rule of law score as partly effective in responding to citizens who are both the direct taxpayers and beneficiaries, though they are also vulnerable to special interests and to bureaucrats degrading oversight and protecting their interests. In the category of management control systems, U.S. departments score about 2.5 points overall, for holding officials accountable through spending authorities and generally about 4.5 points out of 7 for including most kinds of screening, auditing systems, and measurements, though they are not always strong on assuring compliance with law and mission/strategic planning, or benchmarking managerial accountability of outputs and outcomes. Transparency and accountability varies in the effective range of about 4-6 points. It is the protection of evaluator professionalism that can be compromised by weak systems and points range from about 3.5 to 5.5.

Table 7

Type of Evaluation System	I. Functioning Management Control System	II. Transparency and Accountability (Government Watchdog and Public Protection)	III. Professionalism and Political Insulation of Evaluations
5) Useless or Non-existent Evaluation System (Most Likely)	Low	Low	Low
6) Evaluation System that Exists only for Show	Low	High	High
7) Democratic oversight but no professionalism (common to NGOs)	Low	High	Low
8) New or	Low	Low	High

<i>small organiza tion without institutio nal procedur es</i>			
---	--	--	--

United Nations Development Programme (UNDP) - (See charts below, analyzing this organization in detail): In all three categories, UNDP scores in the range of corrupted evaluations, with 0 points for a management control system and only 0.5 for specific measures that a control system would include; 0.5 points for transparency and accountability, and 2 points out of 8 for professionalism and protection against politicization. This is a system with no measures and standards, no oversight, and no protection of professionalism. It is characteristic of a business that sells its position as an intermediary (in this case, access to government officials to promote foreign laws, agendas, or relationships) and takes a percentage of the transactions, rather than as a public body seeking to fulfill a specific public purpose. In this type of system, bureaucrats and other financially interested stakeholders (recipient governments and their contractors) put pressures on evaluators to offer praise and writing evaluations as advertisements for more work, and to hide any negative information from the public. The indicator suggests that the system is corrupted and the evaluation process and evaluators who remain in this system are also likely to be corrupted by it.

Major Foundations (generalized): Major international foundations earn scores that suggest they are easily corrupted in their evaluation systems. The underlying problem appears to be that their main function is distributing money for a specific agenda defined by a donor, over

which beneficiaries and the public have little oversight, making it easy for foundation administrators to serve their own interests with little monitoring and control to ensure compliance with other objectives. With the main function as giving away money and not focusing on results or accountability unless the donor is particularly knowledgeable about management and outcomes, the control system scores 1 point for having financial systems, making it corruptible, and 3-4 points for the specific systems that it may have to monitor funds, making it vulnerable. With internal reporting only, the accountability score is about 2 points, though more if the founder is still in charge. As a foundation, public transparency, however, will be low. Depending on the foundation's goals, evaluation professionalism could actually be effective in looking at specific goals like the delivery of grants for particular purposes, but this information is narrow and may or may not lead to accountability.

U.S. Agency for International Development (USAID) and, similarly, other country donor agencies – USAID is slightly more accountable than the UN system because it is partly accountable to taxpayers and funds are monitored by a legislature but the system also scores in the range of those where evaluations are corrupted. The agency is susceptible to capture by its bureaucrats or by special interests seeking special benefits (implementing agents or political interests who override public law and agenda) because the activities occur overseas where they are hard for the public to monitor, and beneficiaries (and those harmed) by projects do not have rights in the U.S. In the management control system category, overall, USAID scores about 2 points, registering as vulnerable,

but overall control by the public is practically non-existent and earns only 2 points, in the “corrupted” range because of a weak management system that is typical of development aid administration and conflicting and hidden agendas. Professionalism is partly protected by laws, earning 2.5 points, but is vulnerable because of the lack of most protections and the use of contracting procedures that allow government bureaucrats to manipulate evaluations. Transparency is partly protected by laws but only weakly, earning 2 points, and reflecting that foreign policy activities, even aid, are among the least transparent of those of governments. The results here seem to suggest that aid policies are not really intended to produce results that protect beneficiaries, and maybe not even the taxpayer public in the donor countries, but are really discretionary funds to buy influence and prestige with governments and/or citizens of recipient countries.

Other United Nations Agencies such as UNICEF, ILO, WHO: Other UN agencies score about the same as UNDP, with evaluation systems scoring as “corrupted” for the same reasons but with slightly better scores on control systems – 1.5 points of out 7, though still easily co-opted – and up to 3 points for protection of professionalism of evaluators, in the vulnerable category, due to the fact that some of their projects are relief efforts and delivery of aid where there are specific health or other benefit measures that are less subject to politicization and difficulties of measurement as technical advisory services. At the same time, some of these agencies micro-manage evaluations through their evaluation offices in ways that seek to control results directly, and their score on transparency and accountability can drop to 0, denoting

un-accountable and scoring as “corrupted.”

International and local development NGOs (generalized): Although the systems in use in NGOs vary from professional and results-driven cost-conscious management working with very little money and seeking to maximize results, to donor driven systems run by administrators with little or no real management expertise and with an incentive to keep projects running regardless of impact, the lack of any required standards forces many of the most well-known organizations down to the lowest denominator where they are vulnerable to being corrupted. There is a bit of a paradox that small organizations tightly controlled by their founders, have strong adherence to mission and efficiency but lack management skills, while larger organizations have better systems, particularly for measuring goods delivery (like the UN agencies, above) but no incentives to be accountable to their beneficiaries or to taxpayers who donate to them. They may only be accountable to donor controls, ironically making them slightly more transparent than some government agencies might be, but those controls often have the same problems as the donor systems, themselves, with an added layer of self-interested stakeholders with an interest in politicized evaluations and with fewer protections of professionalism than in government. In all three categories, a typical international development NGO scores in the range of corrupted evaluations, with 1 points for a management control system (noting some possible controls) but 0 points for specific measures that a control system would include; 1-2 points for transparency and accountability, and 0.5 points for professionalism and protection against politicization.

European Commission (EC): It might be surprising that the EC's evaluation system of its development projects scores among the most "corrupted," given that many countries in Europe have very strong domestic accountability and evaluation. The problem is that the EC is currently not really accountable to citizens or even to its member governments, let alone recipients in developing countries. This indicator confirms many of the criticisms that have appeared in the press about the EC, including reports by the European Court of Auditors that the EC budget violated basic accounting standards every year over the past decade, with not only no attempt at reform, but firing of whistleblowers like the EC's chief accountant, Marta Andreasen (The Times, 2004), and harassment of newspaper reporters like *Stern's* Hans-Martin Tillack, who have investigated its corruption (2009). Countries deliver money to the bureaucracy but there is no oversight and there is no political agreement on standards or even on mission, leading to little clarity of mission other than that of trade promotion (the original goal of the European Union) replacing even development goals. In all three categories, the EC scores in the range of corrupted evaluations, with 0 points for a management control system and 0 points for specific measures that a control system would include; 0 points for transparency and accountability (since neither the European Parliament or even the European Audit have any real power to oversee the Commission, with citizens powerless to directly challenge the bureaucracy), and 1 point out of 8 for professionalism and protection against politicization. The EC uses the same contracting procedures that USAID uses, through contractors, that enables it to pressure evaluations, and its supposedly

"independent" Anti-Fraud unit is not accountable to citizens but is itself a bureaucracy with officials in the same union as those whom they are supposed to hold accountable.

## Post-Script: Solutions

The irony of exposing the flaws in evaluations systems today is that the "experts" who are in the position to make changes have little incentive to change, while those who are best protected by change are the least informed and organized about where or how to begin to push for reforms. An indicator can facilitate change, but like other improved tools, they must be in the hands of those willing and able to use them.

Professional evaluators in the organizations that score the worst on the new indicator in this article will likely not even recognize their failures.

They will have a hard time understanding or admitting that they have been trying to perform functions that require several sets of professional skills that are outside their individual or department repertoires. They are likely to say that cost-benefit and efficiency analyses and benchmarks do not apply. They may say that this business-like approach that introduces a variety of professional expertise takes the artistry and "humanistic" or "human" judgment out of their work, though in fact it does the opposite by supplementing it with additional expert analysis from multiple fields across the larger spectrum of human achievement.

They are likely to respond defensively to suggestions for more public oversight of their work, and to claim that accountability is a form of "policing." They are likely to say that oversight implies

“mistrust” and that their good faith is being questioned when evaluations promote accountability. They may say that holding a government official accountable for results is unfair because there are “too many factors.”

Overall, such responses from many “professional evaluators” will demonstrate exactly why many of the people in place in current systems are part of the problem and not the solution. Indeed, the only real solution is mobilization of the public.

This author has suggested the formation of Donor Monitor NGOs that act as public advocates (Lempert, 2008a) and has designed a full set of other governance reforms in media, organizational oversight, private attorneys general and other citizen powers that would promote professionalism and accountability at the level of constitutional changes (Lempert, 1994) as well as educational and cultural reforms (Lempert, et al., 1995). But who will fund and promote them? It is as if we are at a stalemate in attempts at real human progress.

In short, foxes have entered the henhouse in design and evaluation of international development projects as well as many other governmental systems, and there is a need to devise better oversight systems and architecture of henhouses. The foxes and failed architects, who have allowed the foxes to enter, see approaches like the one in this article as unfairly stigmatizing them and not praising them for simply doing their best in their roles. But the foxes currently have neither the incentive nor the consciousness of their role in causing the harm, to achieve solutions. The only way that change can really occur is if farmers and consumers who have an interest in the oversight and partly the hens act collectively to protect

their interests. This article offers one tool to at least facilitate that effort.

Table 8  
Scoring of UNDP on the 25 Component Questions of the Indicator

Preliminary Information for Assessment	
<p>Legal Structure of the UNDP Evaluation System, According to Law (from the UNDP Website, referencing U.N. General Assembly Resolution 59/250 of 2004)</p>	<p><i>1) The Evaluation System</i>  <i>Evaluation Mandate describing the systems and goals:</i></p> <ul style="list-style-type: none"> <li>▪ “the systematic evaluation of United Nations system operational activities by assessing their impact on poverty eradication, economic growth and sustainable development of programme countries. It further mandates the United Nations system to promote national ownership and capacity development and to make system-wide progress in collaboration in evaluation.”</li> </ul> <p>☞ Note that: The UNDP mission is not referenced in this document, nor are other international laws.</p> <ul style="list-style-type: none"> <li>▪ “strengthen partnerships with governments and key stakeholders.”</li> </ul> <p>☞ Note that: Key stakeholders are not defined. There is no clarity on whether bureaucrats and implementers using UN funds are the key stakeholders or whether the beneficiaries and taxpayers are. There is no explanation of how the competing goals of “growth” and “sustainable development” are to be achieved and no reference to other legal obligations and treaties that would put sustainability and cultural rights as the priorities, with growth only one of the choices of sustainable human development that is consistent with international laws and UNDP mission. There is no definition of “poverty” or reference to international standards and laws.</p> <p><i>Implementation and utilization of evaluation findings:</i></p> <ul style="list-style-type: none"> <li>▪ 32. Units responsible for functions or programmes that have been evaluated should systematically implement evaluation recommendations when these have been agreed to by management.</li> </ul> <p>☞ Note that: No enforcement mechanism is specified other than the Executive Board.</p> <ul style="list-style-type: none"> <li>▪ 41. Country offices, regional bureaux, and practice and policy bureaux will be required to: (a) prepare an evaluation plan, based on guidelines established by the Evaluation Office, which will include mandatory and other evaluations; and (b) cost this plan, and allocate the requisite funds from appropriate project and programme budgets.</li> </ul> <p>☞ Note that: There is no measurement of cost-benefit that links evaluations to value-added or that sets any quality standards for the evaluation.</p> <p><i>2) The Watchdog Role</i>  <i>Oversight:</i></p> <ul style="list-style-type: none"> <li>▪ <i>The Executive Board of UNDP/UNFPA</i> is the custodian of the evaluation policy. The Executive Board: (b) ensures the independence of the evaluation function;</li> <li>▪ <i>The Administrator of UNDP</i> is accountable for UNDP results, and: appoints the Director of the Evaluation Office, in consultation with the Executive Board,</li> <li>▪ <i>Evaluation units of the associated funds and programmes</i> are the custodians of the evaluation function in their organizations</li> </ul> <p><i>Accountability and Transparency:</i></p> <ul style="list-style-type: none"> <li>▪ 33. All UNDP evaluation reports will be made public.</li> <li>▪ “Evaluation units” will, respectively, for their fund or programme: (i) ensure the maintenance of a publicly accessible repository of evaluations;</li> </ul> <p>☞ Note that: No enforcement mechanism is specified</p> <p><i>3) Objectivity and Professionalism</i>  <i>Independence.</i>                      The evaluation function should be structurally independent from the operational management and decision-making functions in the organization so that it is free from undue influence, more objective, and has full authority to submit reports directly to appropriate levels of decision-making. Management must not impose restrictions on the scope, content, comments and recommendations of evaluation reports.</p> <p>☞ Note that: No standards or enforcement mechanism is specified other than the Executive Board. The role and funding of the Executive Board are not specified.</p> <p><i>Protection of Professional standards and Evaluators:</i>                      38. EO will support a network of evaluation practitioners who will be responsible for nurturing a culture of evaluation in the organization by integrating evaluation into training programmes; promoting the use of quality standards; disseminating methodologies; providing technical guidance; and sharing evaluative knowledge.</p> <p>☞ Note that: No mention is made of protections of evaluators or reference to their professional standards and no enforcement mechanism is specified</p>



	<p>to ensure breadth and accountability of the “network.”  <i>Source:</i> Executive Board of the United Nations Development Programme and of the United Nations Population Fund, <i>The Evaluation Policy of UNDP</i>, DP/2005/28, May 2006.  <a href="http://www.undp.org/ea/documents/Evaluation-Policy.pdf">http://www.undp.org/ea/documents/Evaluation-Policy.pdf</a>                  Commentary provided by author.</p>
<p>Mandate of the Evaluations Office (from the UNDP Evaluation Office Website)</p>	<p><i>1) The Evaluation System</i>  <i>Evaluations Office Mandate describing the system and goals:</i>                  “to support the Administrator in his substantive accountability function and contributing to organizational learning through provision of systematic and independent assessment of results, effectiveness and impact of the substantive activities of the programme, including the special purpose funds under the Administrator’s responsibility.                  - Refocus EO’s policy, strategy and priorities to help in the repositioning of UNDP and enhance its relevance as a knowledge-driven global advisor and catalyst in the context of rapidly changing global development conditions.                  - Support improving UNDP’s development effectiveness through introduction of methodological innovations and application of Results-Based Management.                  - Support implementation of the organization’s change initiatives through mainstreaming the culture of measurement, monitoring and evaluation.                  ☞ Note that: There is no definition of the mission of the office in terms of sustainable human development as defined in treaties and measures (including the U.N. Conference on Environment and Development, The Rio Declaration, 1992), by international treaty enforcement, or by promoting development or protecting specific resources in keeping with sustainable development. There is only a reference to the Millennium Development Goals that offers a set of general indicators that are not balanced by an organizational mission  <i>2) The Watchdog Role</i>  <i>Oversight:</i>                  ☞ Note that: there is no mention of the public role or public over funds or assuring conformity with international law or with the interests of those who fund the UN system or the beneficiaries.  <i>Implementation and utilization of evaluation findings:</i>                  ☞ Note that: there is no statement of the office having any role in accountability of funds or other incentives to hold UNDP employees accountable for performance.  <i>Accountability and Transparency:</i>                  “To maintain and develop partnerships with the UN system organizations (as chair of IAWG), multilateral banks, OECD-DAC on evaluation related work.”                  ☞ Note that: there is no statement on how the office assures public transparency or accountability. The only relations mentioned are not those with the public but with other international organizations that actually have different missions and goals, such as the international banks.  <i>3) Objectivity and Professionalism</i>  <i>Independence.</i>                  ☞ Note that: there is no mention of how independence of evaluations is protected. There is no definition or explanation of how the differences between local control and international standards, between donor goals and UNDP standards are resolved and whether there is any oversight by the Executive Board or whether the Executive Board has actually delegated its oversight role to the bureaucrats, to monitor themselves. (See below)  <i>Protection of Professional standards and Evaluators:</i>                  ☞ Note that: there is no statement on professionalism and protections.  <i>Source:</i> UNDP Evaluation Office website:  <a href="http://www.undp.org/ea/ea-mandate.htm">http://www.undp.org/ea/ea-mandate.htm</a>                  Commentary provided by author.</p>
<p>Critique of UNDP Evaluations as Presented on the UNDP Website (Direct quotes from Evaluation Executive Summary)</p>	<p><i>1) The Evaluation System</i>  <i>Evaluation Mandate describing the systems and goals:</i></p> <ul style="list-style-type: none"> <li>▪ “[I]n practice, mobilization of resources and delivery are more powerful drivers of individual performance among programme staff than achievement of outcome level results.</li> <li>▪ “[A] single, simple and consistent message on resource mobilization [in place of other missions and goals] ... was communicated to both internal and external audiences; [with] development of systems to track, measure and report managers’ success in mobilizing resources; and a clear perceived link between successful resource mobilization and advancement within the organization.</li> <li>▪ “The country programme outcomes against which UNDP is expected to deliver are poorly defined; the logic linking outputs delivered by UNDP</li> </ul>

	<p>with achievement of the outcomes is often not explained; and therefore, this reporting system fails to report on UNDP performance relative to what it is accountable for [other than the goal that the new UNDP administrator has substituted for substantive goals; those of raising money and transferring it to governments and implementers].</p> <ul style="list-style-type: none"> <li>▪ “[Of 166 UNDP countries, only] Twenty-five offices have appointed monitoring and evaluation (M&amp;E) specialists [15%] and 10 have an M&amp;E unit. [only 6%]</li> </ul> <p><i>Implementation and utilization of evaluation findings:</i></p> <ul style="list-style-type: none"> <li>▪ “Results-based management assumes that managers have the flexibility to allocate programme resources to maximize results. That assumption does not generally hold true for UNDP.</li> <li>▪ “A notable omission is the lack of oversight systems that focus on tracking whether programmes use results to adjust resources such as people, money and partnerships in order to improve future results.</li> <li>▪ “The evaluation concludes that UNDP is largely managing for outputs rather than outcomes and that the linkages between outputs and intended outcomes are not clearly articulated. The introduction of corporate systems and tools, which have had some efficiency benefits, have not however, strengthened the culture of results in the organization or improved programmatic focus at the country level. The current approach of defining and reporting against centrally defined outcomes tends to undermine UNDP’s responsiveness and alignment to nationally defined outcomes and priorities.</li> <li>▪ “With the advent of results-based management, the focus of results shifted to outcomes, but apart from the results-oriented annual report, no specific tools were developed to help monitor results. The ATLAS system is steadily gaining in importance, but its primary focus is financial management and project monitoring.</li> <li>▪ “Adoption of outcome-level evaluations in the countries visited has been slow. In general, they have been underbudgeted and poorly timed to influence the content of country programme documents.</li> </ul> <p><i>2) The Watchdog Role</i></p> <p><i>Oversight:</i></p> <ul style="list-style-type: none"> <li>▪ “Four relationships stand out as the most critical [and that are currently missing]: at the Executive Board, to ensure the programme is held to account for development results; between the Administrator or Associate Administrator and the Bureau Directors; between Directors of Regional Bureaux and Resident Representatives or Country Directors; and between Resident Representatives or Country Directors and staff within country offices.</li> <li>▪ “The oversight and management roles of the regional bureaux do not focus on results.</li> <li>▪ “Decentralization has been accompanied by delegation of authority over the country programme. Under current procedures, country programmes are not scrutinized for development potential by regional management, an abdication of responsibility.</li> </ul> <p><i>Accountability and Transparency:</i></p> <p>“There is no evidence that the Resident Representative/Country Director is held accountable for managing for outcomes. In practice, the corporate goals and service lines set by headquarters have proved too numerous, with very permissive definitions. This has led to country offices manipulating their programmes to fit into corporate service lines, diverted attention away from country needs and made reporting to the Executive Board more about process than substance.”</p> <p>☞ Note that: The UNDP Evaluation Office website also presents the “management’s response,” written as “UNDP believes.” While the response disagrees with the evaluation, the way it is presented actually supports the evaluation findings and re-enforces the report’s criticisms. The response offers no actual evidence of measurements or systems and makes it clear that the management can disregard the evaluation findings and recommendations without any sanction. It is also clear that the Evaluation Office controls the site and what is presented, without any public comments or rebuttal either in the preliminary or post-evaluation phase. The office is accountable only to itself.</p> <p><i>3) Objectivity and Professionalism</i></p> <p><i>Independence.</i></p> <ul style="list-style-type: none"> <li>▪ “Independent evaluation is an important element of results-based management to validate performance. Yet country evaluations conducted by the UNDP Evaluation Office have emphasized learning over accountability and have not measured performance against stated intentions. Validation occurs only through infrequent audits.</li> <li>▪ “Roles and responsibilities are generally clear, but country programme outcomes and indicators are not subject to quality assurance and there is little independent validation. Individual targets in the Results and Competency Assessment are self-selected and are often applied retrospectively and poorly linked to incentives. Despite the intended shift to managing for outcomes, individual staff remain tied to a project orientation and accountability for outputs.</li> </ul> <p><i>Protection of Professional standards and Evaluators:</i></p>
--	--

	<p>☞ Note that: there appears to be evidence on the UNDP Evaluation site's web page that even this critical evaluation was also controlled and possibly pressured by the Evaluations Office. The report itself lists the "Evaluation Office Task Manager" on the front page and includes a Forward from the head of the Evaluation office; something that would not appear if the evaluation were independent.</p> <p>Source: UNDP Evaluation Office, <i>Evaluation of Results Based Management at UNDP: Achieving Results</i>, 2007. Derek Poate and 9 staff.  <a href="http://www.undp.org/eo/documents/thematic/RBM/RBM_Evaluation.pdf">http://www.undp.org/eo/documents/thematic/RBM/RBM_Evaluation.pdf</a>          Commentary provided by author.</p>
<p>Putting UNDP Monitoring Systems to Simple Tests of Transparency and Professional Protection</p>	<p><i>1) The Evaluation System</i>  <i>Implementation and utilization of evaluation findings:</i>          Though the Executive Board of UNDP, composed of representatives from 36 countries, is deemed to be the oversight body of UNDP spending, it appears that the group exists in name only and does not meet regularly or have any real staff to perform oversight functions. The UNDP's website does not list any addresses for its members and its Secretariat has only one staff member listed without an email contact. It appears that it delegates its oversight work back to the UNDP bureaucracy, thus short-circuiting any real oversight.</p> <ul style="list-style-type: none"> <li>▪ The "Contact" link on the UNDP Executive Board website goes immediately to the UNDP bureaucracy, meaning that it is impossible to complain about the UNDP Administrator or the investigations office under the Administrator. The link leads to the following:             <ul style="list-style-type: none"> <li>- "If you wish to file a complaint against a UNDP staff member or about a UNDP project, programme or operation, you can do so using this online form and/or the Investigation Hotline." <a href="http://www.undp.org/comments/form.shtml">http://www.undp.org/comments/form.shtml</a></li> <li>▪ According to the UNDP information public information office, there is no regular oversight function of the Executive Board.</li> <li>- "I had never heard of individuals wanting to send a complaint to all members of the Board. I am not aware of the EB having its own investigative unit. All I know is that frauds and violations to UNDP's regulations, rules and procedures are taken care of by the Office of Audits and Investigations [under the UNDP Administrator and part of the UNDP bureaucracy]. Mariana Gonzalez, UNDP Internet Feedback Message, August 22, 2008</li> </ul> </li> </ul> <p>The documents that are the basis for projects and evaluations stem from a universal Project Document format that is now designed to reduce information and to speed the process of project acceptance. The form, itself, is not public. The amount of space on the cover page to describe the indicators for the project is limited to 37 characters for outcome indicator(s) and outcome targets. The focus of the document is on activities, partners, deliveries, costs, and timetables, without any focus on the problem, root causes, chain of solutions, comparative value of interventions, screening to assure compliance with the UNDP mission, screening to assure compliance with international law, or explanations of strategy. It is essentially a sales agreement for the transfer of funds rather than for the achievement of particular objectives.</p> <p>Source:          United Nations Development Programme, <i>Project Document: Required Formats and Contents</i>, part of Results Management Guide, created by Dien Le, December 20, 2006, internal document. (Some applications are available on the web under this title but the form, itself, is not public)</p> <p><i>2) The Watchdog Role</i>  <i>Oversight:</i>          The UNDP Executive Board does not have its own monitoring office equivalent to a watchdog office of funds/ investigative audit or managerial accounting office. (See additional explanations of the Executive Board, above.)</p> <p><i>Accountability and Transparency:</i>          Several provisions in evaluation contracts could be considered as attempts by UNDP to violate accountability and transparency given that there is no independent public review of what constitutes confidential and what is public information that UNDP bureaucrats and others are seeking to withhold in violation of the public interest. Among them are the following:</p> <ul style="list-style-type: none"> <li>▪ "all maps, drawings, photographs, mosaics, plans, reports, estimates, recommendations, documents and all other data compiled by or received by the Individual contractor under the Agreement shall be the property of UNDP, shall be made available for use or inspection by UNDP at reasonable times and in reasonable places, shall be treated as confidential and shall be delivered only to UNDP authorized officials on completion of work under the Agreement"</li> <li>▪ "Information and data that are considered proprietary by either UNDP or the Individual contractor ... and that are designated as confidential ("Information"), shall be held in confidence"</li> </ul> <p>Source:          UNDP, "SSA General Terms and Conditions"  <a href="http://europeandcis.undp.org/files/hrforms/SSA_General_Terms_and_Conditions.doc">http://europeandcis.undp.org/files/hrforms/SSA_General_Terms_and_Conditions.doc</a></p> <p><i>3) Objectivity and Professionalism</i>  <i>Independence.</i>          "The Individual contractor [UNDP evaluator] shall not take any action ... that may adversely affect the interests of UNDP, and the Individual contractor</p>

	<p>shall perform its obligations under the Agreement with the fullest regard to the interests of UNDP.                  ☞ Note that: there is no definition or professional review of the “interests of UNDP” and this can easily be interpreted to mean “the (corrupt) interests of bureaucrats working for UNDP” in ways that compromise evaluations rather than protection of international law and the public, as it should.  <i>Protection of Professional standards and Evaluators:</i>                  Disagreements between evaluators and UNDP are to be resolved by UNCITRAL Arbitration Rules, with costs that are proportionately less to the UNDP than to an evaluator.  <i>Source:</i>                  UNDP, “SSA General Terms and Conditions”  <a href="http://europeandcis.undp.org/files/hrforms/SSA_General_Terms_and_Conditions.doc">http://europeandcis.undp.org/files/hrforms/SSA_General_Terms_and_Conditions.doc</a></p>	
Overall Assessment and Interpretation	<p>While some evaluation regulations exist, the main potential for short-circuiting of the system appears to be the lack of any enforcement or accountability to the beneficiaries (taxpayers who pay or the cultures and individuals in the developing world who are to benefit). The broad scope of goals without a defined focus allows the organization to become politicized and for outcomes to be manipulated. Without accountability, an agenda like fundraising and transfer of funds, that meets the interests of bureaucrats, can easily replace accountability to public laws or to other missions. Decentralization of measures and controls also makes the organization susceptible to manipulation by governments and financial stakeholders at two levels: donors seeking outcomes for their own interest and recipients also seeking to manipulate goals and outcomes. There is a danger that additional regulations and directives that are presented as “reforms” could continue to disguise the overall weaknesses of systems that allow evaluations to function independently (under the control of self-interested management) and without any real control function for impact and for accountability.</p>	
<p><b>Analysis: I. Functioning Management Control System: Utilizing Proper Evaluation Roles, Expertise and Data Collection</b></p>		
<p><i>I.A. Managerial Controls that Apply Evaluation Systems to Rewards and Punishments: Professionalism in the Various Management Functions in the Cycle of Activities of the Organization so that the Evaluation is Meaningful:</i></p>		
Question	Indicator	Scoring
I. A. 1.	The evaluation acts as a control linked to personnel systems and audits, for which employees are rewarded or punished?	No, it does not. The UNDP Evaluation office in New York offers guidance on evaluations but does not exercise any control. Nor is any control exercised through the regional offices or the UNDP Executive Board. The UNDP missions, themselves, control their own work. Moreover, the standard that now seems to run the system is one of using the evaluation to appeal for donor funds rather than achieving any outcomes that are linked to the UNDP mission. Evaluations are not used for personnel reviews and performance is based on fundraising to sell the UNDP to donors for their objectives. <i>0 points.</i>
I. A. 2.	The evaluation function is appropriately measured, itself, for its value added and investments in it are related to calculated organizational income and savings	No, it is not. The evaluation “plans” are simply pro-forma budgeting exercises to fulfill evaluation requirements and they have no measure of adding value. <i>0 points</i>
I. A. 3.	There is a backstop and protection of the integrity of the evaluation system, itself	Theoretically, there is the UNDP Executive Board that represents a rotation of member countries, and each country can also challenge the UNDP Administrator. These mechanisms, however, are indirect, without any direct oversight by member countries or their citizens and no judicial body that can review challenges. Oversight is on paper, only, and the reality is that the UNDP runs itself. <i>0 points.</i>
I. A. 4.	There are specific and distinct measurement systems that can be evaluated and that create individual accountability?	The new “Results Based Management” / ATLAS system is failing to focus on outcomes or performance measures and the Millennium Development Goals that have been set as targets are vague, contradictory, and inconsistent with the organization’s legal mission that has been poorly defined (many say by design). Measurements fail. <i>0 points</i>
Total and Score		0 points. Non-existent controls. The only “system” that exists in the UNDP is an accounting system for money received and spent, which is a sales system for a business, not a system for government programs. Only fundraising is being rewarded. There is no real measurement system to speak of. The “communities of practice” and “knowledge” sharing approach is not a management system but a network.

<i>I. B. Quality of Seven Individual Oversight Systems that Evaluations Monitor:</i>		
Question	Indicator	Scoring
I. B. 1.	System 1. There is protection of organizational mission and strategic planning to fill it efficiently and competitively?	No, there is not. International law and the organization's mission appears to have been replaced by an agenda of fundraising. The mission of "sustainable human development" as defined by the Rio Convention has been lost and replaced with competing slogans that can justify any intervention, anywhere, without real measures. The project design formats are much like sales agreements, with no rigorous screening to assure consistency with international treaty protections for cultures and peoples, or benefits to public stakeholders. <i>0 points.</i>
I. B. 2.	System 2. There are professional design of project and of management systems for implementation?	No, there are not. Most project design is done by project officers with no real research or research skills, or with too little time to really understand needs and how an intervention solves a problem. The project document required formats of the UNDP do not follow standard practice of requiring root cause analysis and problem trees or any other justification for the strategy of intervention other than that partners agree to accept funds for treatment of symptoms through approaches like "capacity building." There are no benchmarks of outcomes. Logframes lack correct links of outputs to outcomes, often misclassifying inputs (trainings or manuals) as outcomes. There is no attempt to show how inputs change behaviors or incentives to lead to long-term changes that are part of the organization's mission (sustainable development) that will improve the lives of beneficiaries. Though UNDP asserts that this is not the case, they are unable to produce real data to justify that they meet professional standards. <i>0 points.</i>
I. B. 3.	System 3. Cost accounting and cost control systems for procurement and for personnel, with spot checks of spending, exist?	No, they do not. Financial systems are book-keeping systems that simply follow transfers of money with due-diligence audits rather than managerial accounting audits that ask what real costs are, compare quality of different purchases, and justify spending as the highest quality for the lowest cost. In many countries, the UNDP works to subsidize government salaries (through per diems or payments to attend workshops that are part of their jobs) or to subsidize organizations that could raise funds from other sources or through the market. These expenses escape real oversight. <i>0 points.</i>
I. B. 4.	System 4. There is a managerial accounting for efficiency of outputs for measurable technical levels (quality per unit cost of output benefit)?	No, there is not. The one thing that UNDP does measure is outputs and for some kinds of what the UNDP calls "outputs" but that are really processed inputs and not systems or behavior changes that are really outputs – conferences, "training," "handbooks," draft laws, service delivery – there may be some measures of how much production costs. But since they do not include quality measures per unit cost, they are not really the kinds of managerial accounting inputs that allow for useful evaluation. <i>0 points.</i>
I. B. 5.	System 5. There are clear benefits and externalities measures linked to the organization's mission and solutions that address root causes (for public policy)?	No, there are not, and this is a clear failure of the UNDP system whose very mission is to apply international treaties to assure that cultures, environment, and equal opportunities are protected in each project and the overall balance of projects, rather than to claim that supporting one intervention at a time in each area (promoting productivity, for example, even at the expense of culture or without labor and environmental protections) is part of the development mission. <i>0 points.</i>
I. B. 6.	System 6. There is planning for organizational sustainability/ maintenance/ growth and avoiding co-dependency on the problem	No, there is not. UNDP does not currently screen its inputs to determine which functions government, NGOs or the private sector can perform on their own and how they should be measured and funded, or how systems advocate for and plan their own funding for these needs. The incentive is actually for recipients to continue to seek technical assistance and foreign resources so that projects keep running, since it is more costly for UNDP project officers to design new projects with new partners than it is to continue funding of previous projects. No system in UNDP works to challenge this co-dependency. <i>0 points.</i>
I. B. 7.	System 7. There is a scientific research and development system into which evaluations feed?	No, there is not. Although UNDP does seek to publish "best practices" and promote them among "communities of practice" networks, the reporting of individual cases lacks any kind of scientific design that would test hypotheses, and lead to meaningful measures. There is no system of scientific measures and at

		best the UNDP relies on measures of outsiders like the World Bank, with different goals to promote. Comparisons are rarely, if ever, done with control groups for comparisons. Networks that share ideas that lack scientific basis are reinforcing isolation and protecting mindsets, not doing research and development. <i>0 points.</i>
Total:		0.5 points. Non-existent measures and standards. The characteristic of the UNDP system is one that has no measures and justifies itself on political grounds and “smile sheets” (whether implementing partners are “happy” that they receive benefits). Each project is defined as unique in itself such that no real standards apply or can be used for accountability. Evaluations transfer a burden to evaluators to “find” justification for projects such that bureaucrats can advocate for additional projects.
Overall Score:		A. 0 points out of 4. B. 0.5 points out of 7. This is a system with no measures and standards and no oversight and is characteristic of a business that sells its position as an intermediary (in this case, access to government officials to promote foreign laws, agendas, or relationships) and takes a percentage of the transactions, rather than as a public body seeking to fulfill a specific public purpose.
<b>Analysis II. Transparency and Accountability of the Evaluation Itself (Government Watchdog and Public Protection)</b>		
<b>Question</b>	<b>Indicator</b>	<b>Scoring</b>
II. 1.	There is real beneficiary (including taxpayers for public projects) control over the process with evaluation reporting to the budget authority?	No, there is not. The UNDP Executive Board does not work directly with evaluators but delegates its authority to the bureaucracy under evaluation, thus short-circuiting the process. At no stage can the public hold the UNDP Executive Board or UNDP directly accountable through any mechanism. UNDP evaluations are run by the project officers whose work is under review and only nominally by the contracting offices who sign the checks with no accountability to the public. Governments that fund the UN system can only cut funds or buy projects but their citizens do not directly oversee the spending. The project officers write the Terms of Reference and manage the process, with review only by the Deputy Resident Representatives. Though there is a UNDP Evaluations office in New York, it exercises no control over evaluations other than to post guidelines for the process. Though the governments in host countries can control the local UNDP offices and though donor countries paying for specific projects can now set evaluation agendas (in direct violation of UN system regulations), the citizen beneficiaries and the citizens of countries funding the UNDP have no direct role in the evaluation process and no rights in it, and their legislative bodies or other public bodies do not commission the evaluations. <i>0 points.</i>
II. 2.	There are no intervening agent or screening authority and no attempts to evade full reporting and openness on public matters to the public, between the individual evaluators and the budget authority (public or public bodies for public spending)?	One positive quality of the UNDP evaluation system is that it directly contracts with evaluators rather than through private contracting firms in the way that USAID, the European Commission, and others do, in a scheme that allows them to exert pressures to falsify and distort evaluations by using the contractors to pressure professionals they hire. This means that evaluators control the product and relationship to the public. At the same time, the UNDP contract seeks to prevent publication and dissemination of materials through a vague standard of what constitutes “private” or “secret” information. Thus, the overall impact is debatable. <i>0.5 points.</i>
II. 3.	Evaluation drafts reach the public/budget authority for discussion?	No, they do not. “Stakeholders” are defined by UNDP as the recipient agencies of funds who have a vested interest in arguing for receipt of more funds, and these stakeholders are invited to see reports and to try to pressure them for changes. At the same time, citizens are excluded from the process and have no way to enforce rights or to even know about projects or evaluations. <i>0 points.</i>
II. 4.	There is enforceable information access and dissemination protecting beneficiary and evaluator?	No there is not. There is no judicial body or oversight body that citizens or beneficiaries can use to hold UNDP accountable to provide information. They must go to their own governments to challenge bureaucrats who they may be unable to hold accountable (often the claim for why UNDP projects are initiated). Requests for documents are ignored. Even consultants interviewing for positions as evaluators are usually unable to see the project documents or previous evaluations unless they are chosen as evaluators and they may not

		even be able to get the names or addresses of UNDP staff. Critical evaluations and annexes can be hidden even from evaluators, let alone the public. There are almost never open meetings between UNDP officials and the public. <i>0 points.</i>
II. 5.	The quality of evaluation reporting is sufficient for beneficiaries to use to make judgments on comparative value and spending?	No, there is not, because outcomes are rarely measured or compared. The typical evaluation document is used as a justification for future funding or a confirmation of delivery of funds with little question asked other than whether someone “received funds” and was therefore “better off.” There are no standards for spending comparisons or other types of decisionmaking between alternatives within projects or between projects other than “numbers of beneficiaries” and some impact on symptoms. <i>0 points.</i>
II. 6.	There is real, regular and measurable information delivery and access?	No there is not. Visit UNDP sites at the country level or regional or New York offices and you will find few if any evaluations, let alone full project documents or responses. <i>0 points.</i>
Total:		0.5 points. Evaluations are transparent in theory and the public is involved in theory, but the reality is that they only rarely see an evaluation or part of one that is critical of UNDP, and even when they do there is little that they can do to push UNDP to change any practices.
<b>Analysis III: Protection of Professionalism and Political Insulation of Evaluations</b>		
Question	Indicator	Scoring
III. 1.	Consultant hiring focuses on professional skills?	No, it does not focus on skills, though it might include them in a peripheral way. UNDP would say that their offices choose specialists on the basis of qualifications, but the selection is almost always done by project officers who have minimal training (often little beyond language skills) and focus on personal qualities, salary levels, or knowledge of the UNDP system rather than on actual problem-solving skills related to the area of work. <i>0 points.</i>
III. 2.	Professional standards of work are protected?	No, they are not. UNDP offers a form contract that does not reference professional standards and that does not allow it to be included. Contract offers can be rescinded if consultants seek to add protections and standards. <i>0 points</i>
III. 3.	Evaluator’s independent judgment is protected through evaluator control of the process?	Determining actual control of the evaluation process is a bit difficult. For most UNDP projects, the partners are government officials and access depends on UNDP, itself. UNDP selects project teams as well as translators. (By contrast, for EC evaluations, evaluators are to hire their own staff and make their own contacts. They have more “control” but no real access or understanding on the ground.) Evaluators select the methods and the questions but UNDP translators and project officers and staff can easily exert control over evaluations (and do) by reporting back behind the back of evaluators. The amount of money spent for evaluations is set by UNDP, often without any rational relation to the size of project or quality of information available and without any connection to real time needed for a full evaluation. It is easy to argue that evaluators have no control at all over the process, but since there is some debate UNDP gets the benefit of the doubt. <i>0.5 points.</i>
III. 4.	Evaluator’s independent judgment and reporting of minority and deviant views are protected?	Yes, it does. This question asks only whether there is pressure exerted on team members of evaluations to come to a consensus or whether individuals can submit separate reports and UNDP is unique among most international organizations in that it does not allow “team leaders” to force a single viewpoint. Contracts and work products are individual. <i>1 point</i>
III. 5.	Evaluations are protected against early termination without professional review?	No, they are not. UNDP offices view their “relationships” with their donors, government agencies and implementing partners as more important than tough evaluations and can stop evaluations at any time to protect their bureaucratic interests. Evaluators are not considered “staff members” and have none of the

		whistleblower protections of UNDP staff. There is no professional oversight over this. <i>0 points.</i>		
III. 6.	Evaluator payment is insulated from pressures?	The typical UNDP evaluation contract conditions payment on acceptance of a written product rather than on days worked, even though work is supervised and labor is expected. UNDP also typically asks to see evaluations before they are finalized and makes “requests” for changes that can be interpreted as pressures. At the same time, there are written regulations assuring evaluators of independence and these are incorporated into the contracts and can be upheld, though it is costly to do so. <i>0.5 points given that the insulation from pressures are debatable.</i>		
III. 7.	Pre-judicial enforcement of objectivity and standards is independent of the implementing agents being evaluated?	No, it is not. UNDP’s “Hotline” that is part of the Office of Audit and Investigations reports directly to the UNDP Administrator rather than to member states or the Executive Board, and the members of the office are also UNDP employees, belonging to the same union. They are not subject to public or professional oversight and have no real incentive to hold the UNDP accountable to its own laws when outside evaluators show they have been violated. <i>0 points</i>		
III. 8.	There are clear standards of performance and rehiring of evaluators is safeguarded through whistleblower protections?	No there is not. If there are records kept on evaluators, they are not public and there is no way to challenge them. The only measures of quality are whether project officers and country offices are “happy” since there is no professional review of evaluators by any body of equal professional credentials. Blacklisting is easy in the UNDP and there is no way to challenge it. <i>0 points.</i>		
Total:		<i>2 points.</i> There is weak protection of professionalism that allows strong, independent minded consultants to stick by their findings in their reports and being paid for doing so. The cost, however, is that such professionals are likely to be blacklisted at least at the country level and must deal with constant pressures and the time/costs required during the process of the evaluation.		
<i>Type of Evaluation System</i>		<i>I. Functioning Management Control System</i>	<i>II. Transparency ad Accountability (Government Watchdog and Public Protection)</i>	<i>III. Professionalism and Political Insulation of Evaluations</i>
Score Summary:		A. 0 points out of 4. B. 0.5 points out of 7.	0.5 points out of 6	2 points out of 8
		<i>Low to Non-existent</i> other than for raising and delivering funds	<i>Extremely weak</i> to the point of irrelevancy	<i>Low protection</i> of professionalism



**Table 9**  
**Final Scoring of UNDP Evaluation System**

	<i>SCORING</i>			
<i>Measures</i>	<i>Corrupted Evaluation Likely Used as a Propaganda and Fundraising Tool</i>	<i>Vulnerable Evaluation</i>	<i>Partly Effective Monitoring but Potentially Vulnerable Evaluation Subject to Influence</i>	<i>Professional, Accountable Evaluation Protecting public spending</i>
I. Functioning Management Control System: Utilizing Proper Evaluation Roles, Expertise, and Data Collection (Weighting of scores of A. and B.)	At most 0 – 1 point in category A or 0 – 2 points in category B The evaluation is just for show and has no impact.	At most 1.5 - 2 points in category A or 2.5– 4 points in category B The evaluation system is probably driven by an outside agenda or co-opted internally.	At most 2.5 – 3.5 points in category A or 4.5 – 6.5 points in category B A strong system needing only some changes	All 4 points in A and All 7 points in B A model professional management system
A. Managerial Controls that Apply Evaluation Systems to Rewards and Punishments: (0 to 4 points)	0 to 1 point An evaluation system that is symbolic but that has no impact.	1.5 - 2 points An ineffective or vulnerable evaluation system	2.5 – 3.5 points A working evaluation system that could be more professionalized	All 4 points— A model professional management system
B. Quality of Seven Individual Oversight Systems that Evaluations Monitor (0 to 7 points)	0 – 2 points No measurements or systems. Evaluation is for show.	2.5 - 4 points Weak organization probably driven by an outside agenda or run for the bureaucrats and friends	4.5 – 6.5 points – Strong management system in most areas	All 7 points – A model professional management system
II. Transparency and Accountability of the Evaluation, Itself: Watchdog Role (0 to 6 points)	0 – 2 points The system is designed for bureaucrats to protect their interest against the public	2.5 – 4 points There is lip service or recognition of public protections that sometimes work but can be thwarted and the public may be fooled by what it is able to see	4.5 – 5.5 points The system is open but quality and some opinions may still be suppressed	All 6 points – The public receives everything and can be assured of quality
III. Protection of Professionalism and Political Insulation of Evaluations (0 to 8 points)	0 – 2 points No real protections at all	2.5 – 4 points Professionals will have to battle against pressures and take considerable risks	4.5 – 7.5 points Either financial pressures or risk of violations of professional standards exist	All 8 points – a model system of protection and professionalism
Total: All columns are the farthest left. The overall score can be read in the top row.	<i>Overall:</i> Corrupted Evaluation: Useless or Non-existent Evaluation System, indicative of a system that has been co-opted by bureaucrats, implementing partners, and sources of funding, with pressures to exert influence over evaluators as well			

## References

- Boyle, R., & Lemaire, D. (1999). *Building effective evaluation capacity: Lessons from practice*. New Jersey: Transaction Publishers.
- Dakin, B. (2003). *Another quiet American*. Bangkok: Asia Books.
- Europeaid. (2007). *Project evaluations guidelines*. Retrieved December 11, 2009, from Europeaid/116548/C/SV
- Emmanuel, C., Merchant, M., & Otley, D. (1990). *Accounting for management control*. London: Chapman & Hall.
- Freedom House. (2008). *Methodology*. Retrieved December 11, 2009 from, [http://www.freedomhouse.org/template.cfm?page=352&ana\\_page=330&year=2006](http://www.freedomhouse.org/template.cfm?page=352&ana_page=330&year=2006).  
[http://www.freedomhouse.hu/images/fdh\\_galleries/NIT2008/NT-Methodology2.pdf](http://www.freedomhouse.hu/images/fdh_galleries/NIT2008/NT-Methodology2.pdf)
- Garrison, R. H., Noreen, E., & Brewer, P.C. (2005). *Managerial accounting*. New York: Irwin/McGraw Hill.
- House, E. R. (1978). Assumptions underlying evaluation models. *Educational Researcher*, 7(3), 4-12.
- Kaufmann, Kraay, & Mastruzzi. (2007). Governance matters VI: Aggregate and individual governance indicators, 1996-2006. *The World Bank Paper Series*. WPS4280 and WPS 4654. Retrieved December 11, 2009 from, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1148386](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1148386)  
<http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTC/OR/0,,menuPK:1740542~pagePK:64168427~piPK:64168435~theSitePK:1740530,00.html>
- Kusak, J. & Rist, R. (2001). Building a performance-based monitoring and evaluation system. *Evaluation Journal of Australasia*, 1(2), 14-23.
- Lempert, D. (1996). Building democracy in the former Soviet Union . . . with mirrors. *Legal Studies Forum*, 19(4), 373-379.
- Lempert, D. (1997). Holding the powers that be accountable to our ethics code to protect our integrity and the peoples we serve. *Human Rights*, 24(2), 22-24.
- Lempert, D. (2008a). Why we need an international development donor monitor. *Policy Innovations*. Retrieved December 11, 2009 from, [http://www.policyinnovations.org/ideas/policy\\_library/data/01464](http://www.policyinnovations.org/ideas/policy_library/data/01464)
- Lempert, D. (2008b). *Democracy building democratic project indicators for NGOs and international organizations*. Under Review.
- Lempert, D. (2008c). A dependency in development indicator for NGOs and international organizations. *Global Jurist*, 9(2), Article 6. Retrieved December 11, 2009 from <http://www.bepress.com/gj/vol9/iss2/art6/>
- Lempert, D. (2009a). A human rights democratic education indicator for NGOs and international organizations. *Interamerican Journal of Education for Democracy*, 3(1), awaiting placement on Website: <http://scholarworks.iu.edu/journals/index.php/ried/>.
- Lempert, D. (2009b). *A quick indicator of effectiveness of "capacity building projects" of NGOs and international organizations*. Under review.
- Lempert, D. (1994). Ukraine's new constitution: Continuity under the banner of change with a proposal for authoritarian to democratic transitions. *Demokratizatsiya*, 2(2), 268-290.
- Lempert, D., Briggs, X., et al. (1995). *Escape from the ivory tower: Student adventures in democratic experiential*

- education*. San Francisco, CA: Jossey-Bass.
- Lempert, D., & Nguyen, H. N. (2008). How to tell if development projects are doing sustainable development: An indicator for NGOs and international organizations. *International Journal of Sustainable Societies*, 1(1), 44-54.
- Poate, D., Riddell, R., Chapman, N. & Curran, T. (2000). *The evaluability of democracy and human rights projects: A logframe related assessment*. Stockholm: ITAD Ltd and Overseas Development Institute for the Swedish International Development Agency (SIDA).
- Tillack, H. M. (2009). *The corrupt republic: Creative fiddling of politics, bureaucracy and economy*. Berlin: Hoffman und Campe.
- The Times* (2004, December 6). Brussels and a culture of corruption.
- United Nations Conference on Environment and Development. (1992). *Rio declaration on environment and development*. Retrieved December 11, 2009 from, <http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm>
- United Nations Development Programme (UNDP). (2006). Project document: Required formats and contents. Restricted access internal document. Some applications are available on the web under this title but the form, itself, is not public.
- UNDP. (2008). *SSA general terms and conditions*. Retrieved December 11, 2009 from, [http://europeandcis.undp.org/files/hrforms/SSA\\_General\\_Terms\\_and\\_Conditions.doc](http://europeandcis.undp.org/files/hrforms/SSA_General_Terms_and_Conditions.doc)
- UNDP Evaluation Office. (n.d.). *EO mandate*. Retrieved December 11, 2009 from, <http://www.undp.org/eo/eo-mandate.htm>
- UNDP Evaluation Office & Poate et. al. (2007). *Evaluation of results based management at UNDP: Achieving results*. Retrieved December 11, 2009 from, [http://www.undp.org/eo/documents/thematic/RBM/RBM\\_Evaluation.pdf](http://www.undp.org/eo/documents/thematic/RBM/RBM_Evaluation.pdf)
- UNDP Evaluation Office. (2002a). *Guidelines for outcome evaluators*. Retrieved December 11, 2009 from, <http://www.undp.org/eo/documents/HandBook/OC-guidelines/Guidelines-for-OutcomeEvaluators-2002.pdf>
- UNDP Evaluation Office. (2002b). *UNDP handbook on monitoring and evaluation for results*. Retrieved December 11, 2009 from, <http://www.undp.org/eo/documents/HandBook/ME-HandBook.pdf>
- UNDP Executive Board. (2006). *The evaluation policy of UNDP*. DP/2005/28. Retrieved December 11, 2009 from, <http://www.undp.org/eo/documents/Evaluation-Policy.pdf>
- Wenar, L. (2006). Accountability in international development aid. *Ethics in International Affairs*, 20(1), 1-23.
- World Bank Independent Evaluation Group & MacKay, K. (2007). *How to build M&E systems to support better government*. Retrieved December 11, 2009 from, <http://www.worldbank.org/ieg/ecd/index.html>