



# DATA QUALITY ASSESSMENT OF PERFORMANCE MONITORING PLAN INDICATORS

Prepared for:  
USAID/Ecuador

Prepared by  
Harry Carr, Team Leader  
Clarence Zuvekas



Checchi and Company  
Consulting/Louis Berger  
Joint Venture



1899 L Street, NW  
Suite 800  
Washington, D.C. 20036

Contract No. AEP-I-824-00-00022-00

January, 2004

# TABLE OF CONTENTS

## EXECUTIVE SUMMARY (English and Spanish versions)

<b>I</b>	<b>Introduction</b>	<b>1</b>
	A. Purpose of the Data Quality Assessment	1
	B. Format of the Data Quality Assessment	2
	C. Performance Indicators Covered in the Data Quality Assessment	2
<b>II</b>	<b>Summary Performance Indicator Quality Assessments</b>	<b>5</b>
	A. Introduction	5
	B. SpO 11: Southern Border Integration	5
	C. SO 12: Democracy	11
	D. SO 13: Northern Border Development	20
	E. SO 14: Poverty Reduction	24
	F. SO 1: Biodiversity Conservation	27
<b>III</b>	<b>Weighting of Indicator Contribution to SO and IR Performance</b>	<b>31</b>
	A. Annual Report Guidance Regarding Weighting	31
	B. Weighting Criteria	31
	C. Recommended Indicator Weights	32
<b>IV</b>	<b>Revision of the Performance Monitoring Plan</b>	<b>37</b>
	A. Introduction	37
	B. SpO 11: Southern Border Integration	37
	C. SO 12: Democracy	38
	D. SO 13: Northern Border Development	39
	E. SO 14: Poverty Reduction	40
	F. SO 1: Biodiversity Conservation	42
<b>ANNEXES</b>		
	A. Southern Border Integration–SpO 11	
	B. Democracy – SO 12	
	C. Northern Border Development–SO 13	
	D. Poverty Reduction – SO 14	
	E. Biodiversity Conservation – SO 1	
	F. Persons Interviewed	

## EXECUTIVE SUMMARY

USAID/Ecuador contracted the Checchi and Co. Consulting/Louis Berger Joint Venture, under its Evaluation Indefinite Quantity Contract No. AEP-I-824-000220-00, to carry out a Data Quality Assessment (DQA) of selected indicators in its Performance Monitoring Plan (PMP). A two-person team visited Ecuador between November 28 and December 20, 2003. The DQA team interviewed Mission and implementing-partner (IP) staff and made three field trips to activity sites. The team assessed the quality of 26 indicators, including 23 that were to be reported in this year's Annual Report and three that the Mission is considering including in next year's Report. The indicators come from all five strategic objectives (SOs) and are at both the SO and intermediate-result (IR) levels.

Section II of the report contains summary findings and recommendations. The annexes provide a more detailed assessment of each indicator. The team's most significant general findings and recommendations regarding the quality of the indicators assessed are the following:

- The indicators chosen by the Mission for its Annual Report are, as a rule, at the IR level, while some are at the sub-IR level. Only one strategic objective included a performance indicator at the SO level. At the other extreme, five of the six indicators for one SO were at the sub-IR level. Although this observation is less about the quality of the indicators assessed and more about how well the "package" of indicators for each SO fully portrays the Mission's performance in achieving it, the team feels obliged to make the finding.
- The indicators generally lack directness in that a fairly large gap often exists between what the indicators measure and the results they are intended to measure.
- Many of the indicators are not defined clearly. Thus in cases where the Mission is receiving data on a particular indicator from more than one IP, the data sometimes lack uniformity. The DQA team has made recommendations for improving the clarity of the Mission's indicators; but in a few cases it seems best to seek alternative indicators.
- The utility of performance indicators as management tools depends in part on the frequency with which they are reported to USAID and analyzed by Mission staff. For a few indicators, data are being reported annually at best. Moreover, a number of indicators are not being reported with the degree of disaggregation necessary for the Mission to conduct a detailed, perceptive analysis of IP performance.

The Mission also asked the DQA team to assist it in weighting the indicators with respect to their validity as measures of progress toward achieving the SOs. In its current *FY 2004 Annual Report Guidance*, USAID is asking the Missions to stipulate in their Annual Reports "the relative importance" of their selected indicators "toward measuring the progress of the SO on a scale of (1) to (5), with a five being the highest score and a one being the lowest." The team recommends the following average weights for each of the five SOs:

- SO 11 – Southern Border Integration 2.4
- SO 12 – Democracy 1.2
- SO 13 – Northern Border Development 1.8
- SO 14 – Poverty Reduction 2.2
- SO 1 – Biodiversity Conservation 1.3

The Mission also asked the DQA team to make any important recommendations resulting from the assessment exercise that would help it revise its Performance Monitoring Plan (PMP). Since PMPs are formulated for each strategic objective, the DQA team’s general recommendations are made by SO.

- The Southern Border initiative of activities, designated as a Special Objective (SpO 11), is due to terminate in 2005. If the Mission intends to continue these activities as a Strategic Objective, it should formally revise its PMP. In this case, the first point to be addressed would be whether the program is to remain “transitional” in nature or take on a “developmental” quality.<sup>1</sup> Once that choice is made, the Mission should hold (a) an in-house strategic planning exercise to determine what it wants to accomplish in the Southern Border region, and (b) a three-day PMP workshop with its IPs to reformulate the results framework and identify performance indicators.
- The results framework for Democracy (SO 12) has an appealing horizontally and vertically integrated structure; but the SO itself is at too high a level for the Mission reasonably to affect with its limited resources. The SO is due to terminate in 2005. The DQA team recommends that the Mission internally question the formulation of the SO and wrestle with bringing it down to a more practical and precise level. Having arrived at a more precise SO, the Mission should formulate a new PMP around it. This task would require a workshop of about five days with the IPs.
- The Northern Border Development program (SO 13) was reformulated in February 2003 and extended to September 2006. Given its conversion from a SpO to an SO and its broader range of activities beginning in FY 2004, the Mission needs to develop a new PMP as a framework for reporting next year’s results. The DQA team supports the addition of new indicators being considered by the Mission, and it recommends participatory workshops with implementing partners and residents of the program’s three geographic areas.
- The Mission is addressing Poverty Reduction (SO 14) through two main activities: strengthening microfinance institutions and improving the macroeconomic and institutional environment for more equitable economic growth, principally by improving tax collections. While the DQA team regards the current strategic focus and results framework as appropriate, it recommends that some indicators be

---

<sup>1</sup> Although their definitions are unclear, the terms “transitional” and “developmental” are generally understood in USAID to distinguish two very different sorts of environments within which it works: “transitional” environments include locations that are in a situation of fundamental flux due to civil/military instability, perhaps moving in and out of violence, where long-term objectives are given up for shorter-term purposes; “developmental” situations are the more traditional environments in which USAID finds itself, where there is an expectation of stability and lack of civil/military violence, and where longer-term objectives can be sought.

clarified and/or amplified, and that the tax revenue/GDP indicator be monitored more closely.

- The Mission's Biodiversity Conservation activities (SO 1) went through a transition period in FY 2003. The new *FY04-FY08 Extension Amendment* (May 2003) reorients the strategy to focus on four themes and reformulates the three IRs to varying degrees. Some of the indicators being proposed need to be reconsidered in the light of the DQA team's observations. The team recommends that the Mission hold a two- or three-day workshop with its IPs to validate performance indicators and rewrite the PMP.

## RESUMEN EJECUTIVO

La Misión de la USAID en el Ecuador contrató a la empresa conjunta Checchi y Cía Consultores/Louis Berger, bajo su Contrato de Cantidad Indefinida para Evaluación No. AEP-I-824-000220-00, para llevar a cabo una Evaluación de Calidad de Datos (ECD) referente a indicadores escogidos de su Plan de Monitoreo de Desempeño (PMD). Un equipo de dos personas realizó una visita al Ecuador entre el 28 de noviembre y el 20 de diciembre de 2003. El equipo ECD entrevistó al personal de la Misión y sus socios ejecutores (SE) e hizo tres visitas de campo a sitios de trabajo. El equipo evaluó la calidad de 26 indicadores, incluyendo 23 designados para reportarse en el Informe Annual del presente año, y tres que la Misión considera incluir en el Informe del próximo año. Los indicadores provienen de todos los cinco objetivos estratégicos (OE) y abarcan tanto el nivel de OE como lo de resultados intermedios (RI).

La Sección II del informe comprende un resumen de resultados, conclusiones y recomendaciones. Los anexos ofrecen una evaluación más detallada de cada indicador. Los más sobresalientes resultados, conclusiones y recomendaciones generales sobre la calidad de los indicadores, son los siguientes:

- Los indicadores escogidos por la Misión para su Informe Anual generalmente se definen al nivel de RI, aunque en algunos casos se los encuentran al nivel de sub-RI. Un solo objetivo estratégico incluye un indicador de desempeño al nivel del OE. Al otro extremo, cinco de los seis indicadores para uno de los OE se encuentran al nivel de sub-IR. Aunque este comentario se aplica menos a la calidad de los indicadores evaluados y más a la eficacia del “paquete” de indicadores para cada OE en describir completamente el desempeño de la Misión en lograrlo, el equipo ECD cree que es necesario reportar esta conclusión.
- Generalmente los indicadores no pueden considerarse medidas directas, en el sentido de que existe a menudo una brecha bastante grande entre lo que mide los indicadores y los resultados que pretenden medir.
- Muchos indicadores no se definen claramente. Por lo tanto, cuando la Misión recibe de más de un SE, datos sobre un indicador específico, los datos a veces no son uniformes. El equipo ECD ha presentado recomendaciones para mejorar la claridad de los indicadores de la Misión; pero en algunos casos parece mejor buscar indicadores alternativos.
- La utilidad de los indicadores de desempeño como instrumentos de gestión depende en parte en la frecuencia con la cual estén reportados a la USAID y analizados por el personal de la Misión. Algunos indicadores se reportan anualmente en el mejor de los casos. Además, varios indicadores no se reportan con el grado de desagregación necesario para permitir que la Misión analice la gestión de sus SE en una manera detallada y perspicaz.

La Misión también pidió que el equipo ECD la asiste en ponderar los indicadores con respecto a su validez como medidas de avance hacia el logro de los OE. En la versión actual de su *FY 2004 Annual Report Guidance*, USAID/Washington pide a las Misiones

estipular en sus Informe Anuales “la importancia relativa” de los indicadores escogidos “en medir el avance del OE según una escala de (1) a (5), donde el ‘cinco’ representa la calificación más alta y el ‘uno’, la más baja.” El equipo ECD recomienda las siguientes ponderaciones promedios para cada una de los cinco OE:

- OEsp 11 – Integración de la Frontera Sur 2.4
- OE 12 – Democracia 1.2
- OE 13 – Desarrollo de la Frontera Norte 1.8
- OE 14 – Reducción de la Pobreza 2.2
- OE 1 -- Conservación de la Biodiversidad 1.3

Además, la Misión le pidió al equipo ECD sus recomendaciones prioritarias provenientes de la evaluación, las cuales ayudarían a la Misión en modificar su Plan de Monitoreo de Desempeño (PMD). Dado que los PMD se formulan para cada objetivo estratégico, el equipo ECD hace sus recomendaciones generales por OE.

- La iniciativa de actividades para la Frontera Sur, designado como Objetivo Especial (OEsp 11), está programada para terminar en 2005. Si la Misión pretendiera continuar estas actividades como Objetivo Estratégico, debería modificar formalmente su PMD. En este caso, el primer punto que debiera enfrentarse sería decidir entre seguir con un programa de “transición”, o convertirlo en uno de “desarrollo”.<sup>2</sup> Una vez tomada esta decisión, la Misión debe celebrar (a) un ejercicio interno de planificación estratégica, para determinar qué quiere conseguir en la región de la Frontera Sur, y (b) un taller con sus SE para reformular el marco de resultados e identificar indicadores de desempeño.
- El marco de resultados para Democracia (OE 12) tiene una atractiva estructura, integrada horizontal y verticalmente; pero el OE mismo se define a un nivel demasiado alto para que la Misión lo afectara con sus recursos limitados. El OE está programado para terminarse en 2005. El equipo ECD recomienda que la Misión cuestione internamente la formulación del OE y luchar para bajarlo a un nivel más práctico y preciso. Una vez definido un OE más preciso, la Misión debe formular un Nuevo PMD en base a éste. Esta tarea requeriría un taller de unos cinco días con los SE.
- El programa para la Frontera Norte (OE 13) se reformuló en febrero de 2003 y se prolongó hasta 2006. Dada su conversión de OEsp en OE, y la gama más amplia de sus actividades a partir del AF 2004, la Misión debe desarrollar un Nuevo PMD como marco para reportar los resultados del próximo año. El equipo ECD apoya

---

<sup>2</sup> Aunque sus definiciones no son claras, los términos “transición” y “desarrollo” generalmente se entienden dentro de la USAID como útiles para distinguir entre dos ambientes muy diferentes dentro de los cuales la Agencia trabaja: ambientes de “transición” comprenden áreas geográficas donde se encuentren situaciones de constante cambio debido a inestabilidad civil/military, talvez con episodios de violencia, y donde es necesario sacrificar objetivos de largo plazo en favor de propósitos de corto plazo; por otro lado, las situaciones de “desarrollo” son los ambientes más tradicionales dentro de los cuales la USAID trabaja, donde se espera un cierto grado de estabilidad y la falta de violencia civil/military, y donde es posible enfocar en objetivos de largo plazo.

la incorporación de los nuevos indicadores bajo consideración por la Misión, y recomienda talleres participativos con los socios ejecutores y residentes de las tres áreas geográficas donde funciona el programa.

- La Misión enfrenta a la Reducción de la Pobreza (OE 14) mediante dos actividades principales: el fortalecimiento de las instituciones de microfinanzas, y el mejoramiento del ambiente macroeconómico e institucional para facilitar un crecimiento económico más equitativo, principalmente a través del mejoramiento de las recaudaciones fiscales. Aunque el equipo ECD considera apropiados el enfoque estratégico y el marco de resultados, recomienda que algunos indicadores se clarifiquen y/o se amplifiquen, y que se realice un monitoreo y seguimiento más atento del indicador que mide la razón impuestos recaudados/PIB.
- Las actividades de la Misión en Conservación de la Biodiversidad (OE 1) experimentaron un período de transición en el AF 2003. El nuevo *FY04-FY08 Extension Amendment* (mayo de 2003) reorienta a la estrategia para enfocarla en cuatro temas, y reformula en varios grados los tres RI. Algunos indicadores propuestos deben reconsiderarse en vista de las observaciones del equipo ECD. El equipo recomienda que la Misión celebre un taller de dos o tres días con sus SE para validar los indicadores de desempeño y reformular el PMD.

# I INTRODUCTION

## A. Purpose of the Data Quality Assessment

The purpose of this Data Quality Assessment (DQA) is “to inform Mission staff of the strengths and weaknesses of each (selected) indicator and the quality of the data being collected.”

The Contractor was asked “to determine if the selected indicators (both qualitative and quantitative) for each Strategic Objective (SO) meet the following seven characteristics of good performance indicators as provided in **ADS 203.3.4.2**.

1. Direct: Performance indicators should closely track the results they are intended to measure.
2. Objective: Performance indicators should be unambiguous about what is being measured; unidimensional (should measure only one aspect at a time); and precisely defined in the PMP.
3. Useful for Management: Performance indicators should be useful for the relevant level of decision-making. Agency-level indicators may be included in the PMP for each SO.
4. Practical: Performance indicators should be selected if data can be obtained at a reasonable cost and in a timely fashion.
5. Attributable to USAID Efforts: Performance indicators should measure changes that are clearly and reasonably attributable, at least in part, to USAID efforts.
6. Timely: Performance indicators should be available when they are needed to make decisions.
7. Adequate: Mission should have as many indicators as are necessary and cost effective for management and reporting purposes. In most cases, two or three indicators per result (SO or IR) should be sufficient to assess performance.”

The Contractor was also asked “to verify that the data are of a reasonable quality based on the following five data quality standards provided in **ADS 203.3.5.1**.

1. Validity: Data should clearly and adequately represent the intended result. While proxy data may be used, the Contractor must consider how well the data measure the intended result. A key issue is whether data reflect a bias such as interviewer bias, unrepresentative sampling, or transcription bias.
2. Integrity: Data that are collected, analyzed, and reported should have established mechanisms in place to reduce the possibility that they are intentionally manipulated for political or personal reasons.
3. Precision: Data should be sufficiently precise to present a fair picture of performance and enable management decision-making at the appropriate levels. One key issue is whether data are at an appropriate level of detail to influence related management decisions. A second key issue is what margin of error is acceptable given the management decisions likely to be affected.

4. Reliability: Data should reflect stable and consistent data collection processes and analysis methods from over time. The key issue is whether analysts and manager would come to the same conclusions if the data collection and analysis were repeated. Contractor should be confident that progress toward performance targets reflects real changes rather than variations in data collection methods.
5. Timeliness: Data should be timely enough to influence management decision-making at the appropriate levels and to comply with required reporting needs. One key issue is whether the data are available frequently enough to influence the appropriate level of management decisions. A second key issue is whether data are current enough when they are available."

The Contractor collapsed these twelve criteria into ten, with the agreement of the Mission. They are: validity, reliability, timeliness, precision, integrity, directness, objectivity, practicality, adequacy, and usefulness.

## **B. Format of the Data Quality Assessment**

Following this introduction and a listing of the performance indicators included in the DQA, Section II presents a brief description of the methodology used by the Contractor. It then covers the indicators assessed for each of USAID/Ecuador's five strategic objectives. Wherever the DQA Team found significant weaknesses or strengths for an indicator related to a particular assessment criterion, summary findings and recommendations are made. In some cases there were no significant findings, and consequently no findings or recommendations appear in Section II. Annexes A through E, however, contain Data Quality Assessment Checklists for each indicator, giving more detailed comments on strengths and weaknesses regarding each assessment criterion.

All missions have been asked to include in this year's annual reports an ordinal scale weighting, between one and five, indicating the relative strength of the indicators included insofar as they measure performance. In Section III the Contractor recommends such a weighting for each indicator to be included in the Mission's upcoming Annual Report.

The Mission also asked the Contractor to make recommendations, based on the DQA, for any possible revision of its Performance Monitoring Plan. These are presented in Section IV.

## **C. Performance Indicators Covered in the Data Quality Assessment**

This Data Quality Assessment (DQA) covers all 23 indicators that USAID/Ecuador had designated for inclusion in its last annual report plus three additional indicators chosen jointly by the Mission and the Contractor. These indicators are shown on the following pages with their respective strategic objectives or intermediate results.

<b>Result</b>	<b>Selected Indicator</b>
SO 11: Improved social and economic conditions of inhabitants along the Peru-Ecuador border thereby promoting border integration	Number of beneficiaries in the target region whose lives are improved by participation access to social services, adoption of improved NRM practices, and/or more effective local governments
IR 11.3: Increased availability and access to social services, with emphasis on health, water, and sanitation	Number of new/improved potable water systems
IR 11.3: Increased availability and access to social services, with emphasis on health, water, and sanitation	Number of new/improved sanitation units
IR 11.4: Improved natural resource management in selected areas along the border	Number of hectares of land under legal title and/or near buffer zone
IR 11.CCR - Improved capacity of local governments to plan projects and implement services	Number of local governments that have improved service delivery management
IR 12.1.1: Implementation of new accusatorial justice system advanced	Progress in meeting selected benchmarks in implementation of new CPC achieved
IR 12.1.1: Implementation of new accusatorial justice system advanced	Percent of justice personnel sub-groups given advanced training
IR 12.1.3: Targeted local governments more effectively respond to community needs	Number of USAID-aided municipalities implementing or improving at least one service chosen with broad citizen input, including marginalized
IR 12.1.3: Oversight capacity of civil society increased	Number of USAID-supported municipalities where citizen groups are overseeing targeted services
IR 12.2: Greater Inclusiveness of Democratic Process	Increased number of defendants, especially from vulnerable groups, receiving defense services from legal services providers that have been strengthened by USAID-funded assistance
IR 12.3.2: Decentralization for strengthened sub-national government advanced	ARD-defined interventions to address key policy bottlenecks
IR 12.3.3: Mechanisms for coalitions and consensus building strengthened	Justice Coalition is (a) established and (b) effectively lobbying for Justice Reform
SO 13: Improved quality of life of the population living along the northern border	Number of beneficiaries in the target region whose lives are improved by participation in access to social and infrastructure services
SO 13.1: Health conditions improved in vulnerable villages and towns	Number of village potable water systems constructed/repaired/expanded in target areas
IR 13.1: Health conditions improved in vulnerable villages and towns	Number of water boards created, legalized, trained and functioning, by category A and B
IR 13.1: Health conditions improved in vulnerable villages and towns	Number of Water Administration Units in municipalities created and/or strengthened

<b>Result</b>	<b>Selected Indicator</b>
IR 13.2: Roads and related infrastructure improved in vulnerable areas	Number of irrigation systems repaired and number of hectares of land irrigated
IR 13.3: Civil society strengthened to better respond to local needs and to the challenges of Plan Colombia	Number of students, teachers, and parents participating in drug education/prevention program
IR 14.1: Increased Access to Microfinance Services	Increased number of microfinance credit clients in selected microfinance institutions
IR 14.1.1: Facilitative financial policy, regulatory and support environment created	Appropriate microfinance, regulatory and supervisory mechanisms established and/or implemented
IR 14.1.2: Strong, sustainable microfinance institutions	Average portfolio at risk maintained below 5%
IR 14.1.2: Strong, sustainable microfinance institutions developed	Financial sustainability of microfinance institutions increased to 100%
IR 14.2.1: Macroeconomic Policies/Environment for More Equitable Growth	Increased tax revenue collections as percentage of annual GDP
IR 1.1: Strengthened capacity of targeted NGOs and CSOs active in biodiversity conservation	Increased financing of local partners by outside (non-USAID) sources
IR 1.2: Economically viable natural resources management (NRM) practices adopted, in selected protected areas and their buffer zones	Hectares of land in selected protected areas and buffer zones under participatory NRM plans
IR 1.3: Key policies and legal framework introduced and/or implemented to conserve biodiversity in selected protected areas and their buffer zones	Key policies, legal frameworks, and enforcement mechanisms prepared, modified, and introduced at appropriate government level to improve biodiversity conservation in targeted protected areas and their buffer zones

## II Summary Performance Indicator Quality Assessments

### A. Introduction

#### 1. Methodology of the Data Quality Assessment

USAID/Ecuador contracted two Performance Monitoring Specialists, Harry Carr as Team Leader and Clarence Zuvekas, from the Checchi and Co. Consulting/Louis Berger Joint Venture, to carry out a Data Quality Assessment for its five Strategic Objectives (SO). The Data Quality Assessment (DQA) Team spent three weeks in Ecuador between November 28 and December 20, 2003. The overall purpose of the DQA has been described above in Section I. The Team interviewed USAID officials at all levels related to each SO as well as the staffs of all Implementing Partners (IP) in Quito. The DQA Team spent a total of five days in the field observing SO activities and facilities, interviewing IP field staff, reviewing field documentation, and interviewing activity beneficiaries. The DQA Team then spent five days synthesizing its findings and recommendations, writing them up, and presenting them to USAID/Ecuador officials.

The DQA Team focused on 26 performance indicators that were either included in last year's USAID/Ecuador Annual Report or considered as probable for inclusion in next year's Annual Report.

Detailed observations, findings and recommendations about the quality of each performance indicator over ten data quality assessment criteria were recorded in Data Quality Assessment Checklists, attached in Annexes A through E. These criteria for judging an indicator's quality, prescribed in ADS 203.3.3 and included in the Scope of Work, are: validity, reliability, timeliness, precision, integrity, directness, objectivity, practicality, adequacy, and usefulness. Several of these criteria or sub-criteria refer to indicators that are measured quantitatively through integral scaled data and/or indicators that use sampling methods to gather data. These criteria generally did not apply to any of the indicators chosen by the Mission for review in the DQA; therefore, most items in the Checklists related to them are left blank.

Summary findings and recommendations are given below.

### B. SpO 11: Southern Border Integration

- 1. Number of beneficiaries in the target region whose lives are improved by participation access to social services, adoption of improved NRM practices, and/or more effective local governments,** referring to the Strategic Objective:

*Improved social and economic conditions of inhabitants along the Peru-Ecuador border thereby promoting border integration*

Only the following data quality assessment criteria apply to this indicator. Findings and recommendations are given for each.

**Validity**

The indicator counts community inhabitants who benefit from program interventions in potable water, garbage collection, or bathroom installation. There is a valid relationship between most program interventions and those identified as having benefited, e.g. homes connected for potable water. There is not such a valid relationship in the case of indirect beneficiaries, for example under lands titled, included in the indicator.

- The SO Team should count only inhabitants who receive direct benefit from program interventions.

The indicator counts inhabitants for each service they receive. Thus there is double counting. The Mission is aware of this and has chosen to continue the practice under the rationale that it is really the level of benefit that should be measured by this indicator. Insofar as one family receives multiple benefits, the Mission feels they should each be counted.

- The DQA analysts agree with this interpretation of the indicator in most cases. However, they recommend not double-counting beneficiaries of two or more interventions related to the same activity, e.g. beneficiaries of a water system also receiving training in the health benefits of potable water.

The indicator counts equally people who receive potable water with people who live on land for which they have received titles. The nature and value of these benefits is clearly different.

- The Mission should work with CARE to develop a scheme for distinguishing levels of benefit, at least between land titling and other social services.

**Reliability**

CARE is at a point where it is changing the way it counts service beneficiaries. Previously it multiplied the number of connections or houses serviced by an estimate of household size – five inhabitants per house – as beneficiaries serviced. After its Mid-Term Evaluation in August 2003, the method was changed so that the inhabitants of all houses served are actually counted. However, these figures come from the household survey taken during the service diagnostic study stage, before the service is installed, and the number of beneficiaries can change by the time the service begins. This difference is not picked up in the count but in the aggregate is probably small.

- CARE field staff should revise its beneficiary count by periodically updating its household survey. In some cases this may need to be done on a sample basis, as for example with solid waste collection.

**Precision**

As noted above under Reliability, until September 2003 the method of calculating beneficiaries was to multiply the number of houses receiving services by an average household size estimator. After this practice was criticized in the Mid-Term Evaluation, CARE started recording the actual number of residents in each house served in the service agreement document.

- The Mission and CARE should continue this practice, as it gives greater precision and permits closer verification.

For community-wide benefits under such activities as land titling and garbage collection systems, CARE now takes an actual survey of the community to ascertain the number of beneficiaries, whereas in the past the IP used the Ecuador Social Indicator System, a secondary source of data. This was also done in response to a recommendation of the Mid-Term Evaluation.

- The DQA analysts concur in this recommendation as adding greater precision and certainty to the data generated.

#### **Adequacy**

The intended objective of the SpO -- *Improved social and economic conditions of inhabitants along the Peru-Ecuador border thereby promoting border integration* – should be clarified in order to ascertain how adequately the indicator measures the SpO. The DQA analysts were told by the SpO Team that this SpO is more transitional or recovery in nature rather than developmental. This distinction comes from its origins in stabilizing the southern border after the last Ecuador-Peru war.

- The Mission should make it clear in the PMP whether the objectives of this SpO are transitional or developmental. If they remain transitional then it would seem that the indicator is adequate to measure its intended objectives. If the objectives have turned more to development of the area and longer-range objectives now obtain, then the Mission should add indicators that measure those achievements on a impact level, e.g. health, economic growth, and environmental conservation of titled lands.

#### **Usefulness**

This is really a service delivery output measure. It does not speak to the impact of those services: whether bathrooms are maintained by the household, whether they have stimulated other “well-being advancement” practices in the household, impact on health, etc. If the Mission only wants to measure coverage of selected services, as the SpO implies, then this is a useful measure.

## **2. Number of new/improved potable water systems, measured by the Intermediate Result:**

*Increased availability and access to social services, with emphasis on health, water, and sanitation.*

This is an input measure of infrastructure provided. The same general findings and recommendations made above for the SpO above apply to this indicator as well, since they stem from a need to clarify the nature of the SpO itself. If the SpO is transitional, then the adequacy and usefulness of the indicator are fine. If, however, it is developmental in nature, then the adequacy and usefulness of the indicator lack some measure of the expected developmental impact of the infrastructure.

#### **Adequacy**

See ADEQUACY under the SpO indicator above.

#### **Usefulness**

See USEFULNESS under the SpO indicator above.

**3. Number of new/improved sanitation units**, referring to the Intermediate Result:

*Increased availability and access to social services, with emphasis on health, water, and sanitation*

This, like the potable water indicator above, is an input measure of infrastructure provided. Indeed, potable water systems and bathrooms are generally installed in the same locations (towns). Consequently, the same findings and recommendations made above for the SpO above also apply to this indicator.

**Adequacy**

See ADEQUACY under the SpO indicator above.

**Usefulness**

See USEFULNESS under the SpO indicator above.

**4. Number of hectares of land under legal title and/or near buffer zone** measuring the Intermediate Result:

*Improved natural resource management in selected areas along the border*

The DQA Team makes the following observation about this indicator.

**Directness**

Although this is a relatively direct measure of the intended result, it is a distant milestone measurement of it. The more relevant measure of the achievement of IR 11.4 is the companion indicator measuring hectares under Integrated Management Plans.

**5. Number of local governments that have improved service delivery management**, referring to the Intermediate Result:

*Improved capacity of local governments to plan projects and implement services*

Findings and recommendations are given for each DQA criterion applying to this indicator.

### **Validity**

The Mid-Term Evaluation concluded that this indicator is too complex, required too much “paperwork” and should be eliminated. A great deal of the validity of this measurement approach – using benchmark accomplishments as indicators of improvement – stems from a judgment about how well the benchmark events actually measure an important dimension of “improvement.” Some benchmarks are better, in the judgment of the DQA Team, than others. Some are very good ones, representing “proxy” measures of multidimensional variables. For example, the verified ability of a municipality to have an effective, working cost-accounting system that can produce an estimate of the real costs for any given service is quite a powerful achievement; moreover, to have achieved it implies the accomplishment of a series of lower-order achievements, such as an effective accounting system, that are also major accomplishments. Similarly, a working, centralized management information system requires several other lower-order abilities in the municipal administration, such as computer literacy and capacity of all related municipal staff.

It is recommended that the measures be improved, not eliminated. Rather:

- A more appropriate way could be found to measure the various characteristics seen as together representing an advanced stage of institutional capacity for municipalities at different levels of development;
- Improvement could be made in some of the definitions of characteristic benchmarks, e.g. “strengthened municipal structure” and “focus of equality in municipal action” for Level Three municipalities;
- The practice of counting any eight out of ten benchmark events to qualify a municipality in Level Three as “improving” should be made more rigorous. All the threshold events for Level Three municipios are not of the same “degree of difficulty,” so to speak. An effective cost-accounting system, as noted about, is harder to achieve than “an improved municipal structure,” particularly when the achievement of each benchmark is itself a bit of an arbitrary judgment; i.e. an ordinance restructuring the tax administration division by adding a particular sub-division to cover the collection of a new tax would qualify as accomplishment of this benchmark.

### **Reliability**

The qualification of a municipality as having been “improved in its capacity to manage service delivery,” as measured by CARE, is a complex benchmark-type indicator. It recognizes that, because of the differences between municipalities regarding geographic and population size and other variables, all municipalities cannot be measured in the same terms. Improvement in larger municipalities that offer more services is a different kind of thing from improvement in smaller ones. So CARE has arbitrarily established three levels within each of which “improvement” is measured or “benchmarked” differently. Smaller – Level One - municipalities, would have to accomplish three things to be judged as having “improved” from one reporting period to the next. (See the PIRS in Annex B for the precise definition of this indicator.) For Level Two municipalities the bar is raised to five particular accomplishments, while for Level Three municipalities it is raised to any eight out of ten particular accomplishments. Thus, the placement of the municipalities with which CARE is working into these levels is arbitrary but reflects a valid reality.

The accomplishment of each particular threshold is marked by a particular, concrete, and verifiable event or document. The CARE team is aware that there are “sub-threshold events” leading up to each of these measured benchmarks. As the CARE team works with each municipality, the responsible staff member is able to judge whether and to what extent the municipality is reaching each prescribed benchmark by observing the sub-threshold events.

- After site visits, the DQA Team finds that the CARE team has made reliable judgments based on their expertise and close participation with subject municipalities, and that the systematic, although apparently complicated, way CARE uses to break down its benchmarks is well implemented and produces reliable performance measurement results.

### **Directness**

How closely is "Improved capacity of local governments to plan projects and implement services" measured by "Number of local governments that have improved service delivery management"?

The indicator, as formulated, measures improved municipal capability:

- in its facility with and use of computerization in administration;
- in its administration, particularly financial administration;
- in its capacity to plan; and
- in its capacity to administer selected services.

As such it is a direct measure of the intended result. However, it is essentially an input measure. It reflects how well a municipality is prepared to implement services without also measuring the quality or coverage of those services. It does not measure the implementation part of the intended result.

- A periodic household survey would provide insight into the coverage and quality of services delivered.

### **Objectivity**

The indicator identifies particular kinds of documentation, e.g. *actas de entrega recepcion*, *certificados de aprobacion*, or *ordinances*. In some cases these documentary benchmarks may not actually represent the establishment of a plan or implementation of a service. There is a "case by case" nature to the objectivity of these measures. During the DQA, for example, a school sanitary unit (bathroom facility) and a potable water system were observed. The *actas de entrega recepcion* for these projects were also obtained. They affirm all the terms of the work, including size and components, number of intended beneficiaries, costs contributed by the various contributing implementers, etc. They contain all the relevant signatures. They seem to be valid and objective "markers" for the beginning of implementation of the projected sanitary unit and water system. Their dates correspond to the periods of construction. There may have been other steps, however, prior to implementation, that the DQA is not aware of and that would have made these measures approximate only.

- CARE staff should indeed treat these documentary benchmarks on a case-by-case basis and be assured that they are measuring what they appear to be measuring.

### **Practicality**

The aspect of timeliness, a principal sub-criterion of practicality, most relevant to this indicator is how close in time the benchmark measure is to the actual implementation of a service or a management information system or to the putting in practice of an equity classification system. For example, the DQA observed a municipal management information system that had been counted as being accomplished in the April-September Progress Report as per its *acta de entrega recepcion*. That document would have been dated sometime between April and September 2003. The system itself was not fully operational in December during the DQA and its was estimated to be "launched" in January 2004.

- For each benchmark reported as having been achieved, CARE should note some estimate for the beginning of implementation of the related system or service.

### Usefulness

The Mission should ask the Implementing Partner to describe constraints or problems encountered during the reporting period, along with the detailed description it now provides about benchmark accomplishments in each municipio. For example, CARE field staff may be aware of major barriers ahead for the approval of a particular strategic plan. The plan may have been completely formulated but may have hit a snag. That information should be included in the textual descriptions of the indicator as the indicator itself does not provide the information.

## C. SO 12: Democracy

### 1. Progress in meeting selected benchmarks in implementation of new CPC achieved, measuring the Sub-Intermediate Result:

*Implementation of new accusatorial justice system advanced*

The following observations, findings and recommendations, stated in summary form, were found regarding the seven indicators on which the DQA focused in the Democracy area.

### Validity

This indicator makes the major, implicit assumption that the final benchmark, "reformed CPC be evaluated," is a significant advancement in the "implementation of the new accusatorial justice system."

- It would be instructive if the Mission would describe what the significance of this stage really is and provide some idea of how much farther there is to go.

Although the indicator is a valid measure of the result, it measures (acknowledges) the passage of various process benchmarks. Since this is the Mission's major indicator for progress in the justice area:

- The Mission and IP should provide a very detailed description in the Performance Indicator Reference Sheet for this indicator (a preliminary version of which is attached to the DQA) of what is represented by the passage of each benchmark. What does it mean? What things have happened in the justice system because of the benchmark being reached? How has it affected citizens?

### Directness

Although the indicator is a direct measure of the result, the result itself is not self-evident in its significance. What is the significance of the "advancement of the new accusatorial justice system"? There is a certain amount of evidence in the DQA that advancing this process, aside from any real impact on citizens, may be what the Mission wants; that the official bodies in this process are reformed. This needs focus to be explicit.

- The Mission should meet with its IPs, consider carefully what is intended to be captured by the IR, and reformulate its wording accordingly.

### Objectivity

It is difficult to know when each benchmark has been reached. How does the Mission mark the passage of the benchmark in precise and objectively measurable terms? This is a matter of

giving very precise definition to each benchmark. As of the preparation of this DQA, the following is the most precise definition available for each benchmark:

- (i) "Inter-institutional committee coordinating implementation of CPC is functioning," with functioning being manifested by:
  - a. Production of reports, studies, and procedures manuals, and
  - b. Emanation of decrees (not laws);
- (ii) "Textual & rapid-appraisal CPC diagnostic completed," with no further clarification;
- (iii) "Results of diagnostic studies disseminated to policymakers," with dissemination meaning more than simply distributing the study results but also ensuring an understanding and discussion of them among entities, manifested by:
  - a. Physical dissemination of information, and
  - b. Holding presentations, conferences, meetings, etc. bringing entities together to promote coordination;
- (iv) "Reforms of new CPC are proposed," reflecting presumably a realization that the CPC needed refining, this benchmark means a reformulation or refinement of the CPC manifested by the document itself;
- (v) "CPC reforms passed," with no further clarification; and
- (vi) "Reformed CPC evaluated," also with no further clarification.

- Specifically, what are the concrete ways in which these events are defined and measured? What sort of "reform" qualifies for inclusion in the measurement of benchmarks (iv) and (v)? What fora and conferences count for benchmark (iii)? The purpose of precision in definitions is to demonstrate the significance of the benchmark and to show that it was objectively marked.

#### **Adequacy**

The adequacy of this indicator and the others that measure performance against the IR really depends on a more precise definition of the IR itself, as described above under "Directness."

#### **Usefulness**

The Mission would gain considerably in its ability to monitor and really know if it were on track if the precision recommended above were given to the indicator.

## **2. Percent of justice personnel sub-groups given advanced training, referring to the Sub-Intermediate Result:**

*Implementation of new accusatorial justice system advanced*

The DQA Team makes the following summary recommendations regarding this indicator.

#### **Validity**

- Target and train justice personnel that directly contribute to the implementation of the CDC, in order to ensure that the activity and measure contribute as much as possible to the IR achievement;
- Have the training be specific to the CDC and its procedures in order to ensure contribution to achievement;
- Consider the possibility of weighting justice professionals who are trained, so that the indicator more accurately measures the IR's achievement;

- Consider weighting the kind of training received – varying degrees of “advancedness” – for the same reason as above, to increase the accuracy of the measure;
- Be sure of the baseline for measuring the indicator, to ascertain just how far off “appropriate” training levels the justice personnel were (also checking the assumption of this activity and indicator); and
- Define a precise, effective working definition of and criteria for training that is classified as “advanced.”

#### Reliability

The types of training that might be classified as advanced must be consistently applied. They should be defined exhaustively and written up so as to ensure consistent application.

#### Directness

The Mission and IP assumption about a low skill level among justice personnel should be tested. Certainly such a study or research would help establish the baseline. The assumption is so central to the activity and measure that verification would add considerable weight to the reliability of the indicator.

### 3. Number of USAID-aided municipalities implementing or improving at least one service chosen with broad citizen input, including marginalized, measuring the Sub-Intermediate

Result:

*Targeted local governments more effectively respond to community needs*

The following summary recommendations were made by the DQA Team regarding this indicator.

#### Validity

- Citizen participation is not a big part of this indicator, as it really is a measure of municipal improvement. Citizens are not formally part of the MOU and do not have formalized roles in overseeing service-improvement implementation. The program does allow for citizen involvement and appears to encourage it, holding workshops and meetings. In municipios visited the involvement of the MPA Groups, as they are called after ARD’s citizen participation process, was very active indeed, and the groups seemed well integrated into the municipal management of the process. But progress itself does not depend on the involvement of citizens. The following recommendations are made in an attempt to either make the measure more precise or reformulate its wording.
  - Greater involvement could be formalized and even required by participation in the MOU.
  - There could be a requirement that only groups form membership in the Working Group (not individuals), thus ensuring continuity and broader-based citizen involvement. The Working Group would be an association of associations interested in and actively overseeing the service improvement process.
  - ARD has no way to ascertain if or which marginalized parts of the municipal population are involved in the process, so some objective way should be designed to ascertain this involvement.

- Reformulate the wording as *Targeted local governments more effectively provide selected services.*
- It is recommended that the Mission work with ARD to adopt a new measure as its principal IR 12.1.2 performance indicator. Two suggestions are:
  - ARD has another, activity-level indicator, reported to USAID/Ecuador on a semi-annual basis, that would be a better measure – more precise, direct and valid – of improvement in municipal management of a selected service: ***Increase in average (mean) score of all USAID-aided municipalities on institutional strengthening report card.*** (See the *3D Project Performance Monitoring Plan, Semi-Annual Review, April-September 2003.*)
  - In one of the early stages of development of a selected service, a diagnostic study is produced. These typically identify detailed work plans and milestone outputs, sometimes with deadlines, which might be used as the basis for measuring the municipality's performance in improving the service. (The disadvantage of this indicator is that it normally will end up measuring the performance of the contractor hired to carry out the service improvement.)

**Reliability**

Quantitative and qualitative data for activity-level indicators are collected systematically. However, it also appears that they are not used, calculated, or combined in any systematic or disciplined way so as to yield reliable composite and comparable measurement over the various municipalities. In the end, the judgment of "improvement" is made by ARD field staff based on all this available and observable information.

- Make the judgment as systematic and comparable as possible.

**Precision**

The most precise way of measuring an improvement would be a milestone measure of selected municipios' progress in improving selected services. Progress could be measured either by the "report card" grade already taken periodically or by advancing in "intervention stages," although the Report Card is partially a function of advancing in stages of intervention. The IP collects enough information, quantitative and qualitative, and has enough close working contact with the municipios, to make such a systematic judgment as a measure of this indicator.

**4. Number of USAID-supported municipalities where citizen groups are overseeing targeted services, referring the Sub-Intermediate Result:**  
*Oversight capacity of civil society increased*

The following recommendations can be made regarding this indicator.

**Validity**

Right now the indicator counts oversight as occurring if ARD is working with the municipality and if a Working Group has been formed. It makes the assumption that if a Working Group has been formed then it is effectively overseeing the municipality. But there is no good definition of what oversight is. And the only way that oversight is measured is through the number of meetings held in which both municipal officials and Working Group members are present. Finally, the membership of the Working Group can change over time, so continuity can be lost. Even if a good definition of intended oversight existed and there were ways to measure it, if the membership changes then oversight can fall backwards.

- Develop a better working definition for “oversight” based on the Performance Indicator Reference Sheet (PIRS) prepared during the DQA.
- Develop ways to effectively measure it, i.e. how well Working Groups are providing that “oversight” beyond meetings held.
- Find ways to require or encourage continuity of membership in the Working Groups.
- Add the qualifier “effectively” to the oversight provided and measure it in a qualitative way, e.g. participation in decision-making meetings, advancement in the Working Group’s Action Plan, etc.

**Directness**

Most of the observations and recommendations made under Validity apply to Directness. “Yes” it is direct, but by definition.

- Give the indicator more precise definition, which may mean actually reformulating it.

**Practicality**

Its lack of precision makes the indicator impractical for the Mission.

**Usefulness**

It is recommended that it be replaced.

**5. Increased number of defendants, especially from vulnerable groups, receiving defense services from legal services providers that have been strengthened by USAID-funded assistance, measuring the Intermediate Result:**  
*Greater Inclusiveness of Democratic Process*

The DQA Team is able to make the following summary recommendations on this indicator, which has only recently been identified and is not yet precisely defined. Because it is new and still being defined, the DQA Team has made extensive recommendations.

**Validity**

This is a new indicator for the Mission, which has not designed its exact content or definition. Nor has the activity been designed as of December 2003. The Mission and Implementing Partner (IP), however, have decided that the activity will go forward and consequently have agreed to do a DQA for it. The indicator has the following weaknesses in terms of its validity.

- The SO itself ends in 2005. This activity is therefore getting a late start. Presumably it will be designed in early 2004, with the Mission and its IP having done a fair amount of definitional thinking. But in the best of all worlds the activity itself – strengthening legal services providers – will not begin until mid-2004. Assuming the strengthening, through training and TA at least, will take several months to reach fruition, the Mission is looking for this activity to start taking effect no sooner than late 2004. If the indicator has not been reformulated over the course of those months in 2004 and if baseline data have been collected, then it will be a matter of “good luck” if the indicator remains a valid measure of the activity, i.e. if the indicator as formulated matches the activity as it becomes

implemented.

- It is recommended that the Mission and IP begin the activity and indicator formulation process as soon as possible and in tandem, refining the indicator where it may be necessary to match the activity's design.
- The validity of this indicator is also questionable regarding the extent of its impact. The essence of the Intermediate Result – *greater inclusiveness*—implies that the Mission knows something about the extent of the problem of non-inclusion of defendants unable to pay for legal services. Just how much of the universe of vulnerable defendants can this activity be expected to address? And just what is the universe, i.e. what is the range of criminal cases covered? Using “number of defendants” is probably less powerful an indicator than “percentage,” but whether it is possible or makes sense is uncertain prior to the baseline study. However it is measured, the uncertain and potentially large “waterfront” to be covered will make the validity of the indicator suspect.
  - The Mission should consider using percentage of vulnerable defendants if this is possible and if the baseline study shows it would make sense.

The Mission has asked that the DQA be carried out for the indicator as formulated and that the DQA process produce guidance to the SO team on any implications for the activity's design stemming from the indicator's formulation. Those implications are shown below in summary form.

- The activity will have to have an effective “vulnerability test” or “means verification method” for a defendant receiving legal services to ensure that he or she qualifies for the legal services.
- There may end up being an array of “capacity strengthening” actions taken by the activity *vis a vis* service providers. Exactly which of those strengthening actions and how many of them will qualify a service provider to fall within the definition of the indicator should be determined early in the activity design, i.e. for a service provider to be counted as having been “strengthened,” he or she should have received actions *x*, *y* and *z* in order to achieve some predetermined amount of capacity to represent defendants effectively.
- In identifying the capacity strengthening “threshold” levels, the Mission should keep in mind the need to have data flowing as early as possible due to the late start the activity will have.
- The Mission and IP need to be clear about who a “defendant” is, not just for the sake of measurement but as a key decision as to where the activity will intervene in the criminal justice process.
- The Mission and IP should do the baseline as soon in 2004 as possible. For this indicator, target setting will almost certainly depend on baseline data.

### **Reliability**

It is important that there be consistent application of clearly elaborated definitions in this indicator. The indicator, and presumably the activity it will be measuring, does a lot of critical targeting, e.g. defendants appears to mean (1) those who have formally gotten caught up in the criminal justice system as opposed to those at risk, in some way, of getting involved, and (2) those who cannot afford legal services.

### **Timeliness**

Early availability of performance data on this indicator will be of critical importance because of the late implementation start (see “validity” above) it is getting (2004 effectively) *vis a vis* the end of the SO in 2005.

- The Mission should keep this need for quick data in mind when it establishes the “package” of strengthening actions and “threshold” level of capacity building that the Mission will consider as qualifying a service provider for counting under this indicator.

### Directness

This indicator measures the extent to which the criminal justice system is open to defendants who cannot otherwise afford the necessary legal services. The activity seeks to increase the number of legal services providers, and thus provide those services to needy defendants. However the criminal justice system is only one part of the “democratic process” referred to in the IR that seeks “greater inclusiveness.” Although the indicator is a direct measure of a more inclusive criminal justice system, it is only a partial measure of more inclusive democratic processes in general, which is the intended result. Therefore its directness is reduced.

- The Mission should consider refining the wording of the IR to better match the activity.

(This reflects a larger problem with the Mission’s Democracy Results Framework. The SO – *Increased Support for the Democratic System* -- is very high. It really is focused on –addresses – the criminal justice system and local government. This focus is shown at the IR and sub-IR level of the framework. Making the reformulation recommended above will make that focus explicit for IR 1.2 and complement the implied focus shown in IRs 12.1 and 12.3 with the words “of key democratic institutions” and “in key democratic areas,” respectively.)

### Adequacy

The activity for which this indicator measures progress works outside of the official criminal justice system, in that it seeks to increase the supply of legal services provided by non-governmental entities. Yet the Intermediate Result seems to address the official processes and institutions in the criminal justice system, seeking to make them more “inclusive.” Increasing the supply of legal services for indigent defendants certainly makes the criminal justice system more inclusive, but without doing anything to the formal system such as working to establish a “public defender’s office” within the justice administration infrastructure.

- The Mission and IP should consider collateral efforts under this IR that address the official justice system and at the same time complement its “supply-side” effort. For example,
  - study and identify constraints in the formal system to using the legal services it is developing and strengthening;
  - work to diminish those constraints; or
  - work to make formalized linkages between the legal service providers USAID is working with and the formal system, such as a referral service within the formal court system.

## 6. ARD defined interventions to address key policy bottlenecks, referring to the Sub-Intermediate Result

*Decentralization for strengthened sub-national government advanced*

The following recommendations are made regarding this indicator.

### Validity

The indicator marks and affirms the first major step in implementing effective decentralization: That ARD/3D identifies major bottlenecks to implementation and recommends actions to be taken. Once having been achieved, however, there is no value in the indicator for measuring progress. The “Bottleneck Report” identified six constraints to decentralization and described ARD/3D strategies for overcoming these constraints.

- ARD could track its success in applying its strategies as a way of continuing to measure the intended result.

As another alternative to more measurement of the achievement of IR 12.3.2, ARD/3D has zeroed in on at least one particular aspect of the constraints identified – the need for citizen participation in the implementation process – and has recommended the establishment of a Citizens Advisory Commission.

- Since the “bottleneck report” recommends actions to be taken to overcome constraints to decentralization, and since ARD/3D has focused its attention on advancing the citizen participation aspect, the IP could define a follow-on indicator that tracks its success in advancing citizen participation. Indeed, ARD states in its semi-annual report, *“Advisory Commission exists and is advising ARD on key policy bottlenecks.”* This could be worked into a complementary indicator for indicator IR 12.3.2.

Not only production of the “bottleneck report” but also its dissemination is implicit in the indicator. Dissemination is understood by the IP and Mission in this case to mean not just distribution of the printed report but also “public education” among policy makers and civic leaders as to what it means and how the recommended interventions are to be carried out. However, dissemination is not defined (how broadly distributed? how thoroughly should recommendations be understood?) nor is it measured.

- Since understanding is an important, although implicit, part of this indicator for the Mission, there should be an effort to define, measure, and track it over time. This would also “stretch” the impact of the indicator farther into the future.

#### **Directness**

As was discussed under the subject of validity, this indicator is a direct but very distant measure of progress against the intended result is at the sub-IR level – “Decentralization for strengthened sub-national government advanced.” It is more distant from achievement of the IR: *“Increased policy consensus in key democratic areas.”*

#### **Objectivity**

It is recommended that the Mission vet the “bottleneck report” among technical experts in decentralization in Latin American countries.

#### **Usefulness**

Because the indicator only measures one major accomplishment early in the process of decentralization, it really fails to allow the Mission any ability to “track” or monitor progress against its results. “Yes, the bottleneck report was completed and disseminated.” “And then what?”

- The Mission and its IP possibly have an opportunity to take the bottleneck report farther, as discussed under validity. Having identified bottlenecks and interventions, they can select a particular intervention (possibly an aspect of citizen participation or interventions focused on local government) and design activities for advancing that intervention as well develop an indicator for measuring its progress. In designing the activity and indicator:
  - Be certain that the particular intervention selected is directed at areas of intervention the Mission is following anyway, for example the two mentioned above;
  - Be sure that the intervention is within the Mission’s manageable interests; and
  - Be sure that any change in the indicator can be attributed to Mission inputs.

**7. Justice Coalition is (a) established and (b) effectively lobbying for Justice Reform**, referring to the Sub-Intermediate Result:

*Mechanisms for coalitions and consensus building strengthened*

The following observations recommendations are made regarding this indicator.

**Validity**

There is a valid relationship between the indicator and the result its measures but the relationship is not so much one of causality, that one causes the other, as identity, that one is an example of the class named in the other. Its validity is superficial.

- The indicator should be reformulated so that it measures, perhaps, the establishment by the Coalition of other consensus-building, civil-society associations, or the influence the Coalition has had on the building of consensus and action on justice issues.

The indicator, as formulated, actually has more validity as a measure of the Sub-IR under IR 12.3 -- *Increased policy consensus in key democratic areas* – rather than this Sub-Sub-IR.

- It should be moved to the sub-IR level if
  - (1) its measurement were made more precise and reliable (see below) and
  - (2) it were focused on policy achievement of the Coalition rather than its functioning.

**Reliability**

The indicator measures two major milestones: the establishment of the Coalition and its functioning as a lobbying entity. These have both been classified as achieved in 2003. The activity extends to 2005. Fundacion Esquel has an extended list of sub-milestones that manifest “establishment” and “functioning.” These include:

- at least 3 launch events that the Coalition has organized;
- at least 4 working sessions with the Commission for the Implementation of theCodigo de Procedimiento Penal (Criminal Procedures Code);
- at least 5 documents about subjects related to justice that the Coalition has produced and distributed;
- at least 4 working meetings of the Coalition itself;
- at least 4 workshops for the discussion of specific justice themes that the Coalition has organized;
- an agenda of themes necessary for justice-sector reform elaborated and distributed; and
- a functioning Secretariat.

However, these milestones are not systematically reported on as such to USAID. Events, documents, workshops, etc. of varying nature and size are reported on a quarterly basis. Whether they are considered as counting towards the indicator is not noted. The actual classification, in 2003, as “Yes – Coalition established” and “Yes, Coalition functioning,” was a judgment made by the Fundacion Esquel (USAID’s IP) Director.

- If the indicator were to be revised and maintained at the sub-sub-IR level, then the specific events counted as qualifying as milestone measures of the “establishment” and “functioning” of the Coalition should be explicitly identified and counted as such in reports.

### **Integrity**

See OBJECTIVITY below.

### **Objectivity**

The Fundacion Esquel serves as the Secretariat of the Coalition. This means that the "functioning" of the Coalition is dependent mainly on itself -- the Fundacion Esquel. This makes for a severe "sustainability" issue and a constraint to the long-term accomplishment of the intended result.

- Fundacion Esquel should formally relinquish its role as the Coalition's Secretariat and help set up a new one, using the methods and procedures it has already developed.
- The "establishment" and "functioning" of a new Secretariat might, indeed, be a more reliable and certainly meaningful measure of the achievement of the sub-sub-IR.

### **Usefulness**

The fact that milestone accomplishment data is reported so frequently AND in an unspecified manner (see RELIABILITY above) makes it really useless to USAID.

- A. See the recommendations under RELIABILITY above.

## **D. SO 13: Northern Border Development**

### **1. Number of beneficiaries in the target region whose lives are improved by participation in access to social and infrastructure services,** measuring the Strategic Objective:

*Improved quality of life of the population living along the northern border*

The DQA Team is able to make the following summary recommendations regarding this SO-level indicator. This indicator was not included in the Mission's reporting for 2003, nor will it be in the 2003 Annual Report. However, the Mission is considering using it in future years, first, because it is at the SO-level (see recommendations made in Section III below regarding the modification of the PMP) and, second, because a similar variable is reported for SpO 11 – Southern Border Integration – and the Mission is seeking comparability between them.

### **Validity**

Double-counting of beneficiaries of closely related activities, such as the water-systems example referred to in detail in the DQA Checklist, should be eliminated. In other cases, it seems best not to eliminate double-counting (\*) but rather to regard the indicator as consisting of "units of person-benefits," recognizing that these units have unequal value. At the same time, the Mission might indicate in a footnote the rough order of magnitude of the double-counting it believes is likely to exist. (\*Experience in other countries suggests that eliminating double-counting of

beneficiaries of unrelated projects in a particular community can be difficult and costly, even if the database is good.)

#### **Directness**

See recommendation for VALIDITY above.

#### **Objectivity**

The Mission will need to examine disaggregated figures by individual activity to see whether, e.g., an overall figure that exceeds the target hides shortfalls in targets for some specific activities.

The Mission should work closely with IOM to reach agreement on what additional double-counting can be eliminated.

#### **Usefulness**

See recommendations for VALIDITY and OBJECTIVITY above.

The Mission should seek uniformity in the definition of its SO-level indicators for Northern Border Development (SO 13) and Southern Border Integration (SpO 11).

## **2. Number of village potable water systems constructed/ repaired/expanded in target areas, referring to the Intermediate Result:**

*Health conditions improved in vulnerable villages and towns*

The following recommendations are made regarding this indicator.

#### **Validity**

Add to the statement of the indicator: "and providing water to connected households on a regular basis."

Since the water systems vary significantly in size, it would be useful to disaggregate the indicator by "towns" and "villages," using a locally appropriate distinction. Also, the indicator itself should refer to "town and village" systems, not simply "village" systems.

The Mission should consider developing a new indicator(s) to measure some aspect(s) of performance (e.g., average number of days per year with no interruption of service, except for normal maintenance such as cleaning tanks; and (weighted) average percentage of users paying their bills on time).

Another possible indicator is average consumption per household, a figure that should be readily available given the presence of meters. Trends in this indicator, together with those in the percentage of households paying their bills on time, would provide a good indirect measure of trends in households' well-being. IOM can easily keep track of the number of water systems it has helped construct/repair/expand. Once the works are completed, IOM collects technical and other data monthly.

**Directness**

The Mission should consider adding a sub-indicator showing the number of beneficiaries (persons) actually obtaining water from the new/improved/expanded systems. (These data are already included as part of the SO-level indicator.)

The Mission should also investigate the feasibility of establishing an indicator that would measure improvements in health status. It is not clear whether adequate baseline data are available.

**Objectivity**

See recommendation for VALIDITY and DIRECTNESS above.

**Usefulness**

See recommendations for VALIDITY and DIRECTNESS above.

**3. Number of water boards created, legalized, trained and functioning, by category A and B, measuring the Intermediate Result:**

*Health conditions improved in vulnerable villages and towns*

The following summary recommendations are made regarding this indicator.

**Validity**

"Functioning" should be replaced by "functioning in a way that significantly improves water-service delivery," or some similar phrase. The indicator should show not only how many water boards are classified as "A" or "B," respectively, but also how many fall into categories "C" and "D." In addition, the definitions of these categories should not be relegated to a separate page.

**Directness****Objectivity****Adequacy****Usefulness**

See recommendation for VALIDITY above.

**4. Number of Water Administration Units in municipalities created and/or strengthened, referring to the Intermediate Result:**

*Health conditions improved in vulnerable villages and towns*

The DQA Team has made the following recommendations regarding this indicator.

**Validity**

The indicator should be rephrased and stated in such a way that it reflects the outcome of what the WAUs are expected to accomplish, according to a pre-established set of quantitative and qualitative criteria.

**Adequacy**

See recommendation for VALIDITY above.

**5. Number of irrigation systems repaired and number of hectares of land irrigated**, measuring Intermediate Result:  
*Roads and related infrastructure improved in vulnerable areas*

The following summary recommendations are made regarding this indicator.

**Validity**

An additional indicator might be added to show the number of irrigation systems that have established funds to provide resources for maintenance, watershed protection, and protection against theft (water diversion). IOM is promoting the establishment of such funds in 2004.

**Directness**

The Mission should consider developing an indicator of water availability or water use, depending on what kind of data are readily available.

**6. Number of students, teachers, and parents participating in drug education/prevention program**, referring to the Intermediate Result:  
*Civil society strengthened to better respond to local needs and to the challenges of Plan Colombia*

Regarding this indicator the DQA Team makes the following summary recommendations.

**Directness**

Now that the educational programs are well established, the Mission should seek to develop an indicator of their effectiveness, even though it is planning to drop this activity

**Usefulness**

See recommendations for DIRECTNESS above.

## E. SO 14: Poverty Reduction

### 1. Increased number of microfinance credit clients in selected microfinance institutions, measuring the Intermediate Result: *Increased Access to Microfinance Services*

The DQA Team reached the following summary recommendations regarding this indicator.

#### Validity

The indicator should reflect only the unit of measure and should not refer to "increases." In the Nov. 2002 PMP, the targeted increase is incorrectly stated as "approximately 350%"; the figure should be "approximately 250%."

#### Adequacy

Consideration might be given to disaggregating current MFI credit clients by loan account in constant (e.g. 2003) dollars. Possible categories are: <\$300; \$300-999; \$1,000+.

The Mission should consider further disaggregating the number of current credit clients (by loan amount) into male and female clients. WOCCU can now provide these data very easily.

Although the indicator is appropriate for measuring access to microfinance services, the IR is weak in that it does not focus on how clients' incomes have changed as a result of increased access. Ultimately, the Mission should want to know whether the MFI programs it is supporting are having a significant impact on poverty reduction. Therefore it should consider developing an impact indicator of this nature, recognizing that reliable indicators will probably have to be based on a sophisticated and relatively costly survey design.

### 2. Appropriate microfinance, regulatory and supervisory mechanisms established and/or implemented, referring to the Sub-Intermediate Result: *Facilitative financial policy, regulatory and support environment created*

The following summary recommendations can be made for this indicator.

#### Validity

Now that the three planned mechanisms (supervisory manuals, normative framework, credit referencing services) have been established, it would be appropriate to define, as clearly and objectively as possible against pre-established criteria, benchmarks that would indicate a desirable degree of implementation.

#### Timeliness

Progress toward any future, related indicator(s) dealing with implementation should be reported according to a regular schedule (at least quarterly). The nature of the indicator(s) would probably call for narrative reports specifically focused on movement toward benchmarks rather than quantitative measures of progress.

**Directness**

Any future indicator(s) related to implementation should include quantitative (if possible) or at least explicitly qualitative dimensions against which informed judgment would have to be exercised according to a pre-established set of criteria.

**Objectivity****Adequacy****Practicality****Usefulness**

See recommendations above under DIRECTNESS.

**3. Financial sustainability of microfinance institutions increased to 100%,** measuring the Sub-Intermediate Result:  
*Strong, sustainable microfinance institutions developed*

The following summary recommendations can be made about the quality of this indicator.

**Validity**

The indicator should reflect only the unit of measure and should not refer to increases. It could appropriately be stated as: "Degree of financial sustainability according to international standards (100% or more indicates achievement of financial sustainability)."

**Reliability**

DAI and WOCCU should calculate the average indicator for their respective cooperating institutions in the same manner. The consultants recommend that a weighted-average indicator be used, with each partner using the same weighting system.

Partners should make greater efforts to ensure that they are obtaining full coverage of significant financial and in-kind donations being received by the individual MFIs.

Data-quality problems should be described in more detail.

**Precision**

See first recommendation under RELIABILITY above.

**Objectivity**

The statement of the indicator should reflect (briefly, of course) the fact that it is based on an international standard of financial sustainability. It should also clarify what "100%" refers to—in this case the average degree of internationally-defined financial sustainability of the 15 participating MFIs. Adding "at least" before "100%" would further clarify the nature of the indicator. Again, we recommend that both partners use weighted-average indicators.

#### **Adequacy**

As noted under RELIABILITY above, the consultants recommend that both partners use a weighted-average indicator.

A sub-indicator would be useful to report how many MFIs fall below the target of 100%. This indicator could be reported quite simply as, e.g., "3/15," in preference to a percentage indicator that would not make clear how the number of MFIs was changing over time.

Regarding the impact of MFI lending on borrowers' incomes and quality of life, see the last recommendation under ADEQUACY for the first SO 14 indicator.

#### **Usefulness**

The Mission should regularly examine disaggregated data by individual MFI, since this is probably the most important broad indicator available to USAID management for indicating whether strong, stable MFIs are being developed.

#### **4. Average portfolio at risk maintained below 5%,** referring to the Sub-Intermediate Result:

*Strong, sustainable microfinance institutions*

This indicator was not reported in the Annual Report for 2002, nor will it be included in the 2003 Annual Report. However, the Mission is considering using it for the indicator in the 2004 Annual Report. The DQA Team agrees with the Mission's plan to include it. The following recommendations are made regarding its quality.

#### **Validity**

The consultants recommend that both DAI and WOCCU calculate a weighted-average figure for their respective cooperating institutions, using the same weighting system.

A useful sub-indicator would be the number of MFIs in relation to the total (e.g. 4/16) that exceed the 5% target (i.e. have an undesirable percentage of their portfolio at risk).

The indicator should reflect only the unit of measure and not incorporate the target.

#### **Reliability**

See first recommendation under VALIDITY above.

Data-quality problems should be discussed in more detail.

#### **Precision**

See first recommendation under VALIDITY above.

#### **Objectivity**

Especially for new MFIs entering the program, partners should ensure that the indicator is interpreted correctly. The potential problem here does **not** lie with the precision of the indicator.

**Usefulness**

The Mission should regularly examine disaggregated data by individual MFI.

**5. Increased Tax Revenue collections as percentage of annual GDP**, measuring the Sub-Intermediate Result:  
*Macroeconomic Policies/Environment for More Equitable Growth*

The DQA Team makes the following summary recommendations for this indicator.

**Validity**

The indicator can be retained, given the nature of USAID-supported activities under IR 14.2; but its limitations as a contributor to the SO should be recognized. An indirect indication of changes in the equity of the tax structure could be provided by disaggregating the tax/GDP ratio (which covers only the taxes collected by the SRI) into its three constituent parts: income tax, value-added tax (VAT), and excise and other minor taxes. The tax system would probably become more equitable if the share of income taxes in the total were to rise.

The indicator should reflect only the unit of measure and not refer to increases. The increases will be evident in the targets.

**Reliability**

The indicator should be based on tax revenues as reported by the SRI.

The tax ratio for all years, including the baseline year, should be calculated using the newly revised (upward) GDP figures as the denominator.

**Directness**

The indicator should be based on tax revenues as reported by the SRI. The tax ratio for all years, including the baseline year, should be recalculated using the newly revised (upwards) GDP figures as the denominator.

**Adequacy**

USAID should obtain data for this indicator on a quarterly basis.

**F. SO 1: Biodiversity Conservation**

**1. Increased financing of local partners by outside (non-USAID) sources**, referring to the Intermediate Result:  
*Strengthened capacity of targeted NGOs and CSOs active in biodiversity conservation*

The DQA Team has made the following recommendations concerning this indicator.

**Validity**

Given the limitations of this indicator, and the confusion surrounding its meaning, the consultants recommend that USAID and its partners jointly develop a set of criteria that would provide an organized, qualitative/quantitative measure of institutional strengthening. A pre-established minimum improvement in this indicator, expressed perhaps in percentage terms, would be required in any year to claim that an organization has been strengthened. Chemonics's indicators for adequate and improved management, more directly related to the BIO 1.2(b) indicator (natural resource management practices), could be used as a model. TNC's standard scheme for measuring institutional capacity would also provide useful inputs.

**Timeliness**

USAID should seek to obtain data from TNC and WWF on a semi-annual basis.

**2. Hectares of land in selected protected areas and buffer zones under participatory NRM plans, measuring the**

Intermediate Result:

*Economically viable natural resource management (NRM) practices adopted, in selected protected areas and their buffer zones*

Recommendations regarding this indicator are given below.

**Validity**

Representatives of two USAID partners questioned the validity of claiming that NRM plans could adequately protect large tracts of land. One interviewee pointed out that good protection on 100 hectares means little if the full 10,000 hectares under the plan cannot be adequately protected. He suggested that NRM plans concentrate on watersheds or micro-watersheds rather than larger areas. As USAID considers revising its indicators following the transition year 2003 for this SO, the DQA Team recommends that it "think small" in reformulating this indicator and establishing targets for it.

USAID partners should make greater efforts to ensure that no double-counting occurs of land under several different types of NRM plans.

USAID should clarify what it means by "participation."

The revised indicator should also include a dimension that measures the adequacy of NRM plan implementation. Chemonics's indicator for "adequate" management (based originally on achieving 5 of 9—now 10—milestones) and "improved" management (an annual increase of at least 20% in the management index) could be used as models. One weakness of these indicators, as Chemonics itself recognizes, is that the milestones are given equal weights. Devising a weighting system for these or similar milestones would seem to be feasible in a collaborative effort between USAID and its partners. TNC has a standard scheme for measuring

institutional capacity that is more complex and implicitly assigns weights to a large number of indicators for which qualitative judgments are sought. This scheme should provide useful suggestions for an indicator that could be agreed upon by all parties.

**Reliability**

See second and third recommendations under VALIDITY above.

**Objectivity**

See recommendations under VALIDITY above.

**Timeliness**

USAID should seek to obtain data from TNC and WWF on a semi-annual basis.

**3. Key policies, legal frameworks and enforcement mechanisms prepared, modified and introduced at appropriate government level to improve biodiversity conservation in targeted protected areas and their buffer zones, referring to the Intermediate Result:**

*Key policies and legal frameworks introduced and/or implemented to conserve biodiversity in targeted protected areas and their buffer zones*

The following observations and recommendations are made concerning this indicator.

**Validity**

The indicator refers to different types of documents (policies, legal frameworks, enforcement mechanisms) at different levels of government. Thus it combines things whose nature and relative importance are different. As the Mission considers revising (or replacing) this indicator, as it intends to do as a result of experience during the transition year of the SO (FY 2003), at the very least it should disaggregate the total figure by both type of document and level of government (national and municipal/community). At the national level, it might be best to identify documents the Mission would like to see prepared, submitted, approved, and effectively implemented (according to a pre-established set of criteria). The entire process might well take several years, so that the target for any one year might include only one of two stages for any particular document. The national-level documents reported under a modified indicator should be few in number: only those with the greatest potential impact.

**Timeliness**

USAID should seek to obtain data from TNC and WWF on a semi-annual basis.

**Directness**  
**Objectivity**  
**Adequacy**  
**Usefulness**

See recommendations for VALIDITY above.

**Practicality**

See recommendations for TIMELINESS above.

### III Weighting of Indicator Contribution to SO and IR Performance

#### D. Annual Report Guidance Regarding Weighting

The *FY 2004 Annual Report Guidance* (September 30, 2003) states:

Rather than use the mission's self-assessment of "failed to meet," "met," or "exceeded," you will be submitting two pieces of information from which Washington will assess SO-level performance... First, you must include the actual indicators against which you decided to evaluate your SO performance. In addition, you will be asked to state the relative importance of your indicators according to their relevance (*sic*) the operating unit places on them toward measuring the progress of the SO on a scale of (1) to (5), with a five being the highest score and a one being the lowest... The point is that we hope to establish some measure of data validity: how closely do the indicators measure what they are supposed to measure... Some indicators may be ranked equally, but all should not be.

The guidance goes on to suggest various criteria for weighting, including the "age" of the indicator relative to the "life" of the SO, how direct the indicator measures the intended result, and characteristics of good indicators included in ADS 203.3.4.2. Otherwise the weighting rationale is left to each mission.

USAID/Ecuador has asked its DQA consultants to take the opportunity of their analysis of the quality of its indicators to recommend weights for those that will be reported in the 2004 Annual Report.

#### E. Weighting Criteria

##### 1. Measurement of SO-Level Performance

The *Guidance* asks the Mission to rank its Annual Report indicators according to how well they measure SO-level performance. How accurately does an indicator, by itself, reflect progress against the Strategic Objective? It is important to note that several SOs in USAID/Ecuador's Results Framework are somewhat general and at a very high level, while many indicators are at a relatively low level.

##### 2. Age of the Indicator

The *Guidance* suggests age as a criterion. Several of the indicators chosen by USAID/Ecuador for its Annual Report suffer the weakness that they were accomplished early in the life of the SO and afterwards do not measure progress.

## F. Recommended Indicator Weights

Weighting Recommendations		
Indicator	Weight	Comments/Rationale
<b>SO 11: Southern Border Integration:</b>		
1. Number of beneficiaries in the target region whose lives are improved by participation access to social services, adoption of improved NRM practices, and/or more effective local governments	2	<ul style="list-style-type: none"> <li>➤ Although the indicator is direct, it is a very distant measure of the SO. However the SO – <i>improved social and economic conditions</i> - is at a high level.</li> <li>➤ The validity of the indicator is diminished by the fact that it counts beneficiaries of land titling and direct hygienic public service equally and adds them together.</li> </ul>
2. Number of new/improved potable water systems	3	<ul style="list-style-type: none"> <li>➤ This is a very direct, valid, objective and useful indicator of its intended IR but a more distant measure of SO-level performance. It should be noted that the SO is transitional in nature and does not look for longer-range impact.</li> </ul>
3. Number of new/improved sanitation units	3	<ul style="list-style-type: none"> <li>➤ This is a very direct, valid, objective and useful indicator of its intended IR but a more distant measure of SO-level performance. It should be noted that the SO is transitional in nature and does not look for longer-range impact.</li> </ul>
4. Number of hectares of land under legal title and/or near buffer zone	1	<ul style="list-style-type: none"> <li>➤ The indicator is an important milestone measure towards lands under Integrated Resource Management Plans, and thus a measure of progress against the IR – <i>improved natural resource management in selected areas along the border</i>. However, it is very distant from, if not wholly unrelated to, measuring achievement against the SO.</li> </ul>
5. Number of local governments that have improved service delivery management	3	<ul style="list-style-type: none"> <li>➤ The indicator is an excellent measure of its intended IR. However, it is distant from and not so direct a measure of SO-level performance.</li> </ul>

Weighting Recommendations		
Indicator	Weight	Comments/Rationale
<b>SO 12: Democracy:</b>		
1. Progress in meeting selected benchmarks in implementation of new CPC achieved	1	<ul style="list-style-type: none"> <li>➤ The indicator is only a fairly direct and valid measure of IR 12.1.1 (a sub-IR) and hardly a direct and valuable measure of progress of the SO at all.</li> <li>➤ As observed in the DQA, the meaningfulness of this indicator is uncertain since it measures milestones in a process the end of which (or the distance) is unknown.</li> </ul>
2. Percent of justice personnel sub-groups given advanced training	2	<ul style="list-style-type: none"> <li>➤ The indicator is a somewhat direct and valid measure of IR 12.1.1 (a sub-IR) but not a very direct and valuable measure of progress of the SO.</li> <li>➤ It is marginally a valuable measure of the relevant IR.</li> </ul>
3. Number of USAID-aided municipalities implementing or improving at least one service chosen with broad citizen input, including marginalized	1	<ul style="list-style-type: none"> <li>➤ This is a sub-sub-IR indicator, really measuring progress at the activity level. It is extremely distant from and has almost no relationship with progress at the SO-level.</li> </ul>
4. Number of USAID-supported municipalities where citizen groups are overseeing targeted services	1	<ul style="list-style-type: none"> <li>➤ This is a sub-sub-sub-IR indicator, really measuring progress at the activity level. It is extremely distant from and has almost no relationship with progress at the SO-level.</li> <li>➤ From site visits made during the DQA, it appears that the indicator may be a fairly good measure for one particular aspect of the IR – <i>greater inclusiveness of democratic processes.</i> (Otherwise it might be ranked zero.)</li> </ul>
5. Justice Coalition is (a) established and (b) effectively lobbying for Justice Reform	1	<ul style="list-style-type: none"> <li>➤ As in the case above, this is a sub-sub-sub-IR indicator, really measuring progress at the activity level. It is extremely distant from and has almost no relationship with progress at the SO-level.</li> <li>➤ Moreover, it is “old” in that the “establishment” and “functioning” of the Coalition was achieved in 2003 and yet the activity runs through 2005.</li> </ul>
6. ARD-defined interventions to address key policy bottlenecks	1	<ul style="list-style-type: none"> <li>➤ The indicator purports to measure progress at the sub-sub-IR level only and it does it poorly.</li> <li>➤ It does not seem very directly related to achievement at the IR-level, let alone the SO-level.</li> <li>➤ It is “old” in that it was achieved early in the activity and has no relevance in the later years.</li> </ul>

Weighting Recommendations		
Indicator	Weight	Comments/Rationale
<b>SO 13: Northern Border Development:</b>		
1. Number of village potable water systems constructed/repaired/expanded in target areas	3	<ul style="list-style-type: none"> <li>➤ The indicator does not focus on number of beneficiaries, although this figure is known and included in the SO-level indicator.</li> <li>➤ It does not directly address the IR (improved health conditions), because impact on health is not measured; but improved water systems, once operational, do contribute to improved quality of life, which is the SO.</li> </ul>
2. Number of water boards created, legalized, trained and functioning, by category A and B	1	<ul style="list-style-type: none"> <li>➤ "Functioning" does not adequately describe what the water boards are expected to accomplish.</li> <li>➤ The indicator does not directly measure the IR (improved health conditions). Its impact on the SO is indirect and probably small.</li> </ul>
3. Number of Water Administration Units in municipalities created and/or strengthened	1	<ul style="list-style-type: none"> <li>➤ What constitutes an acceptable level of "strengthening" is unclear.</li> <li>➤ The indicator does not directly measure the IR (improved health conditions). Its impact on the SO is indirect and probably small.</li> </ul>
4. Number of irrigation systems repaired and number of hectares of land irrigated	2	<ul style="list-style-type: none"> <li>➤ The indicator closely measures the result as stated but gives no idea of how the efficiency of irrigation systems has been improved, or what maintenance and other preventive actions are needed to maintain efficiency.</li> <li>➤ Properly functioning irrigation systems should affect quality of life (the SO) indirectly by raising incomes, but this impact is not being measured.</li> </ul>
5. Number of students, teachers, and parents participating in drug education/prevention program	2	<ul style="list-style-type: none"> <li>➤ The indicator is directly related to its intended IR but does not focus on results.</li> <li>➤ An effective program should affect quality of life indirectly by reducing family violence, theft, and other criminal behavior, and by improving health.</li> </ul>

Weighting Recommendations		
Indicator	Weight	Comments/Rationale
<b>SO 14: Poverty Reduction:</b>		
1. Increased number of microfinance credit clients in selected microfinance institutions	2	<ul style="list-style-type: none"> <li>➤ This indicator is directly related to the IR but provides no direct evidence of changes in beneficiaries' incomes or other aspects of well-being.</li> <li>➤ Greater access to sustainable microfinance institutions (MFIs) will increase opportunities for the poor, but not all MFI borrowers will be able to take effective advantage of these opportunities. The overall impact on poverty reduction (the SO) is likely to be relatively small.</li> </ul>
2. Appropriate microfinance, regulatory and supervisory mechanisms established and/or implemented	1	<ul style="list-style-type: none"> <li>➤ The indicator provides no direct evidence regarding the appropriateness or likely effectiveness of these mechanisms.</li> <li>➤ The planned results do not seek to measure any progress in implementation, and the relationship with the SO is distant.</li> </ul>
3. Financial sustainability of microfinance institutions increased to 100%	4	<ul style="list-style-type: none"> <li>➤ This indicator is based on an internationally accepted measure of financial sustainability, a key goal for such programs.</li> <li>➤ Financial sustainability is essential for long-run expansion of MFIs, which will allow more people to have access to microfinance. Thus it has a logical, long-term relationship with the SO.</li> </ul>
4. Increased Tax Revenue collections as percentage of annual GDP	2	<ul style="list-style-type: none"> <li>➤ Tax revenue increases, by improving the fiscal situation, have a potentially major indirect impact on poverty reduction through their positive effects on fiscal stability and thus economic growth. But increased revenues do not guarantee greater delivery of social services to the poor.</li> <li>➤ Disaggregating the tax/GDP ratio would provide an indication of whether the tax structure is becoming more equitable (an IR concern).</li> </ul>

Weighting Recommendations		
Indicator	Weight	Comments/Rationale
<b>SO 1: Biodiversity Conservation:</b>		
1. Increased financing of local partners by outside (non-USAID) sources	1	<ul style="list-style-type: none"> <li>➤ While strengthened institutional capacity (the IR) should help conserve biodiversity (the SO) the current indicator is a weak measure of institutional capacity. USAID's partners have been confused about its meaning.</li> </ul>
2. Hectares of land in selected protected areas and buffer zones under participatory NRM plans	2	<ul style="list-style-type: none"> <li>➤ This indicator, which in principle is closely related to the SO, could be given a higher weight if reformulated to concentrate on NRM plans for small areas and to focus more on improved implementation. Like the financing indicator, it too has created some confusion among USAID's partners.</li> </ul>
3. Key policies, legal frameworks and enforcement mechanisms prepared, modified and introduced at appropriate government level to improve biodiversity conservation in targeted protected areas and their buffer zones	1	<ul style="list-style-type: none"> <li>➤ Adding different types of documents, prepared for different levels of government (with different potential scales of impact), results in a weak indicator.</li> <li>➤ Furthermore, the indicator does not include the approval and implementation stages, thus making it distant from achievement of the SO.</li> <li>➤ The Mission is considering dropping it or replacing it with a different type of indicator, but it could be reformulated.</li> </ul>

## IV Revision of the Performance Monitoring Plan

### A. Introduction

The Mission is aware that it will probably be necessary to revise the PMP for at least some of its SOs. Some are being extended. Some are ending and may be extended. There are new, long-term activities for Democracy and Northern Border.

The Mission has asked the DQA analysts, taking advantage of the data validation exercise focused on many of the Mission's current indicators, to provide observations and recommendations that would help direct the PMP revision exercise. Moreover, even though the DQA Team has analyzed the quality of selected USAID/Ecuador indicators, it has, within that process, made certain findings regarding the quality of some of the Mission's Intermediate Results and even Sub-IRs as well.

### B. SpO 11: Southern Border Integration

#### 1. Background

Activities under the Special Objective began before the SpO itself was formulated. In October 1998 the presidents of Ecuador and Peru signed a Peace Agreement, ending the latest phase in an historic border dispute. As part of that agreement the countries formulated and are now implementing a Bi-national Border Development Plan. In support of the plan, USAID contributed \$1.5 million in border activities in FY 1999, increasing its support to \$20 million over the period FY 2000-2005. The Special Objective itself – ***Improved social and economic conditions of inhabitants along the Peru-Ecuador border, thereby promoting border integration*** – was approved in September 2000 and Programa Sur (PSUR), or Southern Border Integration activities, began.

The Special Objective recognized the immediate need to improve living conditions of border inhabitants. Thus it focused on infrastructure of public services and the ability of local groups and municipal governments to deliver them. PSUR will be coming to an end in 2004 after five years' duration.

If the Southern Border program will continue in some way and the objective is to be reformulated, one of the principal issues the Mission needs to consider is: Will the program now be more developmental than transitional? Several implications for the content of the program – its intended results and activities – depend on the answer to this question.

## 2. Recommendations Regarding Revision of the Southern Border Objective

The Mission needs to set time aside and consider: What does the Mission want to accomplish now, given what it has already accomplished in the South and the changes that have taken place in the last five years?

- The Mission should carry out a strategic planning exercise within the SO Team and related staff:
  - Decide if the program has become more developmental than transitional:
    - Does it have longer-term objectives than it did before?
    - Is it still a Special Objective or a Strategic Objective?
  - Validate (or reformulate) the wording of the Special (Strategic) Objective itself.
  - Based on the above, validate the Intermediate Results and the Results Framework that holds them together:
    - Are results still related to the Bi-national Development Plan?
    - How has that progressed and how might its progress reflect on the new SpO or SO?
    - What results are other donors trying to achieve that complement what the Mission is doing?
- Then the Mission should hold a two- or three-day facilitated PMP Revision Workshop with the SO Team and Implementing Partners with the following purposes:
  - Based on the in-house strategic planning exercise,
    - Present and explain the reformulated Strategic Objective, Results Framework and Intermediate Results to the IPs;
    - Define the performance indicators together;
  - Reach an agreement with the IPs on the above; and
  - Re-write the PMP for the Strategic Objective.

### C. SO 12: Democracy

#### 1. Findings

During the Data Quality Assessment exercise, one finding about the SO and Results Framework for Democracy stood out: it is formulated at too high a level for the Mission to appreciably affect, within its manageable interests.

- The SO itself is at a very high level, amorphous, and difficult to understand concretely.
- The IRs are formulated at levels high enough to be strategic objectives themselves.

However, the Mission may be saddled with the SO, as it is difficult to revise an SO once approved.

## 2. Recommendations Regarding Revision of the Democracy Results Framework

If the Mission decides it does not wish to revise the SO and bring it down to a manageable level, then the DQA Team can recommend the following steps:

- Bring down the IRs to a lower level and focus them on the justice system and municipal governments.
- Narrow the focus of IR 12.2 as per the DQA Checklist observations and recommendations.
- Clarify new and/or continuing municipal partners given the geographic shift in that activity.
- Refine the Sub-IRs. (Does the Mission really want all the results that currently appear in the Results Framework? And what is the Mission doing to cause them?)
- Refine indicators at the SO and, particularly, the IR levels.
- Report on some of these higher-level indicators in next year's Annual Report.

Methodologically the Mission could follow the same steps described for Southern Border: In-house strategic planning followed by a facilitated PMP design workshop with the IPs.

## D. SO 13: Northern Border Development

### 1. Background

The Northern Border program was initially approved in October 2000 as a two-year effort funded at a level of \$8 million in Plan Colombia funding. Since FY 2002 it has been an integral part of the Andean Regional Initiative, with funding to-date of more than \$25 million. It is designed to contain the spread of the Andean coca/cocaine economy into Ecuador, in particular from across the Colombian border to the north. It seeks to provide rapid improvements in the quality of life of local residents, in large part through the construction/repair/expansion of basic services such as water and sanitary services. A February 2003 amendment to the program extended it to September 2006. The framework of the SO was reformulated as part of this process.

The current program seeks to (1) increase citizen satisfaction with the performance of local-government institutions; (2) expand the scope for licit income and employment opportunities; and (3) undertake drug education/prevention activities. Given the extension of the Northern Border

program to September 2006, its conversion from a Special Objective to a Strategic Objective, and the much broader range of programmatic activities beginning in FY 2004, a new PMP needs to be developed as a framework for reporting next year's results.

## 2. Findings

A member of the DQA Team visited program activities in Sucumbios Province in the Oriente region, one of the three northern provinces in the program (the Others are Carchi, predominantly in the Sierra, and Esmeraldas on the Coast). His impressions of three water projects and of a meeting with a group of 15 students participating in drug education/prevention activities were favorable.

An important concern, however, is that none of the indicators is designed to measure the impact of program activities on households' quality of life, which is the Mission's SO. Likewise, the three indicators related to water projects do not measure improved health, the dimension of quality of life that serves as the IR for these indicators.

## 3. Recommendations

Among the DQA Team's recommendations are the following:

- Instances of double-counting of persons benefiting from more than one intervention related to basically the same activity (e.g. the water-systems example identified in Section II) should be eliminated in the SO-level indicator.
- The DQA Team recommends revisions in each of the indicators related to potable water systems to provide greater clarity. Additional indicators that might be considered would measure water use, reliability of service, and on-time payment of bills.
- The Mission has been considering adding indicators on local government development and the expansion of licit economic activities. The team supports the addition of these indicators, as they are important for measuring performance under the new framework for the SO.
- Separate, 2-3 day participatory workshops with partners and residents of each of the three provinces would be helpful in formulating plans for stimulating licit economic activity and developing an appropriate indicator for measuring its expansion.

## E. SO 14: POVERTY REDUCTION

### 1. Background

This strategic objective is addressed through two programs: strengthening microfinance institutions (MFIs) to provide the poor greater access to credit and savings instruments, and improving the macroeconomic and institutional environment for more equitable economic growth. The key activity in the latter program is strengthening the Servicio de Rentas Internas (SRI) so that tax evasion and tax avoidance can be reduced and the tax revenue/GDP ratio increased.

The microenterprise program was motivated by the increase in the incidence of poverty in Ecuador in the recent past, and the related increase in the percentage of Ecuadorians working in the so-called informal sector of the economy. Improved access to microfinance (savings as well as credit) is expected to help reduce the incidence of poverty.

Activities under SO 14 are planned to continue through September 2006.

## 2. Findings

Among the DQA Team's findings are the following:

- Some of the key indicators in the microfinance program, notably financial stability and the at-risk portfolio, are showing good results, and the number of microfinance clients is growing significantly. None of these indicators, however, provides any direct evidence of how borrowers' incomes have changed as a result of their improved access to credit.
- Both DAI and WOCCU are using internationally accepted definitions of financial sustainability, a key indicator for determining the likelihood that MFIs will prosper and grow over the long run. The definitions they use are not identical but seem close enough for the difference not to be of concern.
- The tax revenue/GDP indicator is not a direct measure of poverty reduction. A strong fiscal situation should stimulate more investment and economic growth, thus reducing the incidence of poverty indirectly; but higher tax revenues do not automatically translate into greater social-service delivery, which would benefit the poor more directly.
- Neither the Mission nor DAI has been closely monitoring the tax revenue/GDP indicator.

## 3. Recommendations

The DQA Team regards the current strategic focus and results framework of this SO as appropriate. However, we do recommend that some indicators be clarified and/or amplified, as well as greater USAID involvement in monitoring the program of assistance to the SRI.

- The indicator for the number of microfinance credit clients should be disaggregated into male and female clients. WOCCU can now provide these data quite easily. The Mission should also consider disaggregating the data by size of loan.
- The microenterprise survey that DAI plans to carry out in early 2004 provides a good opportunity to establish a baseline for regular surveys that would seek to measure the impact of microcredit programs on borrowers' incomes.
- Benchmarks, defined as clearly and objectively as possible against pre-established criteria, should be set for implementation of the three regulatory and supervisory mechanisms that are the subject of IR 14.1.1.
- The financial sustainability indicator should be defined to indicate that an index of 100% is only the minimum desirable figure. The overall indicator should be a weighted average, and the number of MFIs not achieving the target should be shown in relation to the total (e.g. 4/15). These last comments apply also to the at-risk indicator.
- The Mission should monitor more closely the progress made by the SRI in raising the ratio of tax revenues to GDP. At the very least, this figure should be monitored semi-annually. Interpreting the results will require an understanding of seasonal patterns in tax collections as well as of the overall state of the economy, which has a significant effect on the ratio.
- In addition, the tax ratio should be disaggregated into its three component parts: income tax, value added tax, and excise and other minor taxes. If the share of the income tax in the total is rising, it is likely that the tax system is becoming more equitable (a concern of the Mission's IR).

## **F. SO 1: Biodiversity Conservation**

### 1. Background

This SO was designed at a time when the Mission was scheduled to be closed within 2-3 years, so its initial time horizon was relatively short. That time horizon was subsequently lengthened, and the range of activities broadened, after USAID decided to continue its program in the country.

The Mission regarded FY 2003 as a year of transition to a new strategy. In May 2003 it presented an *FY04-FY08 Extension Amendment*, reorienting its strategy to focus on four themes: (1) creation of direct incentives for conservation; (2) adequate management of protected areas or indigenous territories; (3) good local governance of natural resources; and (4) sustainable financing of conservation activities. Details of this strategy, and its proposed results framework (see below), are still being discussed within the Mission.

## 2. Findings

The DQA review of the three SO 1 indicators discussed in the Mission's Annual Report raised a number of concerns, including the following:

- Each of the three indicators suffered from lack of clarity, which resulted in the Mission's partners providing data that were not comparable;
- The indicators were focused relatively little on implementation, possibly in part because FY 2003 was a year of transition to a new strategy;
- Some double-counting of hectares under natural resource management (NRM) plans likely occurred.

## 3. Toward a New Strategy

The proposed Results Framework for the new strategy may be summarized as follows:

**SO 1: Biodiversity conserved in protected areas, their buffer zones, and indigenous territories**

**IR 1: Governmental and non-governmental institutions develop adequate capacity to govern natural resources management, demonstrating transparent, accountable governance**

Indicators:

- Number of these organizations meeting the indices identified in the development literature; and
- Ability of these organizations to acquire non-USAID support.

Comments:

- The first of these indicators moves in the direction of the kind of institutional-strengthening indicator the DQA Team has recommended. But the discussion in the new document suggests that the Mission needs to move farther in this direction, toward a more rigorously formulated indicator (see Part II above).
- The second indicator seems similar to the indicator for the current IR 1, which is fraught with many problems and which we suggest be replaced with something similar to what the Mission has in mind for the indicator discussed immediately above.

## **IR 2: Sustainable natural resource management practices adopted**

### Indicators:

- Hectares of land under participatory natural resource management plans;
- Number of men and women participating in these activities; and
- Number of economic alternatives (to non-sustainable resource use) under implementation.

### Comments:

- The first of these indicators seems similar to or identical to the current indicator for IR 2, which we have recommended be substantially revised to provide greater precision in measuring appropriate management practices.
- The second indicator seems to be of a relatively low order of priority unless one can closely link participation with improvements in well-being.
- The third indicator appears on the surface to lump together a variety of activities without indicating their relative importance or their likely economic impact. It does not seem promising.

## **IR 3: Sustainable financing mechanisms contribute to conservation efforts**

### Indicator:

- Income generated or dollars saved through economic activities that conserve biodiversity

Illustrative sources of income include those from capitalization of a water fund; unspecified long-term mechanisms for sustainable finance of conservation in indigenous territories; ecotourism on Isabela Island (Galapagos) as an alternative to fishing; a larger endowment fund for the Galapagos Marine Reserve; and pricing schemes for selling ecological services (presumably including carbon sequestration and protection of genetic resources).

### Comments:

- This IR has a number of attractions. The Mission may have to go through a period of trial and error to establish appropriate monetary targets.
- For the SO as a whole, we repeat the recommendation made for SpO 11:

- Hold a two- or three-day facilitated PMP Revision Workshop with the SO Team and Implementing Partners with the following purposes:
  - Based on the strategic planning, validate the performance indicators;
  - Reach an agreement with its IPs on the above; and
  - Re-write the PMP for the Objective.