IGES Policy Report, No. 2012-06

Monitoring & Evaluation of Education for Sustainable Development

A FRAMEWORK OF THE MAIN FACTORS AND IMPORTANT LEVERAGE POINTS IN THE IMPLEMENTATION OF EDUCATION FOR SUSTAINABLE DEVELOPMENT IN THE ASIA-PACIFIC REGION

Produced as part of a joint research project by UNU-IAS and IGES to develop Indicators of Education for Sustainable Development to conduct monitoring and evaluation of its implementation in the Asia-Pacific Region

UNITED NATIONS UNIVERSITY
UNU-IAS
Institute of Advanced Studies

IGES
Institute for Global Environmental Strategies

IGES Policy Report, No. 2012-06

Monitoring & Evaluation of Education for Sustainable Development

A FRAMEWORK OF THE MAIN FACTORS AND IMPORTANT LEVERAGE POINTS IN THE IMPLEMENTATION OF EDUCATION FOR SUSTAINABLE DEVELOPMENT IN THE ASIA-PACIFIC REGION

Authors: Robert J. Didham and Paul Ofei-Manu

Additional Research Team Members:
Abel Barasa Atiti, Sachiko Yasuda, Mario Tabucanon, Katie Vanhala, Akira Ogihara

Project Consultants and Reviewers:

Hironori Ozawa, Asia-Pacific Cutural Centre for UNESCO
Vinai Veerawatnanond, The Association for Environmental Education of Thailand
Fumiko Noguchi, ESD-J
Sun-Kyung Lee, Cheongju National Univ. of Education
Athapol Anunthavorasakul, Chulalongkorn University
Mark Elder, IGES
Hironori Hamanaka, IGES
Hidefumi Imura, IGES
Weena Namcharoensombut, International Sustainable Development Studies Institute
Jeanine Schreurs, Maastricht University
Benjalug Namfa, Ministry of Education – Thailand
Wongduan Suwansiri, Ministry of Education – Thailand
Zainal Abidin Bin Sanusi, Ministry of Higher Education – Malaysia
Savitree Srisuk, Ministry of Natural Resources and Environment – Thailand
Monthip Tabucanon, Office of the Parliament – Thailand
Tinsiri Siribodhi, Southeast Asian Ministers of Education Organization Secretariat
Robert Steele, Systainability Asia – AtKisson Group
Masahisa Sato, Tokyo City University
Masaru Yarime, Tokyo University
Ampai Harakunarak, UNEP - Regional Office for Asia and the Pacific
Mahesh Pradhan, UNEP - Division of Environmental Policy Implementation
Justin Alick, UNESCO - Asia Pacific Regional Bureau
Jennie Ekedahl, UNESCO - Asia Pacific Regional Bureau
Danny Padilla, UNESCO - Asia Pacific Regional Bureau
Aurelie Acoca, UNESCO Institute of Statistics
Ngamnet Aektasaeng, UNESCO Institute of Statistics
Nyi Nyi Thaung, UNESCO Institute of Statistics
Kazuhiko Takemoto, UNU-IAS
Manoon Manukhachorn, Valaya Alongkorn Rajabhat University
Darwina Sri Widjajanti, Yayasan Pembangunan Berkelanjutan
Soon-Yong Pak, Yonsei University

Institute for Global Environmental Strategies (IGES)

2108-11 Kamiyamaguchi, Hayama, Kanagawa, 240-0115, Japan
Tel: +81-46-855-3700 Fax: +81-46-855-3709
E-mail: iges@iges.or.jp
URL: http//www.iges.or.jp

United Nations University Institute of Advanced Studies (UNU-IAS)

6F International Organizations Center, Pacifico-Yokohama
1-1-1 Minato Mirai, Nichi-ku, Yokohama, 220-8502
Tel: +81-45-221-2300 Fax: +81-45-221-2302
E-mail: unuias@ias.unu.edu
Website: www.ias.unu.edu

Monitoring and Evaluation of Education for Sustainable Development: A framework of the main factors and important leverage points in the implementation of education for sustainable development in the Asia-Pacific region

Authors: Robert J. Didham and Paul Ofei-Manu

IGES Policy Report

ISBN: 978-4-88788-118-1

Printed in Japan
Printed on recycled paper

TABLE OF CONTENTS

FOREWORD

The declaration by the United Nations (UN) in 2002 for a Decade of Education for Sustainable Development (DESD), 2005-2014, reflects a global commitment to the role of education and learning in transition to sustainable development. As the Decade comes to its conclusion in 2014, it is important to carry out stock-taking of ESD implementation with a view to developing strategies for the way forward. Documenting achievements, mechanisms and indicators for measuring progress is central to monitoring and evaluation of ESD as envisaged in the UN DESD International Implementation Scheme. However, devising suitable monitoring and evaluation frameworks and relevant indicators to measure ESD progress is a challenging task. This is because ESD processes are complex, dynamic, emergent and contextual in nature.

Many countries in Asia-Pacific are implementing DESD with a view to provide everyone with the opportunity to benefit from education and learning processes that enable societal transition to sustainable development. Through use of appropriate monitoring mechanisms it is possible to assess progress made on the implementation within the region. Although remarkable progress has been made in Asia-Pacific with regard to DESD implementation, no systematic evaluation of ESD implementation across countries in the region has been fully carried out.

We are pleased to note that the United Nations University Institute of Advanced Studies (UNU-IAS) and the Institute for Global Environmental Strategies (IGES) initiated a collaborative evaluation research project to address this gap. The collaborative research project that has been implemented in close collaboration with UNESCO Asia and Pacific Regional Bureau for Education yielded useful insights into the success factors and constraints to DESD implementation. This policy report shares factors and leverage points that contribute to successful ESD performance. Critical reflections on the process, challenges and possibilities for monitoring and evaluation in the region resulted in valuable lessons that can be drawn upon to improve the quality of education and learning for sustainable development in the region.

We are also happy to note that through this joint research a clearer understanding of the variety of indicators, as well as the impacts arising out of ESD efforts in Asia-Pacific region has emerged. The need to develop capacities for monitoring and evaluation of ESD in the region cannot be over-emphasised. The expected launch of a set of ESD indicators and a guidebook on monitoring and evaluation will further strengthen monitoring and evaluation of ESD initiatives in the Asia-Pacific region. It is our ardent hope that countries in the region will set aside resources to continually monitor and evaluate ESD processes using these tools.

It is important to emphasise that undertaking rigorous monitoring and evaluation of a complex intervention such as ESD requires the commitment and active support of many stakeholders. The UNU-IAS and IGES collaborative evaluation research project involved stakeholders from different cultures, contexts and levels of experience in ESD. We sincerely thank all those who participated in the research project in one way or another. This report's contribution to evaluation is distinctive in its attempt to highlight the many possibilities

available for monitoring and evaluation of ESD. The use of case studies from Regional Centres of Expertise (RCEs) to provide qualitative data for the evaluation is one such possibility. We thus, acknowledge the work of RCEs as a useful mechanism for involving multi-stakeholder groups in the monitoring and evaluation of education and learning processes during the DESD.

Through this policy report, monitoring and evaluation of ESD have been used to extend the knowledge of stakeholders. We envisage that recommendations emerging from the report will inform the thinking of policy makers, ESD practitioners and the general readers to strategise and plan for ESD beyond 2014.

Yokohama, Japan
29 March 2013

Akira Ogihara, PhD
Senior Coordinator
Institute for Global Environmental Strategies

Abel Barasa Atiti, PhD
Research Fellow, ESD Programme
United Nations University Institute of Advanced Studies

LIST OF TABLES, FIGURES AND BOXES

Tables:

Figures:

Boxes:

LIST OF ACRONYMS AND ABBREVIATIONS

21CE	21st Century Education
AfL	Alignment of assessment for learning
ARIES	Australian Research Institute in Education for Sustainability
ASEAN	Association of South East Asian Nations
ASPnet	(UNESCO) Associated Schools Project Network
ATC 21	Assessment and Teaching 21st Century Skills
CIA	Central Intelligence Agency
DESD	Decade of Education for Sustainable Development
DFID	Department for International Development, United Kingdom
EFA	Education for All
EIA	Environmental Impact Assessment
ESCAP	Economic and Social Commission for Asia and the Pacific
ESD	Education for Sustainable Development
ESD M&E	Education for Sustainable Development Monitoring and Evaluation
ESDinds	Education for Sustainable Development Indicators, project on values-based indicators for ESD
ESD-J	Japan Council on the UN Decade of Education for Sustainable Development
FA	Formative Assessment
GCE O/A	General Certificate Ordinary/Advanced
GMEF	Global Monitoring Evaluation Framework
GRS	Global Rating Scales
HDI	Human Development Index
IA	International Assessment
ICTs	Information and Communication Technologies
IDTs	International Development Targets
IGES	Institute for Global Environmental Strategies
IIS	International Implementation Scheme
LSA	Large scale assessment
M&E	Monitoring and Evaluation
M&E of ESD	Monitoring and Evaluation of Education for Sustainable Development
MCQ	Multiple choice questions
MDGs	Millennium Development Goals
MEEG	Monitoring and Evaluation Expert Group

MEXT	Ministry of Education, Culture and Sports
NA	National Assessment
NAEP	National Assessment of Educational Progress
NAPLAN	National Assessment Program – Literacy and Numeracy
NGOs	Non-governmental organisations
NRC	National Research Council
OECD	Organization for Economic Cooperation and Development
OECD-CERI	Organisation for Economic Co-operation and Development/Centre for Educational Research and Innovation
OM	Observational Measurement
OPDCA	Observe-plan-do-check-adjust
PBA	Performance-based assessment
PCAP	Pan-Canadian Assessment Programme
PDCA	Plan-do-check-adjust
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PRA	Participatory rural/rapid appraisal
QAA	Quality Assurance Agency
QCA	Qualifications and Curriculum Authority
RCEs	Regional Centres of Expertise
SA	Summative Assessment
SAIP	School Achievement Indicators Programme
SD	Sustainable Development
SIA	Social Impact Assessment
SMART	Specific, measurable, achievable & attainable, realistic & relevant, timely
SPREP	Secretariat of the Pacific Regional Environmental Program
SWOT	Strengths, Weaknesses, Opportunities, and Threats
TA	Traditional Assessment
TES	Traditional Education System
TIMSS	Trends in International Mathematics and Science Study
UE	University of Exeter
UK	United Kingdom
UMRE	Unidad de Medicion de Resultados Educativus
UN	United Nations
UNAIDS	Joint United Nations Programme on HIV/AIDS

UNDESD	United Nations Decade of Education for Sustainable Development (2005-2014)
UNDP	United Nations Development Programme
UNECE	United Nations Economic Community of Europe
UNESCO	United Nations Educational, Scientific & Cultural Organization
UNESCO APRBE	United Nations Educational, Scientific & Cultural Organization, Asia and Pacific Regional Bureau for Education
UNGA	United Nations General Assembly
UNLD	United Nations Literacy Decade
UNU/UNESCO	United Nations University/United Nations Educational, Scientific & Cultural Organization
UNU-IAS	United Nations University Institute of Advanced Studies
US	United States
USA	United States of America
USAID	United States Agency of International Development

Background of Research on
MONITORING & EVALUATION OF EDUCATION FOR SUSTAINABLE DEVELOPMENT

Beginning in July 2011, the United Nations University Institute of Advanced Studies (UNU-IAS) and the Institute for Global Environmental Strategies (IGES) initiated a collaborative research project in close cooperation with UNESCO Asia and Pacific Regional Bureau for Education. This project focuses on the **Monitoring and Evaluation of Education for Sustainable Development** (M&E of ESD) and aims ***to establish regionally-relevant Indicators of ESD*** for assessment of the implementation that has occurred during the United Nations Decade of Education for Sustainable Development (2005-2014) in individual countries across the Asia-Pacific region.

The **overall goal of the research project** is to contribute to the monitoring and evaluation of the implementation of Education for Sustainable Development (ESD) through development of progressive indicators for piloting in the Asia-Pacific region. The **main objectives of the project** include:

1. To develop an ESD monitoring and evaluation framework;
2. To develop an ESD learning performance-good practice case framework;
3. To gather data for National ESD Status Reports;
4. To collect case reports on ESD good practice and learning performance;
5. To undertake data analysis to identify leverage points, success factors and barriers to ESD implementation; and
6. To draft pilot ESD indicators for future application and assessment.

This project was developed with regards to the fact that the UN Decade of Education for Sustainable Development (DESD) will come to a conclusion in 2014. Although there is a large amount of anecdotal evidence about the successes achieved under the DESD framework, there is currently no systematic way to evaluate the implementation of ESD across multiple countries. Furthermore, one of the seven target implementation goals for achievement during the Decade clearly states the need for systems to monitor and evaluate ESD performance. With this in mind, this research project was established to first try to identify the important context, factors and leverage points that commonly lead to successful ESD implementation, along with identifying the strengths and barriers in achieving effective ESD learning performance. Second, the research team aims to launch a set of ESD Indicators for Asia-Pacific along with a guidebook for implementing an effective monitoring and evaluation process. These indicators will ideally be both regionally relevant while also being suitable for application by individual countries in the monitoring and evaluation of their own ESD systems. Thus, the indicators will need to be both replicable (allowing for annual systematic usage) and also provide comparability between different countries' ESD implementation.

The strategy for this research was developed to take account of both the quantitative and qualitative nature of educational monitoring and evaluation. However, this also highlights the conceptual challenge for M&E of ESD, which is that to provide meaningful and timely information to support effective interventions in ESD implementation it is necessary to demonstrate how specific educational inputs will support better ESD learning performance (i.e. increasing the quantity of a input should ideally lead to increased quality of ESD). The priority sectors and focal areas for ESD monitoring and evaluation addressed in this research were identified during an Expert Consultation meeting on ESD monitoring and evaluation held in July 2011 as part of the International Forum for Sustainable Asia and the Pacific (ISAP). It was agreed by the experts at this meeting that the target users of the outcomes from the envisioned monitoring and evaluation work should be national governments and relevant policy makers (especially those from the ministries of education and environment). Six different sectors were identified for investigation during the research; these include: National Curriculum, Formal Education, Teacher Training, Non-Formal Education, Civil Society, and the Private Sector.

Following the Expert Consultation held at ISAP 2011, an evaluation framework for identifying the target areas of ESD assessment was developed. This framework was then used to prepare a country ESD survey, and a further reporting format was developed to collect good practice cases on ESD in a systematic manner. Having received the agreement and support of our partner institutes, we then proceeded to initiate the country research and data collection phase of this project. The research utilised two distinct but complementary approaches. First, national ESD focal points were targeted for participation in a quantitative country survey regarding the national context of ESD implementation. Second, the Regional Centres of Expertise (RCEs) were targeted for qualitative research to provide good practice case studies for a comparative analysis.

This research phase of the project from June 2011 to August 2012 was conducted as a multi-country scoping process to identify the important areas for which indicators should be developed. The main research and data collection process occurred between December 2011 and July 2012 in two rounds, starting first with selected countries in East Asia and then following a refining process moving on to selected countries in Southeast Asia. During the scoping phase, research was conducted across a total of nine countries. Throughout the year long research process, two sub-regional reporting and capacity building workshops on M&E and ESD were held. Additionally, two meetings were also held with the Expert Consultation group to review the process and findings of the research project

The main purpose of this research process is to enable the movement from a wide evaluation framework towards the identification of a core set of important targets and leverage points for ESD. Thus, the scoping research phase was followed by the refinement of the selected ESD leverage points in order to

elaborate a set of regional ESD indicators. These proposed indicators went through a further review from the expert working group before their final drafting.

Four major outputs are expected as the products of this year's research. First, a compilation and comparative evaluation of ESD Country Status Reports will present the current status of ESD implementation in the seven reporting countries. Second, based on the ten good practice cases submitted by the RCEs, these cases are analysed to identify the important criteria for ESD qualitative achievements and develops a learning performance assessment framework for ESD. Third, a theoretical discussion of the process for monitoring and evaluation of ESD is presented and compared with the identification of specific leverage points for ESD implementation from the previous two reports to present an overall framework of the main factors and contents of effective ESD implementation. Finally, the specific ESD Indicators for piloting will be identified and explained in a guidebook for ESD monitoring and evaluation in the Asia-Pacific region.

With continued usage and development of these indicators, it would be possible to provide substantial reporting on the status of ESD across the Asia-Pacific region and to provide a comprehensive report of the achievements made during the UN Decade of Education for Sustainable Development. These indicators should also serve as a valuable tool for individual countries to analyse their own ESD systems and to conduct a strategic needs assessment for planning future interventions for strengthening ESD implementation. Furthermore, a comprehensive study of ESD implementation in the region would also provide policy recommendations about how to continue to improve ESD into future.

The success of this research owes a significant debt to the generous participation of numerous contributors throughout the entirety of this research project (the specific contributors to this report have been noted on the title page). Over twenty-five people contributed directly to the data collection, country status reports and good practice cases. An additional group of fifteen experts provided review and consultation support for the overall research process. The continued support of UNESCO Asia and Pacific Regional Bureau for Education was invaluable throughout this work. The authors of this work and the members of the research team would like to express our deep gratitude to all of these individuals and organisations who have so eagerly cooperated with this research, and who continue to demonstrate a sincere willingness to improve the global implementation of Education for Sustainable Development.

Dr. Robert J. Didham
M&E of ESD Research Team Leader
Education Policy Specialist, Institute for Global Environmental Strategies

SECTION 1

Education for Sustainable Development

&

The Challenges of Comparative Monitoring and Evaluation

"In confronting the many challenges that the future holds in store, humankind sees in education an indispensable asset in its attempt to attain the ideals of peace, freedom and social justice." – Jacques Delors[1]

"It is in fact a part of the function of education to help us to escape--not from our own time, for we are bound by that--but from the intellectual and emotional limitations of our own time." – T.S. Eliot[2]

Introduction

Education for Sustainable Development (ESD) refers to the provision of education and learning opportunities to enhance learners' abilities to effectively understand and participate in the pursuit of sustainable development, without which the transition towards sustainability will be difficult to achieve. Recognition of the importance of ESD was brought to the global stage by the 2002 declaration of the United Nations General Assembly calling for 2005 to 2014 to be marked as the Decade of Education for Sustainable Development (DESD). UNESCO was subsequently mandated as the lead agency to manage the implementation of the decade with the overall goal of integrating the principles, values and practices of sustainable development into all facets and aspects of learning/education and to encourage changes in individual behavior, organisational and institutional practices that allow for a more sustainable and just society for all (UNESCO, 2007).

Ten years after the DESD declaration and eight years into the Decade, ESD and sustainability in the context of education was one of the few domains at the UN Conference on Sustainable Development (Rio+20) where agreement was easily reached on the need for continued improvements and strengthening of education systems to help meet the achievement of sustainable development. However, monitoring and evaluation (M&E) of ESD which is the final of the seven international implementation goals of the DESD has yet to be achieved at the level of systematic implementation. With only two more years to the completion of DESD and after several years of committing considerable amounts of resources and time across the globe to establish the importance of the decade, many notable achievements have been made but other important opportunities for advancement still remain. However, to fully identify and capitalise on these opportunities, the systematic M&E of ESD implementation across countries and regions remains necessary if strategic assessment of what has been achieved and what should be the focus of future activities is to be conducted.

[1] Chairman of the International Commission on Education for the Twenty-first Century; quoted from his opening words of *Learning: The treasure within* (1996, p. 12). Paris: UNESCO.
[2] Quoted by The New York Times on 5 January 1965, available at: http://www.nytimes.com/books/97/04/20/reviews/eliot-obit.html

Important opportunities for progressive improvement of ESD implementation include formulating appropriate frameworks which clearly define leverage points or points of intervention for effective action; minimising the variation of ESD implementation across countries and regions; finding effective ways of implementing ESD initiatives in order to achieve the desired learning performance outcomes; putting in place effective system(s) to monitor and evaluate ESD implementation; and trying to find answers to questions regarding the extent to which DESD has been able to help society transition to sustainability, particularly beyond 2014 and in line with proposals for a new post-2015 development agenda. It is also appropriate to review the mode/system(s) of assessment of current educational processes and content in order identify the aspects that can still prove beneficial in the shift towards sustainability. Furthermore, it is crucial to identify the unique aspects of ESD that support qualitative enhancement and reform to education systems generally.

The efforts of the UN Decade of Education for Sustainable Development (DESD) have supported both the integration of education for sustainable development (ESD) into existing educational systems and the reform of these systems to better address capacity building so learners are more able to engage in the debates about our societies' development trajectories as active contributors to building sustainable solutions. The outcome document of the recent Rio+20 – UN Conference on Sustainable Development, *The Future We Want* (2012), reaffirms the support by member states to advance the practice of ESD, "We resolve to promote education for sustainable development and to integrate sustainable development more actively into education beyond the United Nations Decade of Education for Sustainable Development" (para. 233). Considering how ESD can be more actively integrated into educational systems, it is necessary to first take stock of the achievements already made during DESD while also identifying the major obstacles and barriers for effective ESD implementation. In order to properly conduct this type of assessment, it is necessary to establish appropriate monitoring and evaluation methods for the systematic review of ESD implementation across various countries.

Purpose and Objectives of Scoping Research on M&E of ESD

The development of an effective monitoring and evaluation system for ESD requires the methodological identification of the important context, factors and leverage points that commonly lead to successful ESD implementation. Against that backdrop, the United Nations University Institute of Advanced Studies (UNU-IAS) and the Institute for Global Environmental Strategies (IGES) undertook a collaborative research project in close cooperation with UNESCO Asia and Pacific Regional Bureau for Education over the past eighteen months to conduct scoping research on monitoring and evaluation of Education for Sustainable Development (ESD) in the Asia-Pacific region.

During the scoping phase between November 2011 and April 2012, research was conducted in a total of nine countries in East and Southeast Asia.

This project focuses on the monitoring and evaluation of Education for Sustainable Development (ESD) and aims to establish regionally-relevant Indicators of ESD to assess the implementation during the United Nations Decade of Education for Sustainable Development (2005-2014) in countries across the Asia-Pacific region. Between November 2011 and April 2012, scoping research was conducted in two rounds starting first with selected countries in Northeast Asia and then following refinements to the evaluation framework moving on to selected countries in Southeast Asia. During this scoping phase, research was conducted in a total of nine countries based on an evaluation framework that was developed during the consultation with international ESD experts. The main purpose of this research phase was to enable the movement from a wide evaluation framework towards the identification of a core set of important targets and leverage points for ESD.

The early concept and proposal for this research project was initiated through a series of consultation discussions with UNU-IAS, IGES and UNESCO Asia and Pacific Regional Bureau for Education. An investigation of relevant strategies for implementing ESD was conducted and followed by the development of an evaluation framework for analysing the important factors and capacities necessary for implementing effective ESD. The development of this framework was supported by a consultation process with international ESD experts. Based on the evaluation framework, a research survey was prepared for reporting on the important elements of ESD implementation status at a country level. Then, a questionnaire was developed for data collection from the RCEs in a case-study report format in order to collect good practice cases on ESD learning performance.

A selection criteria was also developed to identify potential countries for participation. This was a simple criteria to ensure that the country would be able to provide relevant and useful information on their ESD implementation and that the selected countries would provide for good comparability. The criteria was based on three factors: first, the geographic location was narrowed to countries from either the East or Southeast Asia sub-regions; second, the countries were required to have recent active involvement in the monitoring and evaluation of ESD process that has been led by UNESCO's regional bureau; and third, the countries should each have at least one RCE (to enable the corresponding qualitative study of good practice cases). These criteria led to the identification of nine appropriate countries. There were three countries from East Asia: China, Japan and the Republic of Korea. There were six countries from Southeast Asia: Cambodia, Indonesia, Malaysia, the Philippines, Thailand, and Vietnam. Partners from all nine countries participated in the good practice

studies, but unfortunately it was only able to gain the cooperation of partners in seven countries for completion of the country ESD status surveys.

The main research was conducted in two complimentary formats to achieve both qualitative and quantitative findings. The quantitative aspect of the research involved national ESD focal points participation in a quantitative country survey regarding the national context of ESD implementation.[3] The qualitative investigation of learning performance was based on the comparative evaluation of good practice cases.[4] While the qualitative study provided for greater examination of the influential factors in achieving effective ESD outcomes (i.e. learning achievements), the quantitative study aided in identifying the important leverage points for ESD implementation (or inputs). The findings from both studies were then triangulated during a further investigation of the important factors and components for a holistic M&E of ESD framework, and this was further strengthened through a third research format based on multi-stakeholder participation and cooperative inquiry. Throughout the year and half research period, three expert consultations and two reporting and capacity building workshops were held. In total, around fifty different people participated in these five events which served as extremely valuable opportunities for collective testing and application of the findings, proposals and recommendations being generated during this research process.

To complement this research approach, the methodology employed was a mixed-methods research strategy. Mixed-methods enhances construct validity and methodological triangulation in order to substantiate research findings. The research design was guided by grounded theory and the application of selective coding. This is an appropriate approach when research is not based on set hypothesis testing, and instead can be used to identify primary factors of influence. For the quantitative data collection, a capacity analysis was used during the comparative country assessment. For the qualitative data from the RCE case studies collected in the learning performance assessment report, data is assessed through theoretical sampling and analytical induction.

The M&E process should also serve as a capacity building mechanism for implementing the DESD goals by engaging people in a process of learning to change regarding reorienting their own behaviors and practices as well as for wider education systems so that these can make significant contributions to sustainable development (Tilbury, 2010). The successful development of a monitoring and evaluation framework will therefore be used for both the assessment and

[3] For full details of the *ESD Country Status Reports* (2012), the research approach, and the main findings, see: http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=4140

[4] For the full details of the *Assessment of Learning Performance in ESD* (2012), the research approach, and the main findings, see: http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=4172

reorientation of programmes during the remaining course of the Decade and beyond, in order to ensure on-going relevance and effectiveness (UNESCO, 2005a). Monitoring and evaluation is hence the foremost strategy to ascertain the changes and impact of the Decade.

Structure of Report

The initial phase of this research project on M&E of ESD was established with the aim to conduct scoping research to enable the movement from a wide evaluation framework towards the identification of a core set of important targets and leverage points for ESD. In this final report based on the initial scoping research, the main goal is to explain how effective M&E of ESD in the Asia-Pacific region could be established and conducted. While the previous two reports individually present the findings from the quantitative ESD status reports and the qualitative learning performance cases respectively, in this report efforts are made to identify some level of correlative links between system inputs, throughputs and outputs for ESD, and specifically to demonstrate how the two aspects can be concurrently addressed in the same M&E approach.

Before specifically discussing the research approaches used during this project and the recommendations being made for future M&E of ESD in the Asia-Pacific region though, it is necessary to provide some background and context to ESD, previous M&E approaches, educational assessment, and the insights these provide for developing an effective M&E of ESD system. To this end, in the rest of section one the focus is on the nature of ESD, the qualities of ESD that make it challenging from the standpoint of assessment, previous attempts to conduct M&E of ESD, and the main reasons for attempting to measure the progress that has been achieved during DESD. Section two begins by providing an overview of the general applications of assessment methods. A lengthy discussion then reviews modern trends in both educational practice and assessment with the purpose of identifying the benefits and limitations of the approaches especially as they could apply to ESD. In the final part of section two, the focus is specifically on investigating approaches for ESD assessment; this first addresses several of the main questions, criteria and targets that must be considered in identifying an appropriate M&E approach, and second it provides a discussion of several potential methods for conducting M&E of ESD.

The third section begins by providing an explanation of the structure and methodology used during this project to conduct scoping research on ESD monitoring and evaluation in the region. The three main research approaches are introduced in turn with a clear explanation of the processes undertaken to conduct these approaches and an explanation of the findings and contributions made by approach. The final parts of this section turn towards synthesising and integrating the findings

from this mixed-methods research in order to reach the original project aim of identifying a core set of important targets and leverage points for ESD. Proposals and recommendations are then presented for an ESD M&E framework. This is complemented by the presentation of several key outcomes from an expert consultation held at the end of research phase to further define the ESD M&E framework in a participatory manner, which is presented in an addendum to the main text.

Background on Education for Sustainable Development

Box 1: Definition of *Education for Sustainable Development*

(as defined by UNESCO in *Promotion of a Global Partnership for the UN Decade of Education for Sustainable Development: The International Implementation Scheme for the Decade in brief* (2006c: 5))

- It means education that enables people to foresee, face up to and solve the problems that threaten life on our planet.
- It also means education that disseminates the values and principles that are the basis of sustainable development (intergenerational equity, gender parity, social tolerance, poverty reduction, environmental protection and restoration, natural resource conservation, and just and peaceful societies).
- Lastly, it means education that highlights the complexity and interdependence of three spheres, the environment, society – broadly defined to include culture – and the economy.

Importance and Uniqueness of ESD

Recognition for the importance of ESD, considered to be at the core of education and learning towards sustainability, has increased substantially over the past decade. Its uniqueness lies in the expression of an array of concepts, constructs underpinned by theory, and "policy prescripts, practical methods and tools" that link education and learning to the social, economic and ecological dimensions of sustainable development in their continuous dynamic interaction (Lenglet et al., 2010: 93). UNESCO describes ESD as striving to promote sustainable development through its four identified thrusts, namely (1) improving basic education, (2) reorienting existing education programs, (3) developing public awareness and understanding of

Table 1: The Four Thrusts of DESD

Promote and Improve the Quality of Education: The aim will be to refocus lifelong education on the acquisition of knowledge, skills and values needed by citizens to improve their quality of life.	***Reorient the Curricula:*** From pre-school to university, education must be rethought and reformed to be a vehicle of knowledge, thought patterns and values needed to build a sustainable world.
Raise Public Awareness of the Concept of Sustainable Development: Raising awareness will make it possible to develop enlightened, active and responsible citizenship locally, nationally and internationally.	***Educate the Employed:*** Continuing technical and vocational education of directors and workers, particularly those in trade and industry, will be enriched to enable them to adopt sustainable modes of production and consumption.

(UNESCO, 2006c: 5)

sustainability, and (4) training (UNESCO, 2005b: 28-30). All four aspects should address the content, pedagogical and learning processes involved in their implementation. In addition, the effectiveness of existing national/institutional mandates backing their operation, along with the physical and institutional structures, the human resources and logistical inputs all should be taken into account.

Additional "unique" features of ESD include its support of lifelong learning skills and processes, its holistic and interdisciplinary nature (McKeown, 2002), and that it is driven by values and principles of critical inquiry, reflective thinking, systemic thinking and problem-solving. ESD builds on various competencies and forms of collaboration including envisioning, cooperative learning, learning by doing, partnership building and participation in decision making (UNESCO APRBE, 2011: 3). It is multi-methodological, involves different pedagogies, with day to day applicability to personal and professional life (IGES, 2005). Furthermore, various philosophical underpinnings and orientations including both philosophies of education and environmentalism that inform ESD implementation (Babikwa, 2004) are important for understanding the entire sustainability concept (Ofei-Manu and Shimano, 2010). Taking after the 'parent' concept of sustainable development (SD), the breadth and depth of ESD definition makes room for multiple interpretations of ESD philosophy, principles and practices. Consequently, ESD means many different things to different people/constituencies; a fact that is very widespread in the literature and hence exposes ESD to the danger of being interpreted as 'all embracing' or in divergent frameworks. ESD can follow two pedagogical interpretations: "1) ESD as a means to transfer the 'appropriate' sets of knowledge, attitudes, values and behaviour; and 2) ESD as a means to develop people's capacities and opportunities to engage with sustainability issues so that they themselves can determine alternative ways of living" (UNESCO, 2009: 27).

Conceptually and in practice, ESD has evolved from many predecessors that have been promoted worldwide. In fact, it is important to recognise that as ESD has drawn many of its features from relevant aspects of pre-existing educational theories, approaches and pedagogies, one of the main unique characteristics of ESD is not its creation of new concepts but rather its ability to establish a holistic framework for integrating and applying these various perspectives in a comprehensive manner. ESD therefore does not necessarily represent specific pedagogies as expounded by several educational theorists, though it embraces many of their basic principles, and similarly it cannot be equated with one particular, codified educational or instructional method/practice (Lenglet et al., 2010). Although associated with several "adjectival education" ideas such as climate change education, disaster risk/preparedness education and community development education, ESD is not considered exclusive to these alone. In addition to being able to manifest itself in different settings, "ESD goes beyond mere socialisation, knowledge transmission, skills learning or awareness-raising"

to invite learners to engage in the realities and challenges with which they are confronted, while transforming the same realities through the context of sustainability (Lenglet et al., 2010: 95).

With strong future orientation regarding content ('curriculum') and process ('pedagogy') in all educational settings of formal, non-formal and informal education, ESD encourages the use of multiple perspectives in addressing sustainability as that also promotes interdisciplinarity and cross-cultural competencies. These distinctive educational/learning approaches and methods can guide the search for appropriate ESD practice by enabling learners to engage with the three dimensions of SD namely a) natural (environmental) capital, b) manufactured (economic) capital and c) socio-cultural and human capital to recreate their reality on a daily basis using these dimensions for their well-being. This is because the ways and methods with which people, individually and collectively understand their reality, come to grips with it and can act on it is central to ESD and against the backdrop that learning does not happen in isolation but rather it is a social act irrespective of the location. In addition to placing importance on local relevance and cultural appropriateness, ESD also recognises that local action and dynamics may have global consequences and vice versa (UNESCO, 2012a; Lenglet et al., 2010; UNESCO 2009; UNESCO, 2005b).

Scales of Relevance for ESD

With regard to scale, Lenglet et al. posit that ESD is relevant for implementation across scales at the local, national, regional, international and at the interplay between these different levels. They contend that at the local level, ESD is practiced because increasing numbers of communities are faced by several challenges including national resources deterioration, climate change, population growth or shrinkage in urban and rural areas respectively, cultural transitions that need to factor in other cultures around the globe, the so-called globalisation, socio-economic transformations driven by increasing inequities and inequalities, unemployment, education systems that fail to incorporate the knowledge and skills that are locally appropriate and culturally relevant, etc. This leaves local communities in the situation where they must "find appropriate balances and trade-offs between cultural, social, economic and environmental necessities, demands, and aspirations" (2010: 93).

At the national level, against the backdrop of using ESD in practice to possibly address the numerous local sustainability issues, national governments have been taking measures to integrate ESD elements into national educational policies and guidelines, curricula and assessments particularly in relation to formal education. For example in Japan, there is a national policy to increase the number of schools joining UNESCOS's Associated Schools Project Network (ASPnet) to 500 by 2014 and this has already been exceeded. Through increased funding and support from central and local

governments for Enviroschools which currently stands at 547, in addition to the review of the New Zealand curriculum and further support for education for sustainability (EfS) teachers, most progress with EfS in New Zealand has been made within primary and secondary schools (EfS, 2012). Through its Green School project, China is supporting ESD adoption through the whole school approach to environmental management and education. In the Philippines, mandate is given through its Agenda 21 documents in promoting sustainable development along with the promotion of environmental education and ESD as important factors to building capacity of the citizenry towards realising a sustainable society. Thailand is using its own unique Sufficiency Economy Philosophy to integrate ESD into the mainstream of Thai education (Didham and Ofei-Manu, 2012).

Regionally, the *Situational Analysis of ESD* is a regional ESD policy document launched by UNESCO Bangkok to obtain a snapshot view of the current state of ESD in the region and assisting in guiding the regional implementation of the DESD. This is just one of several Asia-Pacific ESD initiatives to help create and foster regional partnerships and networks on ESD, to facilitate the exchange of knowledge and experiences, and for future cooperation on regional coordination of ESD (UNESCO APRBE, 2005). Another example is the Association of South East Asian Nations (ASEAN) Environmental Education Action Plan (2008-12) which includes aspects of learning for sustainability. The main objective of the United Nations Economic Commission for Europe's (UNECE) regional Strategy for ESD in Europe was to incorporate major themes of SD in all education systems. Its adoption coincided with the launch of the UNDESD in the region, and it encouraged UNECE countries to "develop indicators to assess its implementation, organize thematic and sub-regional workshops and compile good practices in ESD" (Filho, 2010: 120). Added to that is the formulation of ESD competencies by an international group of experts delegated by the UNECE for educators which made recommendations for policy makers to develop these competencies with respect to the central role of educators to operationalise the ESD concept. There is also the Pacific Education for Sustainable Development Framework developed as a mechanism to assist in the implementation of the Pacific Plan and the basis for coordinating actions to achieve the regional vision for a prosperous future (UNESCO, 2006a).

At the level of international policy development, the UN DESD with implementation led by UNESCO has become an example of how international momentum is being upheld to promote ESD. The Decade has hence become the most important international platform that seeks to embed sustainable development in all learning spheres (whether formal, non-formal or informal education), and where ESD policy and practice are being presented, shared, debated, assessed and further developed (Lenglet et al., 2010). Furthermore, the introduction of the concept of the Regional

Centres of Expertise (RCE) to play a key role in mobilising and facilitating engagement in ESD collaborative initiatives and knowledge and skill production and sharing as well as value promotion at the local level and linking them (RCEs) up across levels with other centres at the international level by using its increasing networks. The RCE community has also begun engaging in international sustainability policy processes so as to showcase its role and ESD-related activities. The interplay between these different levels due to the global nature of current human-environment system challenges and the simultaneous expression in local practice is evident. Learning about and dealing with these challenges require inputs across scales where local knowledge and practice must inform global understanding and action, and *vice versa*. The result is the conglomeration of rich insights, practice and expertise in ESD across local, national, regional and international institutions.

A Brief History of ESD and the Decade

Historically, the ESD concept can be traced to "two distinct areas of interest of the United Nations, namely education and sustainable development" (UNESCO, 2005b: 25). The education aspect is rooted in several international human rights and education declarations, conventions and frameworks for action that were linked to international processes/strategies including Education for All (EFA), the Millennium Development Goals (MDGs), International Development Targets (IDTs), and United Nations Decade of Literacy, (UNLD, 2003-2012). The SD aspect is rooted in the environmental movement which began with the 1972 United Nations Conference on Human Environment in Stockholm, "internationalised" environmental issues and hence resulted in the formation of the United Nations Environmental Programme (UNEP). A search for a broader strategy to address the emerged tension between the environment and human development led to the launch of the Brundtland Commission report *Our Common Future* in 1987 (UNESCO, 2005b: 25-26). After that, the emphasis of the narrative on education in Chapter 36 of Agenda 21 of the United Nations Conference on Environment and Development (Rio de Janeiro, June 1992) was that education is critical to the achievement of sustainable development which served as a watershed for initiating international collaboration and investing the needed human and material capital as a response to calls for ESD implementation. These strategies which are to be implemented by governments, international agencies, businesses and civil society groups sought to improve access and quality of learning for sustainability and reorient education systems to support more sustainable futures (Tilbury, 2010).

Following its appointment by the Commission on Sustainable Development as Task Manager for Chapter 36 of Agenda 21, UNESCO's roles included expediting educational reforms and coordinating

the activities of all stakeholders in education. The objectives of this wide-ranging programme of work included the following: incorporation of education into national strategic and action plans for sustainable development, promoting sustainable consumption and production patterns using education, identification and sharing innovative practices, and most importantly clarifying the concept and main messages of ESD (Tilbury, 2010).

Ten years after the Rio Earth Summit, the World Summit on Sustainable Development was held in Johannesburg in 2002 where a proposal for a Decade of ESD spearheaded by the Japanese and Swedish governments was included in the Johannesburg Plan of Implementation. The Johannesburg Plan of Implementation went on to remind the world of the place of education and learning in the transition towards sustainability. The adoption of the resolution UN GA 57/254 by UN General Assembly in December 2002 resulted in the establishment of the Decade of ESD. The initiation of the UN Decade of ESD (2005-2014) and the emergence of ESD as a global movement (Nomura, 2009) presents a global vision which provides everyone the opportunity to benefit from education and learning that engenders societal transition towards a sustainable future (UNESCO, 2005a). At the 34th UNESCO General Conference in 2007 a resolution on ESD was adopted to encourage considerably more initiatives be taken by Member States and by UNESCO to reorient teaching and learning towards ESD globally. The initial challenge led to further conceptualisation and focus regarding ESD implementation and the strategic prioritisation of actions towards realising visible results particularly during the second half of the Decade (UNESCO, 2012b: 2). The DESD seeks to strengthen and encourage the lifestyles that place value on environmental integrity, economic viability, and a society that is just and peaceful for present and future generations (UNESCO, 2005a).

The main purpose of the UN Decade of Education for Sustainable Development (DESD) which is being pursued by different stakeholders across scales – including UN organsations and programs, national governments, international and national NGOs, groups and individuals (de Haan et al., 2010) – is to integrate the principles, practices and values of sustainable development into all facets of learning. It is hoped that this will promote behavioral change, encourage critical thinking and also effect changes in organisational and institutional practices that allow for a more just and sustainable society for everyone (UNESCO, 2007). The objectives identified for the DESD for implementation at all levels include the following: “i) To facilitate networking, linkages, exchange and interaction among stakeholders in ESD; ii) To foster an increased quality of teaching and learning in education for sustainable development; iii) To help countries make progress towards and attain the Millennium Development Goals through ESD efforts; iv) To provide countries with new opportunities to incorporate ESD into education reform efforts” (UNESCO, 2005b: 6).

The *International Implementation Scheme* (IIS) for DESD spells out a broad framework that identifies all stakeholders and their potential contributions. To be able to promote and advance the Decade has resulted in developing the following seven key strategies: "vision-building and advocacy; consultation and ownership; partnerships and networks; capacity building and training; research and innovation; use of information and Communication Technologies (ICTs); and monitoring and evaluation" (UNESCO, 2005b: 17). These strategies are seen to provide important opportunities for multi-sectoral and multi-level engagement of stakeholders in dealing with the essential action themes of environment, rural development, sustainable urbanisation, health promotion, gender equality, cultural diversity, peace and human security, and sustainable consumption. The IIS also emphasises the need for multi-national partnerships and coordination to strengthen collective ownership of and commitment to the Decade (UNESCO, 2005a).

Challenges facing ESD Implementation

Among the challenges and constraining factors facing ESD implementation are the following:

- The difficulty in identifying common ground between the various actors conducting ESD-related work across different sectors for proper coordination to guide and monitor relevant ESD efforts;
- The differences in mandates and thematic areas among partners and stakeholders and the challenge of how not to weaken the comparative advantage each partner brings as these mandates are aligned toward ESD;
- Identifying specific projects and activities from the onset to guarantee the maintenance of momentum from launch and initial activity;
- The challenge of transforming existing approaches to education towards ESD and the complexity of its introduction at the national and local level, as ESD is not only about teaching the subjects relevant to sustainable development but also about participatory learning process;
- Lack of cooperation and partnerships between stakeholders in research and development activities and unavailability of appropriate research data that might contribute to ESD regarding changes in the current system of education;
- Inadequate number of qualified and committed ESD personnel to coordinate the strategies and programmes of the ESD implementation and assessment (through monitoring) and evaluation. In addition, the challenge of significantly increasing human capacity through education and training afforded to government officials, school administrators and others in decision-making positions in ESD to integrate ESD into current policies and plans;

- The danger of ESD losing its priority status among most donor countries/agencies regarding funding of its activities as financing is gradually becoming a barrier to moving ESD forward due to the current economic situation (UNESCO , 2005b; UNECE, 2009; Tilbury, 2010; UNESCO APRBE, 2011; Ofei-Manu and Shimano, 2012); and
- In especially the formal education sector, against the backdrop of the current intertwining of the mental and physical structures of education, economic and the socio-politics, the major challenge is whether the move towards ESD should be an incremental reorientation of education towards sustainability or a paradigm shift. If an incremental change, how to reshape the established structures and ways of thinking about education and how to make ESD 'count' through recognition by adopting it as a "mainstream subject" for testing internationally and still uphold its transformational goal.[5]

Previous Approaches for M&E

DESD Implementation across Scales: The lead agencies and the initiatives they are engaged in

The UN DESD is a global platform that seeks to embed sustainable development into all learning spheres, to reorient education and develop initiatives that can showcase the special role of ESD (Elias, 2006 in Tilbury, 2010). According to Tilbury (2010), in 2005 when UNESCO became the official international lead agency for the DESD, it increased its responsibility and efforts in these areas and set in place a series of mechanisms to guide the Decade. Early on, UNESCO released an International Implementation Scheme for the DESD (UNESCO, 2004; 2005b) as well as an Action Plan for its own contribution to the Decade (UNESCO, 2005a). The establishment of a UN DESD Secretariat followed with a global coordination role, then a High-Level Panel to provide guidance and advice to UNESCO on the DESD, a UNESCO Reference Group on the DESD to assist on implementing the DESD strategies, a Monitoring and Evaluation Expert Group (MEEG) to advice on DESD reporting progress and an Inter-Agency Committee to ensure harmonious international coordination (UNESCO, 2007, Tilbury, 2010). The important role the DESD advisory panels had to play was to provide clarity and direction on DESD strategies and advice on how best to engage communities of practice which had yet to engage with the agenda (Mula and Tilbury, 2009).

[5] Additional challenges are discussed in Ofei-Manu and Didham (2012) and references there within.

Global

The declaration of the DESD in 2005 was marked by a series of high-level launches which served to raise awareness about the Decade's objectives with global support from stakeholders across all sectors (Paden, 2007; Tilbury, 2007; Tilbury, 2010). On the ground, governments are encouraged by the UN resolution to voluntarily implement the DESD in their countries (UN, 2005) hence leading to different levels of commitment. With regard to global ESD monitoring and evaluation which is identified as an important strategy of the International Implementation Scheme (IIS) in the context of developing indicators at all levels, UNESCO was given the responsibility to lead in the establishment of the necessary mechanisms (for M&E) and also to report on the progress made in 2010 and 2015 (Tilbury, 2010). As a result, the Monitoring and Evaluation Expert Group (MEEG) was established to provide advice regarding use of appropriate monitoring mechanism to assess the progress made on DESD implementation globally and also to assess the contribution the agency itself made to implementing DESD. After pulling together the experiences in developing monitoring systems and indicators for ESD, MEEG then recommended UNESCO to publish three global implementation reports with different foci during the course of the Decade:

1) For Phase I, the first report came out in October 2009 after it had been reviewed at the 2009 Bonn UN Conference on ESD. It focused on the context and structures, provisions and policies of work on ESD in member states put in place during the first half of the DESD in support of ESD development around the globe (UNESCO, 2012b). Bringing together a considerable amount of data collection and triangulation processes that could help validate the findings, the Global Monitoring and Evaluation Framework (GMEF) was developed to create opportunities for mapping national and regional developments in context and structures for ESD with heavy reliance on meta-analysis and voluntary contributions from key regional and national stakeholder groups as well as expert opinion with little empirical input (Tilbury, 2010).
2) The GMEFs for Phases I and II were expected to share common goals including DESD awareness raising among stakeholders, monitoring ESD progress across a range of sectors, providing opportunities for learning and reflection, assessing changes, providing a regional and global map indicating progress and making assessment of UNESCO's contribution to the DESD as well as the lessons learnt in the process of implementation. The GMEF for Phase II built on the data and lessons learned from Phase I (Tilbury, 2010). Furthermore, MEEG spelt out the following objectives for Phase II: 1) To clarify the learning processes that need promotion in order to facilitate learning in ESD and to identify ESD-related learning opportunities to promote and facilitate sustainable development; 2) To capture the entire

spectrum of educational levels and settings where the processes and learning for ESD are occurring; 3) Through engagement, identify the stakeholders involved in the processes and learning for ESD and to determine how they are involved in it; 4) To find out whether the existing processes for ESD aim to achieve both normative objectives and learning objectives; 5) "To examine (i) what has started to change, (ii) what has been learnt in the process of reorienting education systems towards ESD, (iii) whether opportunities for ESD outside of education systems have increased, and (iv) to what extent it is contributing to advancing sustainable development specifically in the context of processes and learning" (Tilbury, 2010: 104). The GMEF for Phase II has six mutually supportive components, the implementation of which would generate information regarding the different objectives identified for Phase II of the DESD M&E process and contribute to the Global Monitoring & Evaluation Progress Report for 2011. These components are expert literature review of processes and learning for ESD, ESD portal of experiences, case studies, key informant analysis, questionnaire, and assessment of UN contribution to the DESD (Tilbury, 2010). Phase II report comprising a companion literature review by D. Tilbury (UNESCO, 2011) and the main report authored by A. J. Wals (UNESCO, 2012b) which came out in mid-2012 was focused on multi-sectoral ESD-related learning processes occurring in various contexts of education, teaching and learning.

3) The third report to be produced in Phase III and due to be out in 2015 is expected to focus on the impacts and outcomes of the DESD (Tilbury, 2010).

The GMEF 1 faced several limitations that are documented in UNESCO (2009) and Tilbury (2010). Furthermore, the two published GMEF reports provide no clear format for validation and/or comparability, etc. among regions or countries in a particular region.

The UNESCO Education sector has produced Action Learning and Training Tools series to make available to governments, communities and individuals resource materials on ESD issues for teaching, learning, and training and to serve as guidance in practically implementing ESD on the ground. They comprise the following as at the end of 2012:

- The first toolkit firstly prepared for the North American audience but now can be used by all countries is entitled: Education for Sustainable Development Toolkit (UNESCO, 2006b). It is to assist communities develop sustainability goals through use of local educational systems and programs to modify existing curricula or reinforce those goals by creating new programs.
- Piloted in countries selected from Africa, Asia and Latin America and the Caribbean the Education for Sustainable Development Lens: A Policy and Practice Review Tool is aimed at assisting Member States in their attempts to reorient existing programmes in formal

education system, particularly at the school level. The aspect to be addressed by this tool is how "education policies, curriculum and other support processes sufficiently integrate the principles of ESD to inform and strengthen the quality of learning experiences for sustainable development..." (UNESCO, 2010:4).

- Exploring sustainable development: A Multiple-Perspective Approach is the third of the toolkit in the series designed for secondary school students and it is to provide a multi-level approach to education through understanding and working with complexity (UNESCO, 2012a).
- The ESD Sourcebook, the fourth publication in the series and meant for use by primary and secondary school teachers, teacher educators and mid-level decision-makers is to describe ways in which ESD can be integrated into primary and secondary schools and complement other materials already published (UNESCO 2012c).

Regional, Sub-Regional and National Initiatives towards the DESD

The Asia-Pacific region consists of six sub-regions, all of which are impressive in size and diversity. Together, they hold over half of the world's population. Some key challenges for consideration when interpreting ESD progress in the region include: i) cultural barriers, ii) diversity of geography, iii) governance and national coordination, iv) education, v) human resource capacity, and vi) natural disasters (Wals, 2010). In the context of development, some of the specific challenges faced under the four dimensions of sustainability are as follows:

- Environmental: Depletion of natural resources, loss of habitats and biodiversity, climate change, deforestation and desertification and water shortages, unsustainable farming practices including overfishing, pollution and disaster preparedness;
- Social: Access to education and healthcare including improving the quality of education and prevention of spread of HIV and AIDS, respectively, good governance, human rights, peace and human security, and gender equality;
- Economic: Corporate responsibility and accountability, food security, poverty, rural development, urbanization and disaster preparedness; and
- Cultural: Preservation of tangible/intangible heritage including valuation of traditional knowledge; safeguarding cultural and linguistic diversity and promoting intercultural and interfaith understanding (UNESCO APRBE, 2011: 12).

On the other hand, ESD in the Asia-Pacific region is now at a point where member countries are beginning to respond to the DESD initiatives in a variety of ways by engaging in clear strategies for ESD. They include "moving ESD from theory to practice by identifying clear thematic national SD

priorities, linking priorities to existing aims and objectives for education and learning in current policies, building inter-ministerial support, discussing financing and engaging with the right people at the national level" (Wals, 2010: 114).

At the UNU/UNESCO International Conference and Regional Launch of the DESD in Nagoya, Japan in June 2005, the *Working Paper: Asia-Pacific Regional Strategy for Education for Sustainable Development* (UNESCO APRBE: 2005a) was presented to serve as a guide regarding ESD implementation throughout the region. This working paper was based following the findings and recommendations identified in *A Situational Analysis of Education for Sustainable Development in the Asia- Pacific Region* (UNESCO APRBE: 2005b) including a number of core ESD issues and an overview of the region's current ESD status, to what level countries have incorporated ESD policies, programmes and practices into formal and non-formal education settings at local, sub-national and national levels, thus serving as a foundation for further planning and implementation of ESD-related initiatives in the region (UNESCO APRBE, 2007; Tilbury, 2007 and the references therein). Emphasising the importance of partnerships, the working paper further suggests potential roles for a selection of crucial stakeholders from governments, UNESCO National Commissions, communities, the private sector, education institutions, civil society, media, youth and international agencies, in addition to recommendations for engaging these stakeholders in ESD (UNESCO APBRE, 2005a). In pursuance of the previous activities and to assist Member States in the region to address the challenges facing effective ESD implementation, a series of coordination and capacity building workshops to enhance ESD leadership were organised in 2008 and 2009. Some of the lessons learned are to "Sharpen focus to reorient education systems towards national development priorities …. Incorporate ESD into national development and education plans … Establish support for capacity development ….. Internalize ESD within national budget structures …. Establish inter-ministerial support for ESD" (UNESCO APRBE, 2011: 16).

In addition, the Astrolabe – a tool for ESD capacity building and initiatives coordination and designed to contribute to the current attempts to promote the quality of education and learning in the Asia-Pacific region (UNESCO APRBE, 2011) – was launched to assist "countries in taking stock of ESD linkages in national policy, mapping current ESD-related activities and identifying key actors and their scale and scope of involvement in ESD" (UNESCO BKK, 2013). Regarding the sub-regional level initiatives, the Secretariat of the Pacific Regional Environmental Program (SPREP) began efforts in conjunction with UNESCO and UNEP to reorient its Regional Environmental Education Framework towards ESD (Tilbury and Janousek, 2007). Other activities include i) a situational analysis of ESD in the Pacific that was carried out and widely distributed within the region to serve as a baseline from

which to initiate activities (Elias, 2006; Tilbury, 2010), and ii) the endorsement of the Pacific ESD Framework prepared by a regional ESD Working Group at the request of the National Commissions for UNESCO. The Asia-Pacific UN Interagency Steering Committee for DESD and the Asia Pacific Regional Consultative Group for DESD facilitated by UNESCO Bangkok work to direct coordination of DESD efforts in the region (UNESCO APRBE: 2005b).

Member States of the Asia-Pacific region held workshops and symposiums in addition to DESD launches at the national level. These activities which were promotional and celebratory in nature brought together stakeholders and resulted in the development of working groups and committees, which are striving to advance and implement actions for the Decade at the national level. To maintain the continuous momentum of DESD, such initiatives are crucial. They are also vital for the generation of stakeholder support and their engagement with activities that advance progress towards a sustainable future (Tilbury, 2007). With several countries and organisations in the Asia-Pacific region at various stages of development of ESD implementation and assessment tools such as reporting formats, guidelines, techniques and frameworks, etc. a few countries that have finalized their frameworks for ESD implementation in the curriculum mainly in the formal sector include the following:

- Australia: Based on earlier ESD activities and the first Australian Curriculum – which includes sustainability as a cross-curriculum dimension, the Sustainability Curriculum Framework was developed to serve as a guide for curriculum developers and policy makers at the national, state and territory levels and is to provide "information and guidance on how education for sustainability may be structured to support a progression of learning from Kindergarten to Year 10" (DEWHA, 2010: 4) through an effective curriculum integration;
- New Zealand: The present curriculum documents dealing with education for sustainability are:
    - The New Zealand Curriculum comprising 1) Direction for learning – with vision, values, key competencies and learning areas as the subcomponents, and 2) Guidance – with purpose and scope and effective pedagogy as its subcomponents,
    - Te Marautanga o Aotearoa, and
    - The Education for Sustainability Teaching and Learning Guidelines for 11-13 year olds which has its foundations in the Ministry of Education's 1999 Guidelines for Environmental Education in New Zealand Schools and consists of components namely "rationale, key concepts, pedagogy, learning objectives, connections, learning programme design" and resources (NZC, 2013);
- Japan: Backed by several governmental policy plans such as Japan's Action Plan for UN DESD (2006) and Japan Council on the UN Decade of Education for Sustainable Development (ESD-J, 2006) to promote the integration of ESD into lower and intermediate levels of formal

education, the National Institute of Educational Policy Research of Japan embarked on clarifying how "curricula, teaching materials and instruction and evaluation methods should be for the purpose of embedding and strengthening ESD in schools" (Okamoto et al., 2012: 2). The ESD framework consists of six concepts, seven abilities and attitudes for ESD and three guidelines for ESD in school education (Kadoya and Goto, 2012).

Good examples of policy and implementation strategies and mechanisms and implementation at the national level are available.[6] It must also be acknowledged that a number of regions, individual countries and agencies/initiatives have attempted, or are trying, to develop their implementation and/or assessment (M&E) frameworks, strategies, approaches and appropriate indicators (Burford et al., 2013; ESDinds, 2012; several references in UNESCO APRBE, 2007 and ARIES, 2006). This is a step in the right direction because that is what the International Implementation Scheme encourages Member States to do (UNESCO 2005b). One important thing that needs to be addressed will be to find way(s) to facilitate these "country-specific" tools, strategies and approaches in such a way that appropriate information can be accessed by interested countries who can then tailor them according to their situation and use rather than to go through some of the tedious processes already taken by other countries or entities to obtain such information. It is hoped however, that general discussions on ESD indicators will be done in the next report.

Reasons for Measuring ESD Progress and Conducting Monitoring & Evaluation

ESD is dynamic in nature and is constantly evolving in various and different contexts including geographical location, culture, political system type and dynamics, time, etc. even though the core concept remains consistent. Formulating the appropriate monitoring and evaluation (M&E) frameworks (which has the key sectors for investigation, target areas and points of intervention as components) and finding appropriate methods and tools like indicators to measure ESD progress or otherwise is therefore both very challenging and critical, bearing in mind that the choice of appropriate ESD indicators "requires clearly articulated goals for DESD and an understanding of what indicators can and cannot assess" (Tilbury, 2007: 239).

Since the launch of the DESD to achieve the goal of embedding ESD in all learning spheres, attempts have been made to provide exemplars in the forms of capacity strategies, mechanisms, methods, practices and initiatives across various scales (Tilbury, 2010). The total results realised to date have been mixed with modest successes and also failures. A major challenge now is how to evaluate the

[6] Several cases of ESD implementation at the national level can be seen in Didham and Ofei-Manu (2012).

current status of ESD implementation in the context of how education/learning has contributed to sustainability and to systematically identify methods to further mainstream these important learning processes. Being a component of the seven key strategies of the International Implementation Scheme for advancing the UN DESD is indicative of the importance of monitoring and evaluation. M&E during the DESD helps to guarantee continuous relevance and effectiveness of ESD efforts with respect to planning guides, reorienting DESD programmes, increasing understanding of ESD progress, and improving decision making and action for the DESD. In addition, stakeholders are given the opportunity to engage in DESD activities (UNESCO APRBE, 2007). Another challenge besides developing the tools for M&E measurement is to seek to identify and strengthen the framework(s) or platforms that have the potential to effectively strengthen the implementation of ESD as well as its measurement, an example being the Regional Centres of Expertise (RCEs).

As the end of the UN Decade of Education for Sustainable Development draws near, the establishment of a systematic approach to document and assess the progress in implementing ESD becomes increasingly essential. Furthermore, it is necessary to understand what have been the main success factors and barriers in ESD implementation and practice, if we are to properly consider the future needs for improving ESD beyond the end of the Decade. An M&E of ESD process can help to monitor progress, to learn and improve from existing experience, and to influence future policy and practice.

SECTION 2

Approaches to Educational Assessment

&

Application to Education for Sustainable Development

The assessment of education for sustainable development is open to several different trajectories, and selection of one (or multiple) of these trajectories should be made through a clear evaluation of the benefits and deficiencies of each approach. Consideration of these compromises is especially necessary when the target is the establishment of a definitive set of ESD indicators or reporting criteria for the systematic monitoring and evaluation of ESD implementation, performance and/or achievements. Before identifying the approach and structure of an ESD M&E system, the desired scope, breadth and depth of the reporting should be clearly defined. For example, a more localised process will generally allow for a greater depth and relevance for the specific indicators or reporting criteria utilised. While an increase in breadth or multi-country coverage of M&E of ESD either at a regional or global level, will most likely result in the usage of more general and less specific indicators. Next, it is also important to clarify what is the purpose of the monitoring and evaluation and who will be the target user of the information generated from this process, or more simply what is to be learned from the M&E process. Possible considerations include the status of current ESD implementation, the level of existing knowledge/expertise on ESD held within a system, the quality of ESD learning being achieved, identifying areas for future policy interventions, etc.

Again, each of the above considerations come with different advantages and weaknesses, and although not all options are mutually exclusive certain choices do set a specific trajectory that greatly limits the inclusion of other options. There are also several logistical aspects that should be reflected upon in establishing an M&E system that further shape the type of assessment approach that is most favourable. These include the desired timeliness and regularity of data collection/reporting, the sources of information and the level of acceptable burden to be placed on respondents, along with the required time in compiling and assessing the given data. Before describing the approach taken in this project to conduct scoping research aimed at the development of an ESD M&E system for regional application in Asia and the Pacific, it is worthwhile to more generally review the main benefits and challenges of different approaches for education assessment.

General Purposes and Approaches of Assessment

First, in discussing educational assessment more generally rather than monitoring and evaluation specifically, it is important to recognise that the use of assessment tools can be applied for many different purposes. In education, assessment approaches may be used to account for the specific factors in a given context to ensure good project planning or effective interventions, to assess the efficiency of project implementation, to review the outcomes and impacts of a given initiative, etc. Drawing reference to how assessment is applied to the various stages of a project (or planning) cycle provides a useful example for simple explanation purposes. One of the more familiar versions of

such a cycle is the PDCA cycle, most often defined as plan-do-check-act. However, this can lead to some confusion over what “acting” entails as the “doing” is generally regarded as the project implementation while “acting” refers to taking corrective actions when the “checking” phase finds that the actual results do not correspond with the desired outcomes, thus for the use of discussion here it may be more helpful to refer to the A as adjust. Also, as is often advocated in planning processes, it is essential that there is some level of observation and due diligence at the beginning of the project cycle to understanding the context that is being dealt when creating plans. This provides us with a project cycle of observe-plan-do-check-adjust (OPDCA) including a final closing of the loop to create the actual cycling of the project development. In specific regards to education assessment, the concepts of diagnostic assessment, formative assessment and summative assessment will be linked to the observe, do and check stages of the project cycle respectively.

In the initial *observation* process, assessment and evaluation of the current context in which a project is to occur is very important, and there are many forms of assessment tools that can facilitate this type of stock taking and context setting. One form of assessment that is familiar across many professions relevant at this stage is a SWOT analysis (i.e., an investigation of the Strengths, Weaknesses, Opportunities, and Threats of the system being addressed), while more specific to the sustainable development field is a Strategic (or Sustainability) Environmental Assessment (SEA) which can be applied either for observing and stock taking or for evaluating the potential impacts of a given plan or programme in the next stage of the project cycle. At the stage of observation, the purpose is to gain an understanding of the current situation in order to effectively consider what is required to move a project towards its intended outcomes and align the planned actions in accordance with the present context. At an institutional level, another form of assessment that supports such observation is a capacity assessment of the given organisation or system (which will be discussed later in this section). These various forms of observational assessment can help to distinguish problems in the wider system that might need to be addressed before they become serious or opportunities that could be capitalised on if integrated into the planning. In conducting observation in regards to educational assessment, diagnostic assessment is regularly applied. Generally, diagnostic assessment is applied in educational settings to gain a baseline of existing knowledge, skills, values, etc. that the intended learners hold prior to new educational interventions with the purpose of appropriately designing curricula or learning material to build from the point of existing knowledge and attitudes. Thus, diagnostic assessment is conducted in the observation stage, but the collected information is used to strengthen the planning stage.

At the *planning* stage of the project cycle, assuming the previous stage has already provided a good understanding of the present context, then the types of assessment used at this stage are mainly aimed at detecting the potential impacts of the activities being considered and to identify opportunities for mitigating against harmful results. Environmental Impact Assessment (EIA) and Social Impact Assessment (SIA) are two approaches commonly used during project planning not only to mitigate potential negative impacts but also to help create co-benefits from project implementation in regards to positive environmental and social outcomes. Additionally, relational mapping methods such as impact diagram, systems diagram or network diagram can be utilised to help facilitate holistic, integrated planning processes. Relational mapping is just one type of methods among the various types developed and applied as part of participatory rural/rapid appraisal (PRA) methodology established through the work of many development practitioners working to secure greater community participation in both the observation and planning stages of international development projects (see Kumar, 2002 for overview of these methods). The relational mapping methods help to consider the main factors of influence identified during observation, to consider the interconnections between these various factors, and to map potential impacts of proposed interventions with the goal of identifying positive inputs that will support the overall strengthening of the given system.

Planning effective interventions can however be a daunting challenge, as it is necessary to: 1) specify the project objectives and desired long-term achievements, 2) review the existing gaps between the present situation and the desired achievements, and 3) identify potential actions that will support movement towards those objectives while also considering possible undesired impacts from such actions requiring mitigation responses. In relation to sustainable development (in its wider context beyond ESD), one of the major challenges is that appropriate interventions remain context dependent, thus there are no blue-print solutions that can be universally applied for achieving sustainable development. In fact, acknowledging the persistent implementation gap that continues to hinder the movement from broad SD strategies to clear action plans and on to effective implementation of SD initiatives, it can be postulated that one of the real barriers must occur at this stage of planning where ideals for sustainable development must be translated into a clear recipe for its implementation and practice. A similar parallel could also be made at this point for translating ESD ideals into clear and effective curriculum and course content.

It is beyond the scope of this research to theorise a solution to this challenge, but an additional series of methods have been gaining traction in this area as innovative approaches for addressing this challenge. These methods could loosely be categorised as futures assessments, and they include

visioning, scenario building, forecasting and backcasting. A key part of these methods are their use in collaborative processes where groups of multi-stakeholders (or sometimes selected experts) are brought together to create constructive dialogues regarding ideal future scenarios and then to map out the major movements or changes that need to occur between the present context and reaching this future scenario. A similar process for ESD could be conceptualised where the future scenario discussed is about the desired learning society we would like to achieve and what are the major components of knowledge, values and skills that underpin this society; from which the major changes in both education provision and socio-cultural systems would need to be considered in relation to the influences they have on each another.

The next stage of the project cycle is *do*, the *implementation* phase of the project. Assessment often does not feature heavily during this phase, nonetheless assessment at this phase can have a very beneficial role in ensuring that project implementation is occurring in an appropriate manner and leading towards its desired outputs. Potentially assessment at this stage may seem very similar to the goals for assessment during the subsequent check stage, thus leading many practitioners to argue that it is better to leave such assessment until the project is completed and full results are accessible. However, the main difference here is that assessment at the doing stage helps to strengthen project management and accountability in time to react to problems or challenges that arise before they become serious barriers to successful implementation. By identifying target achievements and performance standards within the establishment of project action/work plans, it is easy to facilitate quick assessment at this stage to review if implementation is occurring in a timely and efficient manner.

Although in project cycles, assessment approaches are less frequently applied at this stage of implementation, in the education field it is common for teachers to utilise formative assessment to gain an understanding of what are the levels of learning being achieved and in which areas students are gaining or not gaining proficiency. This allows teachers to adapt their intended lesson plans to better address the needs of their students towards better achievement of the learning objectives. Formative assessments can serve much like a mid-term review and by taking stock of current levels of learning achieved, teachers can both redesign their lesson plans and also reconsider their learning objectives for the overall course or curriculum. It is important to note that such redesigning and reconsidering can be in relation to both identified deficiencies or accomplishments in mid-term learning achievements. In this manner, in some cases formative assessment may highlight certain concepts that students have not fully grasped and must be returned to prior to moving forward, but in other cases it may reveal that students' knowledge attainment has progressed more rapidly than

expected thus the teacher may enhance the knowledge coverage and learning objectives for the course in light of his or her students' needs and capabilities. This type of formative assessment can be crucial in the effective delivery of ESD especially as teachers embark into unfamiliar teaching material, methods and learning approaches in line with the progressive nature of ESD content and pedagogy.

It is at the *check* stage of the project cycle where what is commonly referred to as Monitoring and Evaluation is applied, although some people may argue that "checking" is only akin to monitoring while evaluation falls under the first step of the adjust (or act) stage. Either way it is semantically divided up though, the process still remains fluid in practice; in order to conduct monitoring, data on outputs, outcomes and impacts must be collected and this data must then be analysed and evaluated to identify successes, failures, necessary interventions and corrective actions, etc. Of course, in this report the main focus is on monitoring and evaluation in order to determine the effects of impacts of the ESD practices that have already been implemented, and beyond this short hiatus explaining the wider applications of assessment approaches across the entire project cycle, discussion of M&E approaches remains the main content of this work (and specific methods for M&E will be elaborated later in this section). The purpose of the check stage in the project cycle is to gain an understanding of what has been achieved during the previous implementation stage and to consider if the desired outcomes and impacts have resulted. It is also important to identify potential barriers to successful implementation that have occurred and to acknowledge any unintended impacts or results. The knowledge generated during this project stage will directly inform the decisions taken during the following project stage when the main efforts are to make any necessary adjustments or corrections in the overall project implementation.

In the education field, summative assessment is applied at the end of a learning cycle or course of teaching to identify what has been achieved from the implementation of the course. Often, standard proficiency tests or entrance exams are used as the main form of summative assessment, however teachers can also conduct their own testing of student learning in regards to a single course or learning module. In regards to both M&E and summative assessment, it is valuable to acknowledge the link in the system that this approach creates in a project cycle or an educational process to ensure that a learning cycle is engendered in the overall system rather than it just occurring across a linear path from beginning to end. By conducting "checking" and then "adjusting", this strengthens the next round of planning and implementation by learning from and building on the strengths and weaknesses of the previous round. This type of link in establishing a learning cycle within the project cycle can be paralleled to the learning process that is described in experiential learning theory (Kolb,

1984). This theory proposes that holistic learning occurs both through, "…concrete/real world experiences and abstract thinking/reflection about a given situation as important processes through which we gain knowledge about our environment. Observation of action and reinterpretation of knowledge frameworks based on review of previous outcomes are continuous actions throughout the experiential learning cycle" (Ofei-Manu and Didham, 2012: 110).

It is important to make one distinction between the form assessment normally takes in regards to the check stage of the project cycle and the application of systematic M&E of educational activities. The type of assessment that occurs as "checking" is usually narrowly bounded towards solely addressing the outcomes and impacts resulting from the implementation that occurred in the previous stage of the project cycle. Systematic monitoring and evaluation of educational activities (and ESD specifically) requires a wider approach which not only focuses on the specific outcomes and achievements in regards to learning performance but also reviews the relevant educational policies and curriculum, the planning process for specific education interventions, the systems and institutions supporting implementation, and the modes of implementation. M&E of ESD may even include evaluation of the accountability mechanisms in place, and when comparing across countries this may include comparative evaluation of the specific M&E mechanisms each country has in place.

The final stage of the project cycle, or the last stage before it is reinitiated as a cycle, is *adjust* (although regularly referred to as *act* which can misconstrue its purpose). This stage is generally about taking lessons learned and the knowledge reinterpretations generated from the previous project stage and incorporating them into subsequent rounds of the project cycle for overall improvement. Sometimes, "adjusting" may be seen as a replacement to the original "observation" stage during the subsequent rounds of the project cycle, however if this is the case then it is important that observation of wider contexts are covered during the "checking" and "adjusting" as these contexts can shift/change during the project cycle and are not usually reviewed in these stages that focus more on the project results and achievements. If evaluation is linked in M&E as part of the previous stage, then there is usually no clear assessment methods distinctly applied to this stage. However, since in reality the project cycle is more fluid in nature some forms of assessment in relation to evaluation will support critical reflection on findings and outcomes in this stage. Furthermore, some aspects of observational assessment may also be linked to this stage as referred to above.

In an educational context, although diagnostic, formative and summative assessments are often distinguished from one another, as concepts they do not specifically refer to different assessment methods. Rather, the distinguishing feature of these three types of assessment are the stages within

the learning cycle (or project cycle) where such is being conducted, respectively for observation, for review of implementation, or for identifying outcomes and impacts. There are differences in the various methods of educational assessment that are applied at each of these stages, however it is also completely possible for one standard form of assessment to be reapplied across all three stages of educational assessment and achieve effective results. As such, an education course or curriculum would have set target learning goals and expected levels of knowledge attainment. An assessment method would be applied to judge students' learning and acquired knowledge on the given subject, and this method would be applied prior to, during and following the implementation of the education course/curriculum. This allows for a clear demonstration of knowledge gain over the course of implementation, the ability to assess the level of knowledge attainment and to consider if the learning goals have been achieved. This systematic application of assessment methods though is not commonly applied in educational context, rather the initial diagnostic assessment may only be informal in nature through general discussion with students, while the formative assessment may be a pencil and paper test established by the teacher, and finally the summative assessment may be based on standardised performance-based testing. The differences in the methods of assessment at each level may serve for meeting their individual purposes, but this can also complicate comparability and the ability to assess change across time. As will be discussed later in this section, setting a baseline from which monitoring occurs can be essential for clear evaluation of collected data especially when addressing aspects of ESD such as qualitative improvements and performance.

Standard Approaches for Educational Assessment

In order to establish a framework of the main factors and important leverage points in the implementation of ESD and the prospective future evaluation/assessment in the near future, it will be suitable to take a view of the features of the current "education/learning systems" namely traditional educational system (TES) and 21st century skills education (21CE). Notable differences between TES and 21CE include the following: 1) While TES involves direct instruction led by the teacher, in 21CE there is an interactive engagement with the learner and the teacher acts as a facilitator; 2) While TES is associated with rote learning and memorisation, 21CE is associated with both cognition and practical experience; and 3) While in TES there is competition among the students, in 21CE the students engage in collaboration. Further comparison will be done in relation to the approaches and scale of assessment of such education systems and consequently, the justification for calling for a new form of education/learning namely ESD. Although ESD assessment/evaluation framework(s) should cover the context and content (inputs), process (throughputs) and performance outcomes (outputs) in all three educational settings of formal, non-

formal and informal education if possible, this section focuses mainly on assessing performance outcomes in formal education.

Traditional Education Systems

The traditional education system (TES) has been around for several centuries in various forms in different countries. However, for the sake of this report reference is being made to the typical western style twentieth century education. TES arguably has served its purpose, albeit satisfactorily in the era of small population, adequate natural resources and a relatively smaller percentage of the global population whose large and unsustainable ecological footprints (or carbon footprint) have impacted the environment negatively. However, one of the main purposes of this system of education which is mainly to produce graduates with the knowledge and skills and a worldview to perpetuate the current socio-political structure and an economy that is set on the trajectory of unsustainability has recently been called into question. This is against the backdrop of a rapidly increasing global population especially in economies seeking to emulate the unsustainable lifestyles of the West, the inability of technological advancement to keep pace with resource depletion and increasing pollution, and a society disillusioned by a plethora of sustainability problems.

Although TES has several educational theories, methodologies and methods associated with it that are important because they are relevant and are seen to ground ESD in principle, it also poses several educational problems including the following:

- TES focuses entirely on the cognitive aspect of learning, an approach that is intended to allow students gain individual meanings about the subject matter at hand;
- It teaches students how to succeed on standardised tests largely based on rote learning, information memorisation or what is also called in some places the “chew, pour, pass and forget syndrome”;
- The social structure of TES is considered highly authoritarian. The teacher acts as the custodian of knowledge and the students largely as the recipients with their input in deciding how and what they are to be taught totally ignored, hence resulting in asymmetric power relations. Students are told what to learn, when to learn it, and how to learn it. This authoritarianism is mirrored between the administration and teachers where the administration has power over the teachers who in turn distrust the students assuming they do not want to learn. As a result, students only tend to value education/learning extrinsically and not intrinsically;
- Teachers are under pressure to “deliver” good results through teaching to the test so as not to risk putting the entire school into danger in terms of continuous funding;
- The topics to be covered, standards to be achieved, and curriculum to learn are mostly determined by the central government, textbooks and testing instruments are written by a few individuals/experts. Writing mostly from their offices they decide what is important for students to learn and know;

- Conformity is rewarded while diversity of thoughts and opinions is not. Oftentimes, students who do not fit into the system are failed and asked to leave or they quit themselves;
- Through no fault of their own but rather the result of the value-system and system structure which fails to instill personal values in them, most students think of education primarily as a way to get ahead of their peers and therefore seek to gain a comparative advantage over their fellow students becomes their prime objective. Also, most students are in in school not necessarily to learn but to complete it, get a degree and then a job (Bondelli, 2013);
- TES has been inaccessible to those who cannot afford it and has resulted in significant numbers of "uneducated" people. And even for those who can afford, particularly in the rich economies a considerable number do not possess the necessary skills to fill the available job vacancies, leaving such economies with significant levels of unemployment. In TES therefore, the focus on assessments-based results sometimes fails to supply what is needed in the real world job market hence prompting the search for a reformed education which can provide this.

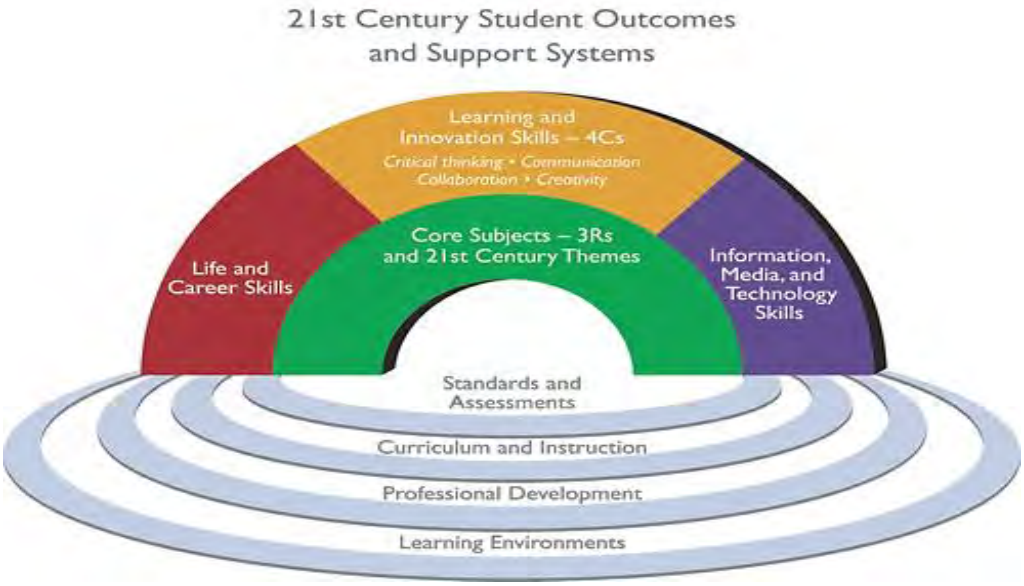
21st Century (skills) Education System and Assessment[1]

Twenty-first century education (21CE), although not necessarily new, has only been agreed upon recently by proponents (educators and policy makers) to be considered as an educational/learning system. It can be viewed as having evolved from TES mainly as a series of incremental reforms to help strengthen the relevance of modern education systems and the skill training they provide, and hence its features place it between TES and ESD. One major problem of TES that 21CE seeks to address is producing graduates that are ready to take up positions/jobs that were difficult to fill formerly due to lack of necessary skills. Developed by the Partnership for 21st Century Skills for conceptualising different skill types important for college and workforce (Kay, 2010), the Framework for 21st Century Learning consists of the following areas that must be assessed: "Core subjects (i.e., reading, language arts, world languages, mathematics, economics, science, geography, history, government and civics, and arts); 21st century themes (global awareness, financial, economic, business and entrepreneurial literacy, civic literacy, health literacy, and environmental literacy)"; learning and thinking skills; ICT literacy; and life skills (Kay, 2010: xv; Dede, 2010). Integrated with the foundation of the framework is an educational support system namely standards and assessments, curriculum and instruction, professional development and learning environments (Figure 1) (Trilling and Fadel, 2009).

[1] Note that there have been multiple recent attempts to incrementally reform/improve education to increase its relevance for required skill-provision in the 21st Century and especially to move beyond the narrow forms of traditional rote learning and assessment. 21CE was selected as an example in this study not as an endorsement by the authors, but as an archetypal example of the type of ameliorative modernisation programmes that have been advocated for education. In fact, this specific version of 21CE advocated by the Partnership for 21st Century Skills originated in the early 2000s in the United States with funding from the U.S. Department of Education. However, almost a full decade earlier, the UN had formalised the International Commission on Education for the Twenty-first Century which was chaired by Jacques Delors and delivered the seminal report on transformative education reform for the 21st Century in *Learning: The treasure within* (1996) and then subsequently *Education for the 21st Century* (2001).

Reviewing over 150 studies, Lai and Viering (2012) posited that critical thinking, creativity, collaboration, metacognition, and motivation are considered to be the core of 21st century skills and are important for reasons including the following: predicting important educational and employment outcomes on various settings (critical thinking), predicting achievement over and above ability (motivation), driving the development of new ideas, inventions and technologies (creativity), compensating for deficits in intelligence or prior knowledge of a subject, say during problem-solving (metacognition), and providing a more lasting learning and higher achievement than individual learning (collaboration).

Figure 1: Framework for 21st Century Learning



(By courtesy of Partnership for 21st Century Skills)

Furthermore, reviewing between 25 and 44 studies for each separate skill, Lai and Viering (2012: 11-26) were able to identify the components of these skills:

- Critical thinking skills:

  a) By definition, analysis of arguments, use of deductive/inductive reasoning to make inferences, judging or evaluating making decisions/solving problems; and

  b) By disposition, open-mindedness, inquisitiveness, the propensity to seek and attribute things to reason, the desire to be well-informed, being flexible, having respect for and the willingness to entertain other viewpoints.

- Creativity: Creative people have a high level of self-efficacy and are willing to take risks through openness to new ideas, sharing ideas and asking questions. They are often intrinsically motivated and also tolerant of ambiguity. Cognitively, they are capable of identifying problems, generating ideas through divergent thinking and actual problem solving.

- Collaboration: Comprises coordination, conflict resolution, communication, negotiation, problem-solving and decision-making.
- Motivation (intrinsic and extrinsic): Consists of an individual's beliefs, perceptions, values, interests, and actions.
- Metacognition: Consists of 1) cognitive knowledge (self-appraisal, epistemological understanding, and knowledge – person and task, declarative, procedural, strategy and conditional), and 2) cognitive regulation (planning, monitoring or regulating, cognitive experiences and evaluating).

The first four skills in particular are similar to some aspects (elemental characteristics) of ESD learning performance (discussed later in the third section) suggesting a common ground for these two educational/learning systems. In addition to the Framework for 21st Century Learning, a couple of closely-related frameworks are available. For example, the National Research Council (NRC) of Canada has developed a framework for categorising the types of knowledge and skills needed by students for college and career readiness. They comprise: 1) Cognitive skills that include critical thinking, problem-solving and critical thinking; 2) Inter-personal skills consisting of teamwork, dealing with diversity, cultural sensitivity, complex communication and other social skills; and 3) Intra-personal skills comprising self-development and self-management, self-regulation and adaptability/resilience and time management (Lai and Viering, 2012). Another framework known as the ATC 21 Framework developed by the Assessment and Teaching 21st Century Skills organisation offers four categories of skills:

> "1) Ways of Thinking …. Creativity and innovation; critical thinking, problem solving, and decision making; and metacognition or learning to learn … 2) Ways of Working, which includes communication and collaboration or teamwork … 3) Tools for Working, which addresses information literacy and information and communication technology (ICT) literacy… 4) Living in the World, which includes citizenship, life and career skills, and personal and social responsibility" (Lai and Viering, 2012: 4).

The report of the UN's International Commission on Education for the Twenty-first Century also proposed a new framework on 21st Century learning areas in the seminal work on transformative education *Learning: The treasure within* (1996) as the Four Pillars of learning, and which in turn ESD has been acknowledged as elucidating a fifth pillar (see box 2 for full details).

Box 2: Five Pillars of ESD

Learning to Know: is the process of mastering learning tools and building the capacity to be a life-long learner.

Learning to Do: focuses on occupational training and educating people to be valuable assets in their employment while also acknowledging the adaptive labour needs of the modern market.

Learning to Be: entails supporting the full development of each individual and their self-expression.

Learning to Live Together: entails education into citizenship and social life thus providing individuals with the abilities to participate in cooperative communities.

Learning to Transform Society & Change the World: is for individuals to gain the skills and knowledge to achieve lofty social goals such as social equality, non-discrimination, social solidarity, transition to a low-carbon society and to live sustainably

NB- The first four pillars were originally identified in *Learning: The treasure within* (Delors et al. 1996), while the fifth pillar is considered a new edition to the learning pillars elucidated by ESD (Shaeffer 2006).

While the 21CE addresses the criteria needed to succeed in the current economic system, it is silent or barely mentions the looming crises facing the very environmental systems it depends on: depletion of non-renewable resources, exposure of global and regional ecosystems to danger and several of the planetary boundaries exceeding their safe limits, current unsustainable production and consumption, waste and pollution, economic inequity and injustice, and short-term view of the world. The learning/teaching objective of ESD however, is to re-orient/redirect the current socio-economic systems which create jobs towards sustainability, an aspect which 21CE fails to address.

While the mission of all the systems of education is to 'develop productive citizens', for TES using mainly traditional assessment, individuals must possess a body of mainly knowledge and some skills which must be taught by the school, and then students tested afterwards to see if they have acquired the knowledge and skills to be counted as productive. In performance-based assessment (PBA) which has become popular with 21CE, the individual, team or group must be capable of performing meaningful tasks in the real world and with the help of teachers/facilitators, become proficient at performing these tasks that they are bound to encounter after graduation and then prove their capability by successfully completing the tasks set by the school. ESD in addition to tilting heavily towards PBA's knowledge and skills approach also takes perspectives and values considerably into account. Table 2 (on the following page) indicates comparisons of these three types of education based on some features.

Focus of Educational Assessment

Assessment traditionally is an activity used essentially to evaluate a student's understanding of factual knowledge (Trilling and Fadel, 2009). The main purposes of assessment include the following: 1) by evaluating the strengths and weaknesses, student learning can be promoted, 2) it can also serve as evidence for meeting accountability requisites. Through sound assessment practices, teachers can obtain appropriate feedbacks on the quality of education they provide. Assessment is also considered to be the third side of the education triangle along with teaching and learning (Birenbaum, 1997: 71), and it is further seen as one of the elements that interact in the teaching-learning process. The other elements include methods adopted by teachers and students/learners, the nature of the learning material, and prior knowledge of the relevant material (Beckwith, 1991). In that regard, assessment, which is often seen as framing learning is incapable of being understood when isolated from it (learning). Assessment is also seen as a form of information gathering as the more test information there are about students, the clearer the picture about achievements and gaps becomes.

Table 2: Comparing Features of Traditional Education, 21st Century Education and ESD Systems

	Traditional Education System (TES)	21st Century Skills Education System (21CE)	Education for Sustainable Development (ESD)
Purpose of learning	Learning for certificate or degree that stops after graduation. It promotes simplicity of problems and offers simple solutions	Learning for "future life", mainly to acquire skills for job and career and fulfilling one's civic responsibilities. It embraces complexity	Life-long learning hinged on the "Five pillars of learning" i.e., "learning: to know, to do, to live together, to be, and to transform oneself and society". It embraces complexity and promotes multi/ inter-disciplinarity
a) Instruction type b) Teacher involvement	a) Direct instruction (didactic) b) Teacher directed	a) Interactive exchange (engaged) b) Learner centred with teacher facilitation	a) Both direct and interactive to fit the situation; it is exploratory b) Active, learner-centred with teacher facilitation
Type of education and place of delivery	Mainly formal education	Formal education and some level of non-formal education	Formal, non-formal and informal education all considered important
Aspect of education emphasised	a) Content (curriculum) b) Highly theoretical	a) Largely content with some amount of process (curriculum and projects) b) Theory with a level of practice	a) Content and process (curriculum and projects) both stressed b) Emphasises practice and relevant theory, i.e. critical praxis
Learning style	Rote learning, information memorisation; mainly cognitive	Mainly cognitive with some level of practical experience	Both cognitive and practical experience including experiential learning (values, etc.) with room created for needed adjustment
Relationship of learners with others	Competitive	Collaborative	Collaborative, cooperative, networking, participatory
Teaching/learning package for target learners	One-size-fits-all. However, subject contents are clearly defined	Personalised but also fits teamwork. Subject content well defined	Personalised and also for team/ group-work. Pedagogical processes and scope of curricular contents yet to be clearly defined
Place and Format of delivery	a) Classroom b) Mainly text-based	a) Classroom + Global community (networks) b) Both web-based and also text-based	a) Classroom + Global community (networks) b) Both text-based, web-based, and field based; with increasing trends towards later two
Flexibility to real-world situations	Relatively rigid, it struggles to provide the workforce needed for changing situations	Has adapted to current situation, hence is able to provide the workforce needed for present situations. It is unclear though if it has ability to continuing adapting for changing situations. It is limited in shaping the systems and structures towards sustainability	Adaptable hence able to provide the workforce needed for changing situations. Structure on resilience, hence is capable of helping shape systems and structures towards sustainability
Underlying philosophy versus sustainability	Not sustainability friendly	Neutral, though some aspects tilt toward sustainability but not without additional information	Sustainability-friendly

Consequently, the use of standards, criteria and assessment schemes as a means to determining student achievement (through external tests or teacher judgment) has become popular and is now considered to be more 'effective' than the grades and marks obtained from traditional pen and paper examinations (Hay and Macdonald, 2008; Tognolini and Stanley, 2007). One reason for this is because it is considered as uninfluenced by achievement levels of other students. Recently however, the scope of assessment has widened to include skills and values when the educational reform proposals by 21st century skills and education for sustainable development are taken into consideration.

Usually conducted by expert external judges, faculty/teachers, peers, employers and parents with particularly students and teachers/faculty as targets, what is usually assessed include:

- The achievement in a subject area or major: Assessing an activity can establish the extent to which students have learned the content, skills, and attitudes of a particular major/subject. It could also involve the assessment of the overall/aggregation of subjects. The quality of a subject/ is easier to assess when overt learning outcomes for the subject have been formulated;
- Performance patterns: Programs may seek for answers to particular aspects of student performance, for example in relation to gender, time of joining the program, etc.;
- Quality of instructional practice where faculty/teachers may want to evaluate the impact of a newly introduced instructional technique with respect to improved learning over other instructional methods. The quality of the assessment measure is dependent on the validity and reliability of the measurement and also how appropriate the targeted learning goal is in line with the mission and goal of the program (Pusateri, 2009: 2).

Some aspects of assessment worth noting irrespective of the education/learning system include:

- Assessing student learning is inclusive of educational values;
- Assessment is maximised when the existing programs it seeks to improve has clear and explicitly stated purposes. "Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time" (Pusateri, 2009: 6);
- Equal attention ought to be paid to the learning outcomes as well as the experiences that lead to those outcomes during assessment;
- Assessment that is ongoing and consistent rather than being irregular hence unpredictable works best. Additionally, the involvement of all stakeholders from the entire educational community ensures wider and deeper improvements in assessment;
- The likelihood of improvement in assessment is enhanced when it is part of a larger framework to promote change; and
- Assessment is considered a form of public service through which educators render themselves accountable to students and to the general public (Pusateri, 2009: 6).

Assessment and Evaluation

The description of assessment and its relationship with evaluation in the literature is quite confused. Even though the two differ from each other fundamentally in purpose and also in the way they use the information gathered, yet the same tools can possibly be used for the two approaches. While assessment is "a set of processes designed to improve, demonstrate, and inquire about student learning" (Mentkowski in Purdue University, 2012), evaluation is "the systematic process of determining the merit, value, and worth of someone (the evaluee, such as a teacher, student, or employee) or something (the evaluand, such as a product, program, policy, procedure, or process)" (Purdue University, 2012). Evaluation occurs as a one-off by marking or grading students' work immediately following completion. Student assessment on the other hand is a longer process where information on student performance is collected throughout the learning process to measure overall learning and understanding.

Table 3: Differences between Assessment and Evaluation

Key Attributes	Assessment	Evaluation
Timing	Mostly formative	Mostly summative
Focus of measurement	Process-oriented	Product-oriented
Relationship between administrator and recipient	Reflective	Prescriptive
Use of findings	Diagnostic	Judgmental
Standards of measurement	Absolute	Comparative
Measures taken on modifiability of criteria	Flexible	Fixed
Relationship between the objects of assessment/evaluation	Cooperative	Competitive

(Adapted from Straight, 2002)

Assessment and evaluation however, can be used concurrently in a particular teaching/learning situation and "they can occur at a rather small scale (e.g., classroom) or a rather large scale (e.g., programs. For example, an instructor can use the results of a midterm exam for both assessment and evaluation purposes. The results can be used to review the students course material related to common mistakes on the exam (i.e. to improve student learning as in assessment) or to decide what letter grade to give each student (i.e. to judge student achievement in the course as in evaluation)" (Purdue Univesity, 2012). However, while assessment might be favorable at the level of individual students, in terms of ESD some level of comparative, multi-country evaluation is quite beneficial to identify what have been the major strengths and weaknesses in ESD during the decade.

Both monitoring and evaluation which make up the core theme of this report address programme performance primarily based on the achievement of goals and objectives. Although monitoring mainly involves operational and strategic issues by asking questions focused on efficiency, tracking continuity in programme performance and examining progress towards the programme objectives, evaluation which is a systematic judgment of the programme itself deals with strategic analysis of say the educational/learning programme that informs practice and assesses impact, appraises results in relation to the programme goals, explores the added value of the programme to inform future work and establishes written record of practice (Liddy, 2010).

Standard Categories of Educational Assessment

Assessment can be divided into a several categories including the following in the table below.

Table 4: Categories of Educational Assessment

Category	Example
By Process	Formative assessment and summative assessment both of which occur in the classroom particularly the former but can also be seen at the policy level (OECD-CERI, 2008)
By Activity Type	Whether paper-test based or task-based (i.e., traditional assessment and performance-based assessment)
By Scale of Conduction	From individual, classroom, school, district/statewide, national to international

The various levels of assessments are interconnected and gathering data at these multiple levels can be used for various decision-making purposes. Assessment may also be quantitative or qualitative in nature. Tests such as multiple-choice or fill-in-the-blank questions are quantitative, while tests involving application or synthesis are qualitative.

Assessment by Process: Summative Assessments versus Formative Assessments

Attempts to distinguish between summative assessment (SA) and formative assessment (FA) have generally been problematic as the important difference between them "is not when they are used but their purpose and the effect that these practices have on students' learning" (Hernandez, 2012: 490). FA can be differentiated from SA as allowing for practice, not holding students accountable for the skills and concepts they might be learning, hence helping teachers to determine the "next steps during the learning process as the instruction approaches the summative assessment of student learning" (Garrison and Ehringhaus, 2012: 1-2). Second, in FA students' involvement in the assessment process is key for effectiveness of the process.

Summative Assessment (SA)

Summative assessment strategies – including "methods that involve a single episode of data collection (e.g., nationally or locally normed tests) as well as those that incorporate tracking student performance over time (e.g., portfolio, case studies, longitudinal studies)" – are normally employed for program quality rather than providing feedback on students' progress (Pusateri, 2009: 19). The importance of summative assessment (SA) lies in the ability to periodically determine what students know (or otherwise) as well as student's learning relative to content standards at a particular point in time (Garrison and Ehringhaus, 2012). Summative assessment therefore helps in evaluating certain aspects of the learning process. However, because their occurrence is far and between, taking place once every few weeks, monthly, quarterly or annually after instruction, SA (relative to FA) "happen too far down the learning path to provide information at the classroom level and to make instructional adjustments and interventions during the learning process" (Garrison and Ehringhaus, 2012: 1). Summative assessments can be used as "tools to help evaluate the effectiveness of programs, school improvement goals, alignment of curriculum, or student placement in specific programmes" (Garrison and Ehringhaus, 2012: 1). In addition to providing feedback to improve quality, SA also promotes coherence in curriculum planning and offers support to earlier curriculum recommendations. Disadvantages of SA include high cost and high labour intensity, and students may not receive direct feedback with regard to their performances, hence resulting in limitation on their own gains from the effort invested in.

Formative Assessment (FA)

Formative assessment forms part of the instructional process. When made to be part of the classroom practice, FA provides the information necessary for the real time adjustment of the teaching and learning process. Descriptive feedback, a component of FA and one of the key instructional aspects of involving students in their own educational process is considered one of the most significant to help students advance in their learning. In the context of a learning-oriented assessment (Hernadez, 2012) or learner-centred assessment (Webber, 2012), continuous assessment supports "formative function for learning and summative function for certification" (Hernandez, 2012: 490) by providing feedback to students during their learning. Timely feedback, revising and proceeding with the set target activities all aimed at achieving the designated learning goals are extremely important (Wilson and Scalise, 2006). It should also be noted that some assessments are designed to be both formative (by producing feedback from students) and simultaneously summative because a grade is awarded which contributes to the overall results of what was learned (Hernandez, 2012; Yorke, 2003). According to Garrison and Ehringhaus (2012), some of the instructional strategies for FA include the following: a) criteria and goal setting – the

need for students to understand the goal and necessary criteria for its achievement; b) observation based on gathering evidence of student learning to inform instructional planning and for feedback for students; c) questioning strategies that have to be part of the lesson, d) creating a kind of learning community within the classroom through self and peer assessment; and e) student-led record keeping to help them better understand their own learning.

Increase in use of modules, curricular unitisation, fewer staff, increase in student diversity and plagiarism are some of the reasons contributing to increase in use of summative assessment over formative assessment (QAA, 2007; Hernandez, 2012). In a balanced conventional assessment system however, both types of assessment are integral parts of information gathering (Wren, 2009), although striking a balance between the two without excessive focus on one over the other which then becomes problematic, is not easy (Hernandez, 2012). In monitoring and evaluation of ESD implementation, while formative evaluation can significantly add to the impact of ESD programmes during the span of its implementation, summative evaluation can provide the written account of the work accomplished.

Other forms of assessment worthy of mention but which will not be discussed in detail are:

- Diagnostic assessment is intended to improve the learner's experience and their level of achievement by looking backwards to assess what the learner already knows and/or the kind of difficulties that could limit the learner's engagement in new learning if undiagnosed;
- Dynamic assessment which measures what the student achieves when learning an unfamiliar topic or field, an approach that is useful to assess one's potential for specific learning in the absence of relevant prior attainment;
- Synoptic assessment normally allows for the demonstration of the learners' capability to combine and apply their knowledge, skills and understanding in the subject by integrating the components of their learning obtained from different parts of a programme and also be able to demonstrate their accrued knowledge and grasp of a topic or subject area;
- Criterion referenced assessment allows each student's achievement to be judged against specific criteria irrespective of the performance of other students. Reliability and validity are especially considered key in this type of assessment compared to all the others; and
- Ipsative assessment is the assessment used against the student's own previous standards to measure against their most recent piece of work how well a particular task has been undertaken against the student's average attainment, against their best work (University of Exeter, 2013).

Assessment by Type of Activity: Traditional Assessment versus Performance-based Assessments

In comparison *w*ith the traditional assessment (TA) which usually takes the form of a paper and pencil tests and includes standardised testing at the national level such as university entrance examinations, performance-based assessment (PBA, also known as authentic assessment, alternative assessment or direct assessment) became popular in the late 1980s-early '90s and requires that students perform tasks instead of selecting answers from a ready-made list. This can range from generating a scientific hypotheses and performing an experiment, writing a topical essay, designing and constructing a model, explaining a historical events orally or in writing, solving math problems to conversing in a foreign language. PBA is therefore a method used to assess/evaluate students' knowledge, concepts or skills by using the said components to perform tasks designed to emulate real-life situations (Wren, 2009). Important components of PBA include the following: 1) defining the purpose of the assessment by knowing which concepts, knowledge and/or skills should be assessed; 2) choosing the type of performance activity by giving consideration to factors including available resources, the amount and level of data needed to cross the evaluation threshold regarding the student's performance and time constraints; and 3) developing the scoring criteria, i.e. whether to use for example a rubric or some other scoring criteria (Wren, 2009). Both TA and PBA involve both formative assessment and summative assessment at one point in time during conduction.

Differences between Traditional Assessment (TA) and Performance-Based Assessment (PBA)

PBA is considered as having a number of advantages over TA (Liskin-Gasparro, 1997; Mueller, 2008; Wren, 2009). A summary is as shown in table 5 on the following page. PBAs are considered as more in line with instruction than multiple-choice tests of TAs and therefore teaching to the test (item teaching) is encouraged by its advocators in spite of it being considered unethical when it comes to preparing for TAs. According to Wren (2009), this justification is based on PBA's flexibility for providing students access to scoring rubrics in advance in order for them to know exactly how their performance will be evaluated. Lastly, "there is limited potential for traditional tests to measure higher-order thinking skills since, by definition, those skills involve analysis, interpretation, and multiple perspectives" (Liskin-Gasparro, 1997).

Table 5: Differences between Traditional Assessment and Performance-Based Assessment

Attributes	Traditional Assessment (TA)	Performance-based assessment (PBA)
Assessment Activity	Student selects a response	Student performs a task
Nature of Activity	Contrived	Emulates real life
Cognitive Level	Knowledge and comprehension (indirect)	Both acquisition and application of knowledge
Testing Domain	Examinee is tested on a single knowledge area in order to prevent ambiguity	Tasks are multifaceted and ill-structured (i.e., both the goals and methods for problem-solving are not clearly defined, and is an initial part of the students' task)
Time-frame of Measurement	Measures taken once and often relying on machine-scoring (i.e., single correct responses); provides no chance for demonstration of student thought processes, student-teacher interaction or revision for improvement	Entail long-term projects and focuses on processes/rationales resulting in multiple 'correct' answers; learners are stimulated to find appropriate solutions
Accessibility to Assessment Information	The content for assessment is securely kept from test takers; improvement of student performance is by studying and memorisation	Students have a prior knowledge of the content (knowledge and skills); as assessment is considered both a learning activity and an evaluation device with preparation occurring through "learning by doing"
Involvement of the Teacher in Assessment	Teachers may be responsible for test format and content (although not usually considered in the development of large-scale, external tests); because teacher's input is not needed to reflect the correctness of a student's answer, greater distance between teachers and students is created and the entire assessment programme is not seen as enriching	Teachers serve as a key contributor by collaborating to create tasks for assessment and by helping to develop guidelines needed for scoring and interpreting the answers
Level of (order) Thinking Skills	Lower order thinking skills	Higher order thinking skills
Reliability[a]	Reliability and efficiency are higher as they can be statistically analysed and compared	Because testing involves students constructing complex, open-ended responses, PBAs are ill-structured, messy, and complex by design and users will have to contend with reliability-related issues
Validity[b]	TA approaches are seen as more valid	PBAs do well by reflecting real-life tasks constructed within curricula and pedagogies that are multi-faceted in character when it comes to validity; consequently, developing PBAs standards for evaluation/assessment and consistent application is much more difficult across a set of tasks than it is for a multiple choice, paper-and-pencil test

[a]The closeness of a score a student obtains in a specific assessment as a reflection on his/her possible "true score"

[b]How well a test measures what it is actually expected to measure (Liskin-Gasparro, 1997)

Assessment by Scale: Classroom, School, National and International Assessments

Classroom/Student Assessment

Assessment of classroom practice can range from judging formal written work of varying content and structure to those based on dynamic interactions in classroom performances (Stanley, 2011). Evidence shows improvement in learning that involves students in their own assessment and the use of classroom assessment for learning (Hill, 2011). With regard to the influence of assessment methods on students' learning approaches, the results "suggest that both the quality of student learning and students' pursuit of higher grades are enhanced by the careful selection of an assessment method that firstly, encourages students' development of higher order intellectual skills and the employment of deeper learning approaches; and secondly, allows students to demonstrate their development" (Scouller, 1998: 470-71). In a study on assessment preferences of students and how they relate to learning strategies and orientations based on two university level academic disciplines which differ in their educational environments, results showed that individual differences in assessment overshadow disciplinary group differences and also differences in assessment preferences were found to be significantly correlated with learning strategies and orientations (Birenbaum, 1997), a finding that has already been established in lower levels of education.

One other conflict between FA which is normally associated with higher order teaching – where teaching for understanding is emphasised as opposed to rote memorisation and multiple choice questions (MCQ) for test scores normally associated with SA (William et al., 2004) is that although it is generally accepted that increased use of FA can lead to higher quality learning, it is not necessarily reflected in external examinations often associated with SA. William et al. (2004) however point out that teaching that is attentive to higher order goals can result in higher attainment even when such attainment is measured principally in terms of lower order goals like tests.

School Assessment

According to Hill (2011), school-based improvement in learning outcomes based on assessment is achieved through school-wide improvement programmes and assessment for learning (AfL) in schools within the formative assessment context which prioritises feedback and promotes the important role that learners play in the assessment and learning process. Also, the need for sharing of power and control in the assessment process is conditional for inclusion of learners as full members of the learning community and to be able to see themselves as capable learners. The AfL-based school level change or improvement is hence affected by several factors: the principal's role, senior staff members' involvement, alignment of assessment for learning (AfL) with teachers' qualification assessments, changing the facilitation model and how assessment for learning is

embedded in the school culture. Consistency of alignment to standards is positively correlated with the richness of the school's source of evidence and some school systems have shown that school-based assessment satisfy rigour in high stakes contexts (Stanley, 2011). Others have not. In a study comparing school assessment with grades obtained in external tests, there was random variation of the external tests within a time period as compared with the little variation within the school assessments suggesting teachers leniency in assigning scores. However, whether involving external tests or teacher judgment in the classroom level or school level, a key design requisite for quality assessment is ensuring alignment to curriculum goals and expected outcomes (Stanley, 2011). Often, school systems with external tests are criticised for teachers narrowing or directing the curriculum to the 'test requirements'. One major driver of this is the pressure on teachers to 'deliver' good student results that are often used to manage teachers and schools with regard particularly to funding (Ofei-Manu and Didham, 2012; Stanley, 2011).

Teacher Assessment

One advantage of teacher assessment is the ability to ensure the teaching and learning programme at the school supports and grounds the evidence about student performance (Stanley, 2011). Currently, teacher competence is assessed using standardised objective tests and certification exams, occasional formal observations, task-based standards-driven and job-related multiple assessment, and portfolios that reflect the day-to-day work of the teacher in addition to standards that define what a quality teacher knows (Wilkerson and Lang, 2004). The primary aim of teacher assessment – underpinned by a manageable assessment regime in terms of appropriate workload, consistent and quality engagement by teachers and students and timely feedback – is to make judgment about students' progress in addition to the next step in learning. It is to ensure that teachers are competent to help students achieve the academic standards to which they are held (Wilkerson and Lang, 2004). According to Stanley (2011), there are a number of issues to address when it comes to effective teacher assessment for an education system. They include: 1) the level of workload – i.e., manageability of the assessment practices, amount of evidence collected/recorded and how it was done, and 2) quality engagement between students and teachers.

External Assessments: National, Regional and International Assessments

External assessments are assessments that use "criteria (rubric) or an instrument developed by an individual or organisation external to the one being assessed" (AAC&U, 2002: 1). In other words, the contents of the assessment have little or no input from either the instructor or the assessed irrespective of the location of the assessment. Primarily "summative, quantitative, and often high-stakes", they range from district/statewide, national, regional to international. Reasons for

introducing external assessment include the following: to assist learning institutions to make improvements, to facilitate accountability to stakeholders by rending public accountability for standards achieved and for use of money, to inform (potential) students and employers about standards, make publicly available information on quality and standards, and to assist government in making funding decisions (AAC&U, 2002; Pusateri, 2009; Greany and Kellaghan, 2008).

National Assessment

"A national assessment is designed to describe the achievement of students in a curriculum area aggregated to provide an estimate of the achievement level in the education system as a whole at a particular age or grade level" (Greany and Kellaghan, 2008: 7). National assessment (NA) is seen as important source of data for national education audits usually undertaken to equip policy makers with information about important features of that country's system of education. Furthermore, NA "is a tool for providing feedback on a limited number of outcome measures that are considered important by policy makers, politicians, and the broader educational community" (Greany and Kellaghan, 2008: 17). Reasons for conducting NAs include the following: 1) To gauge the country's readiness to participate in international assessments; 2) To provide information on the operation of the system of education that is considered "relevant to the work of curriculum developers, examination bodies, teacher educators and teachers" in general as well as for policy makers; 3) To determine whether standards remain static, deteriorate or improve when administered over a period of time; and 4) the possibility of using NA results to effect changes in practice with regard to the teacher's behavior towards students in the classroom (Greany and Kellaghan, 2008:17-21).

Traditionally at the national level, institutions have been evaluated against many criteria (Liu, 2011). For example, in Canada, apart from the province of Nunavut, all provinces and territories administer a type of large scale assessment (LSA), a programme mandated by the government and called School Achievement Indicators Programme (SAIP) which "focuses on school achievement indicators in mathematics, reading and writing, and science" (Volante and Jafaar, 2008: 203). Reporting is at the provincial and national levels only. SAIP has since been replaced by the Pan Canadian Assessment Programme. According to Stanley (2011), the application of educational standards generally requires a mandatory curriculum specified more in detail than before. This has been done with variations across national systems regarding the levels of prescriptions that range from broader curricular outcome statements as seen in the UK and Australia to more specific content in the US system.

Some decisions to consider in conducting a NA include the following:

- Who is to give policy guidance for NA – Usually a national steering committee will be needed to give policy guidance to the NA;
- Who is to carry it out – A team/organisation which is credible and whose work will command respect. Various countries assign responsibility for NA to groups ranging from teams within the ministries of education, autonomous bodies (universities and research institutions) to non-national technical teams (Greany and Kellaghan, 2008);
- Who to administer the tests and questionnaires which varies by country: from use of school inspectors, experienced teachers drawn from non-participating schools and retired teachers to graduate students with the requisite abilities.

The rest are: what population to assess and whether to assess a whole population or sample, what subject areas to assess, how frequently to carry on the assessments, how to assess student achievement and report it, the kinds of statistical analyses to employ, how the results of the NA will be communicated and used, and what the cost components are (Greany and Kellaghan, 2008).

The results of NAs from 15 countries (nine of which were summarised from Greany and Kellaghan, 2008) in Table 6 show both similarities and differences in the parameters considered.

- Similarities: Assessments were performed in language/literacy and mathematics/numeracy in at least one primary-grade level in all countries. Except Singapore and UK which provided no data and Chile, Australia and Uruguay (and a separate long-term national study in the US) where population was used, assessments in all countries were based on samples.
- Differences: There was a variation in frequency of assessment among countries ranging between one and four years. Variation also existed among agencies in charge of implementing the assessments. They include the ministries of education, national research councils, universities, a research institute supported by the government, and a national examinations board, etc. Implementation agencies in several countries received significant non-national support. Countries such as Uruguay, South Africa, Chile and Canada reported change in the implementation agency between assessments. In what could be attributable to the technical capacity of national assessment teams, the methods of analysis also varied considerably. Also, the description of student achievement ranged “from citing the mean and distribution of the number of items to which students responded correctly, to determining the percentage of students whose performance reached ‘expected’ standards or the percentage scoring at varying levels of ‘proficiency’” (Greany and Kellaghan, 2008: 81). Although inconclusive due to the lack of information on the extent to which the results have been used and consequently contributed to shaping policy in most countries, significant variation existed in the use of these assessment results.

Table 6: Summary of National Assessments (NA) of Achievement Levels of Education (of some countries around the world)

Country	Purpose of assessment	Grade Frequency (in bracket)	a) What was assessed b) What kinds of assessment are used?	a) Sample/population b) Analyses	Who did it	Use of results
Chile	To provide: 1) Feedback to the stakeholders namely parents, teachers, school authorities, etc., on the extent to which students are achieving the set learning targets; 2)Data for policy makers to guide allocation of resources in textbook and curriculum development and in in-service teacher education.	Grades 4 and 8 (annual).	a) Spanish (reading and writing), mathematics, natural and social sciences.	a) Practically, all students (except small schools in inaccessible locations) in the stated grades are assessed in Spanish and math. Natural science, history and geography tests are administered to 10 percent of students; b) Schools performance were ranked nationally and also in the same socio-economic category.	First administered by the Pontificia Universidad Catolica de Chile, assessment is now administered by the Ministry of Education.	Results are used extensively in policy discussions. Results also reported nationwide and also by school, location and region.
India	To help government provide baseline data on the quality of education of each state.	Grades 5 (for 2001-2); Grades 3 and 6 later. (Every 3 years).	a) Language and mathematics.	a) Sample b) Reported grade 5 scores for each state in terms of the percentage of items answered correctly	National Council of Research and Training supported by the District Institutes of Education.	To monitor changes in levels of educational achievement and to identify educational and non-educational factors that may help account for the differences.
Nepal	To determine the extent to which student achievement has changed over a 4 year period (1997-2001) during a time of major policy changes.	Grade 3 (1997, 2001); Grade 5 (1999).	a) Mathematics, Nepali and social studies.	a) A sample of 171 schools b) Scores above 75% rated "satisfactory". Reliability studies of each test and comparisons of mean scores for 1997 and 2001.	Educational and Developmental Service Centre.	To monitor changes in achievements from 1997 and 2001.
Sri Lanka	To assess the achievement of pupils who have completed grade 4.	Grade 4 (2003, 2007); Grade 3 (1996); Grade 5 (1994, 1999); Grades 8 and 10 (2005).	a) Sinhala or Tamil, English and mathematics.	a) Sample designed to be representative of the national and provincial populations of grade 4 students.	National Research and Evaluation Centre, Faculty of Education, University of Colombo.	1) For analysis of the education sector for the development of new government strategy for donor support for education; 2) Establish benchmarks for monitoring of students achievement levels in the provinces.
South Africa	1) To provide baseline data for grades 3, 6 & 9 against which future progress could be monitored. 2) To allow comparisons of curriculum achievement in math and science with those in industrial countries.	Grade 8; 1995, 1999, (2003; TIMSS).	a) Mathematics and science.	a) Sample. One intact grade 8 class was sampled in each selected school. b) Comparison of students performance in math and science with that of other countries for performance at the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles.	Human Sciences Research Council (1995, 1999). University of Pretoria (2003).	Policy makers in parliamentary debates.

South Africa	1) To provide baseline data for grades 3, 6 & 9 against which future progress could be monitored. 2) To allow comparisons of curriculum achievement in math and science with those in industrial countries.	Grade 8; 1995, 1999, (2003; TIMSS).	a) Mathematics and science.	a) Sample. One intact grade 8 class was sampled in each selected school. b) Comparison of students performance in math and science with that of other countries for performance at the 5th, 25th, 50th, 75th, and 95th percentiles.	Human Sciences Research Council (1995, 1999). University of Pretoria (2003).	Policy makers in parliamentary debates.
Uganda	1) To determine the level of pupils' achievement in literacy and numeracy, 2) Examine the relationships between achievements and gender, age, school location, and zones of the country, and also patterns of achievements, 3) Compare achievements of grades 3 and 6	Grades 3 and 6. (Every 3 years from 1996-2002; Anually since 2003).	a) English literacy and numeracy. Oral fluency in English once every 3 years.	a) A minimum of three schools per district in all fourteen national zones. b) Pupils' test scores assigned to one of four levels: determined by officials of the National Curriculum Development Centre, Teachers Colleges and Education Standards Agency as "advanced", "adequate", "basic" and "inadequate".	Uganda National Examinations Board.	1) To disseminate key lessons learned to the education community from teachers to policy makers.
United States	To 1) Measure students' educational achievements, 2) monitor changes in achievement at specified ages, 3) examine achievements of sub-populations defined by demographic characteristics and by background experiences.	Grades 4, 8 and 12. There is a separate state-level NAEP tests for Grades 4 and 8. (Every two years for math and reading) and less frequently for the others).	a) Mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history. New subject areas to be assessed are foreign language and world history.	a) Samples of grade 4 and 8 (public schools) at the state level and grade 12 at the national level. Sample size is 2500 students per state for each test. b) A separate long-term trend study national results in math and reading for age samples 9, 13 and 17 drawn from both public and private schools.	The National Assessment of Educational Progress' (NAEP) National Assessment Governing Board. Various agencies contracted to carry out aspects of NAEP.	Results are widely publicized. Sometimes results are used to score political points or criticisms about the quality of the U.S. school system.
Uruguay	To identify: 1) The level of "fundamental understanding " of language and math, 2) the socio-cultural factors that may impact student achievement, 3) provide teachers information about students' performance and help them improve teaching and performance.	Grade 6 (every three years). Grades 1, 2 & 3 (in 2001 for teacher development purposes). Grades 9 and 12 (in 1999 and 2002, respectively). 15 year olds being assessed in PISA since 2003.	a) Math (problem solving) & reading comprehension (Grade 6); Math, language, natural and social sciences (Grades 9 and 12).	a) Population and sample of grade 6 students; population of grade 9 students. Sample of grades 1, 2, 3 and 12; sample for PISA assessments. b) UMRE used 60% correct as an index of adequacy of pupil performance. Individual school scores compared to the national average, regional average and school average.	Early on as part of World Bank- Unidad de Medicion de Resultados Educativus (UMRE) financed project for Grade 6 NA. Secondary level NA was an Inter-American Bank Project. Since 2001, NA is run by Research and Assessment Division of the National Admin. For Public Education.	1) By teachers, principals and school inspectorate. 2) Used by government to identify schools for special support and for large scale, in-service teacher training programmes.
Vietnam	To measure the quality of education with a special focus on primary level student achievement.	Grade 5; previously Grades 3-5. (1998 and 2000).	a) Vietnamese reading and mathematics.	a) Sample designed to be representative of the national population and populations in each of 61 provinces b) Cross tabulation of achievement data and school data by region; correlates of achievements; and factor analyses.	Ministry of Education & Training supported by other national agencies & an international team.	Government officials made 40 policy recommendations based on the overall results.

Australia*		All public and private school students at years 3, 5, 7, and 9. (Annually).	a) Science, ICT literacy, civics and citizenship.	Each school's test results are publicized against national averages and against 60 similar schools across the nation on a socio-economic scale.	National Assessment Program – Literacy and Numeracy (NAPLAN).	NAPLAN resorts to public shaming and parental choice to generate pressure among schools for academic improvements.
Canada*	1) School achievement indicators in mathematics, reading and writing, and science (SAIP). 2) To promote jurisdictional consistency, standards and improvement.	13- and 16-year olds. (1^{st} cycle: 1993-6; 2^{nd} cycle: 1997-99; 3^{rd} cycle 2001).	a) Mathematics, reading and writing, and science.	Sample of more than 35000.	School Achievement Indicators Programme (SAIP) by the Council of Ministers of Education of Canada. SAIP replaced with Pan-Canadian Assessment Programme (PCAP) in 2007.	To reassure the public that the education system is being monitored for effective and efficient use of public funds.
Finland* (secondary school or higher)	To inform curriculum and school investments.	Grades 2 and 9. (Not indicated).	a) Not indicated b)) National – Problem-based tasks that asks student to think critically.	Sample.	Finnish education authorities.	
Japan*	Part of policy effort to improve student performance.	All public school students in Grade 5/6 and 8/9; Annually.	a) Math and Japanese English/science for Grade 5/6 and Math and Japanese English/science for Grade 8/9.	Sample of only 30 per cent of the eligible students since 2010.	Ministry of Education, Culture and Sports (MEXT).	Government releases prefecture-by-prefecture average scores intended to pressurize schools through the municipal boards of education for better performance.
Singapore* (secondary school or higher)	To guide course placement in secondary school. *Information for post-secondary education.	Grade 6. (Not indicated) *GCE O/A levels.	a) Science, mathematics English and Malay/Chinese/Tamil. b) Short and long open ended responses. * Short and long open ended responses and multiple-choice items.	No data available	The Singapore Education Assessment Board. * Cambridge International Examinations Group.	No data available
United Kingdom* (secondary school or higher)	As guidance for school-based formative and progress assessment conducted by teachers.	7- and 11- year olds. * Year 11 at 16 years old.	a) Not indicated b) Observation scales completed by teachers regarding pupils work, oral written and performance tasks & tests. *Essays and open-ended problem solutions, oral language assessments.	No data available	The Qualifications and Curriculum Authority (QCA). *Several examination groups including Oxford Cambridge, Ed Excel, the Assessment and Qualification Alliance.	No data available

Source of Content: Greany and Kellaghan (2008); except those marked with an asterix

International Assessment

International Assessments (IA) such as Trends in International Mathematics and Science Study (TIMSS), the Progress in International Reading Literacy Study (PIRLS) and the Programme for International Student Assessment (PISA) are similar to NA in terms of procedure as they are in purpose: : "a) to determine how well students are learning in the education system; (b) to identify particular strengths and weaknesses in the knowledge and skills that students have acquired; (c) to compare the achievements of subgroups in the population (for example, defined in terms of gender or location); or (d) to determine the relationship between student achievement and a variety of characteristics of the school learning environment and of homes and communities ….. e) attempt to establish whether student achievements change over time" (Greaney and Kellaghan, 2008: 61-2). One major advantage of IA over NA regarding objective is that IA actors and stakeholders namely policy makers, educators and the general public are provided with information about their system of education in relation to other systems (Postlewaite, 2004, Greaney and Kellaghan, 2008), information that sometimes put pressurize policy makers and politicians to make improvement in their services, and which may also contribute to enhanced understanding of the factors that contribute to differences in student achievement (Greaney and Kellaghan, 2008).

International Assessment tests have mostly been carried out at primary and secondary school levels. The curriculum areas that have attracted the greatest number of participants/examinees are: reading comprehension, mathematics and science. Other test areas "less taken" include writing, literature, foreign languages, civic education and computer literacy. On some occasions, national reports on country-level performance can be or have been prepared using the results of PISA, TIMSS, PIRLS and regional assessments.

Firstly administered in 2000 and conducted every three years by the Organization for Economic Cooperation and Development (OECD) the Programme for International Student Assessment (PISA) is increasingly becoming an international standard by which educational outcomes in reading, mathematics and science of 15-year-olds are measured. The primary subject of focus among the three changes every cycle. Emphasizing functional skills that students acquire by the time they complete mandatory schooling, PISA also comprises assessment of general or inter-curricular competencies such as critical thinking and problem solving. Proponents of PISA say it focuses on applying acquired knowledge in reading, mathematics and science to address issues in real-life context and that it meets the demands of many participating countries by measuring aspects including critical thinking and problem-solving skills. PISA is considered unique because the tests it develops are not in direct links with the school curriculum, albeit the context it provides through the

background questionnaires help analysts interpret the results. PISA tests are therefore designed to assess the extent of application of students' knowledge to real-life situations at the end of the compulsory education, so the students then equipped can fully participate in society.

Organised every four years targeting students in grades 4 and 8, the Trends in International Mathematics and Science Study (TIMSS) like PISA aims to "provide comparative evidence on the extent to which students have mastered official school curriculum content in mathematics and science, which is common across a range of countries .... monitor changes in achievement levels over time….monitor students' attitudes toward mathematics and science" (Greaney and Kellaghan, 2008: 64). Though both are developed by content experts from some participating countries, TIMMS' coverage is "[D]esigned to assess official curriculum organized around recognized curriculum areas common to participating countries" while PISA's coverage is "[D]esigned to cover knowledge acquired both in school and out of school, defined in terms of overarching ideas and competencies applied to personal, educational, occupational, public, and scientific situations" (Greaney and Kellaghan, 2008: 64). TIMMS' reading aspect is covered by the Progress in International Reading Literacy Study (PIRLS).

In 1999, the Bologna Process was established through a Joint Declaration of the European Ministers of Education. The goal of the process is "to develop shared standards and articulation agreements for all academic degree programs across all participating countries" (Pusateri, 2009: 7) and its priorities are: introducing the three cycle system, recognising qualifications and periods of study, and quality assurance. Drawing comparisons among three countries regarding the purpose of external quality assurance in higher education, Billing (2004) reported that there is no single external quality assurance model that applies universally though most of the elements of such a model might apply in most countries.

Problems of international comparisons of external quality assessment according to Billing (2004) include: 1) the impact of cultural differences on how 'quality' and 'level' are affected; 2) the availability of data in different forms and the wide differences in opinion on which indicators ought to be used to measure quality; 3) significant differences in fundamental elements of the structure of education systems and programmes; 4) variation in the national educational objectives; and 5) subjective use of individual systems as frames of reference.

Table 7: Advantages and Disadvantages of International Assessments

Advantages of international assessments	Disadvantages of international assessments
• Provision of comparative framework in which student achievement can be assessed and curriculum for a country to formulate methods to rectify perceived deficiencies • Inter-country result comparisons can help a country use IA results to determine what can be achieved, the way it is distributed and the existing links between mean achievement and distribution • IA can capitalise on the existing variability across education systems and hence broaden the scope of potential variables for study than can be seen occurring in an individual country • IA-based hypothesis regarding why student achievements differ among countries serves as a source of information for policy makers and researchers • A conceptualised understanding in education that might have been ignored in a country can be revealed by IAs; they can also lead to questioning assumptions considered as given • IAs usually have the penchant for attracting the attention of the media and an array of stakeholders including policy makers and politicians, academics, teachers and the public, in relation to the differences highlighted in the rankings of countries in terms of their mean level of achievement and result in a major educational policy shift • Data provided by IAs can be utilised individual countries to perform in-country analyses which, in effect, becomes an NA report • Participation in IAs is beneficial particularly for countries that lack the capacity in their universities to develop the skills for NA *(*examples available in reference)*	• Designing a procedure for assessment that is capable of measuring adequately the outcomes of different curricula is difficult – such difficulty in developing such an assessment procedure suitable for all countries increases with increasing difference between the curricula and levels of achievements of participating countries • NA, rather than IA, can test aspects of countries' curriculum that are unique to the individual countries • Since the relative effects of variables are dependent on the context they are embedded in, a high achievement-related practice in one country would not necessarily follow the same trend in another. Hence realising the variation existent in education systems for assessing the relative importance of the school resources and processes of instruction is very difficult • Strict comparisons of the samples and populations of students taking part in IA may be very difficult to achieve; same problem occurs in NAs • Since test score variance is an important factor for describing student achievements and determining correlates of achievements in the education system, national tests that are carefully designed should guarantee a relatively high test score distribution. This however is not observed among students from less industrialised countries as they find some of the items in IA difficult • Focus by media only on the results of IA rankings of countries which can be misleading for a couple of reasons can be problematic, for example ignoring the statistical importance of average differences in achievement • Poor performance in IA can carry with it significant policy and political risks especially the job security of the officers involved when the results contrary to expectations are not that great • Countries with poor human capital and infrastructure for assessment undergo significant strain to meet the demands of deadlines, etc.

(Content adapted from Greaney and Kellaghan, 2008: 66-75)

Benefits and Limitations of Standard Educational Assessment Approaches and Methods

According to the American Psychological Association Board of Educational Affairs Task Force (Pusateri, 2009), a comprehensive list of assessment strategies towards achieving learning outcomes are included in Table 8 (on the following page). Most of the advantages and disadvantages are considered general, a few however are specific to the assessment components/types. Although

earlier in this section assessment was categorised into three types by process, by activity and by scale, the categorisation in Table 8 provides another perspective.

Table 8: Assessment Strategy, Types/Description, Advantages and Disadvantages

Assessment strategy	Description/types	General advantages	General disadvantages
Course data: Used for judgment of classroom (formative) performance and approaches of more recent assessment-driven teaching and learning processes	a) Objective tests (including multiple choice, true-false, etc.), b) Essay tests, c) Embedded questions and/or assignments, d) Classroom assessment techniques (e.g., 1-minute papers, course focus groups, free-writing, etc.	a) Provides immediate feedback to students for improvement and to teachers to gauge effectiveness, b) maximises teachers' freedom to invest in students' learning, c) some develop writing skills and critical thinking, even PBL, d) inexpensive, e) comprehensive	a) Limitation mainly coming from the pedagogical constraints of the teacher, b) Generally seen as testing low level knowledge, c) Results can be affected by teacher bias
Individual Projects/ Performance-based Assessment: Students are provided with the opportunity to apply their learning to projects that enable students to optimally use their potential intrinsic interest in the subject	a) Written products (e.g., term papers, lab reports, critiques), b) New oral presentations (e.g., speeches, role plays), c) Graphic tests and displays, d) Poster presentations, e) Structural/situational assessments	a) Promotion of skill transfer and integration of content, b) A student-centred design approach that promotes investment and motivation, c) Allows for clear expression of knowledge base, d) Makes room for creativity and students benefit directly from experience, e) engages active, in-depth learning	a) Range of content that students are responsible for is narrow b) It is labour intensive and time consuming to design and implement for both students and instructors c) The challenges of variation in students' ability and motivation d) Cost may be expensive
Summative Assessment: Usually used for the purposes of evaluating program quality rather than providing feedback to students	a) Standardised tests, b) Locally-developed exams, c) Capstone experiences, d) Internships/Professional applications, e) Portfolios, f) Assessment center methods (e.g., in-baskets, guided problem-solving), g) Case or longitudinal studies	a) Provides feedback loop to improve quality, b) Facilitates comparisons over time, c) Usually cheaper to conduct that its counterpart assessments	a) Not necessarily student-centred and students often do not receive feedback on their performance, b) Some assessment options are cost and labour-intensive, c) May not reflect the gains/growth with time, d) Verifying bad performance may threaten motivation, e) Anxiety or nervousness might negatively impact actual performance
Self-Assessment/ Reflection	a) Student journals or Self-critiques	a) Reflects quality of content knowledge, b) Empowers students to practice self-evaluation, c) Enhances student transfer of accountability to other situations, d) Flexible format	a) Possible evaluative biases in self-assessment, b) Limited experience of student and hence inaccurate judgment
Collaboration	a) Research teams and group projects (e.g., written and oral), b) On-line group activities (e.g., maintaining print record of interactions in chat room or other internet-based contact)	a) Promotes participations and engagement among students including opportunities to practice group skills, b) Creates a venue where multiple subject/topic content can be synthesised	a) Students with limited training in group dynamics will have some difficulty, b) Erroneous ideas can be time consuming
Interviews and Surveys (Attitude Measurement):	a) Satisfaction measures (e.g., seniors, alumni, employers, graduate school advisors, parents), b) Performance reviews (e.g., alumni, employers, graduate school advisors), c) Exit interviews, d) Focus groups, e) Follow-up alumni interviews, f) External examiner interviews (exit interviews conducted by objective, external expert)	a) Easy to administer, b) Provides relatively quick feedback, c) Can be reliable. d) Promote more engagement, e) Generation of reinforcing feedback to sustain effectiveness, f) Participants may be beneficiaries regarding changes due to the feedback, g) External judges may be more objective in their appraisal of students' abilities/achievements	a) Possible demonstration of response bias due possession of one's own agenda, dishonesty, exaggeration, or pretence, b) May not be valid with respect to bad design, low return rates or participation, c) Labour intensive to interpret, d) Quality of the facilitator/interviewer and protocol may influence results

(Adapted from Pusateri, 2009)

Assessment Framework and Tools for 21st Century Education

Reeves (2010) argues that the traditional educational system (TES) approach to standardisation is based on the assumption of comparability among students which turn out to reward rote memorisation and sticking to the rules when all conditions are controlled. He argues that one can erroneously apply the 21CE using the TES assumption hence in principle the same results will be realised. In contrast, assessment based on 21CE assumptions does not accommodate standardisation but rather embraces variation and "volatility" which are considered a reflection on the complexity of the tasks and what happens in the real world. The 21st century skills are considered as "integral" – with regard to students learning to demonstrate proficiency in core academic standards – rather than alternative to academic content (Reeves, 2010). He argues further that secrecy of the content of a traditional test which is premised on the assumption of fairness and hence often enshrined in test procedures and sometimes codified into law end up encouraging rote learning because students are rewarded for memorisation. The 21CE assumption regarding openness of content means students being partners in an assessment process where they end up contributing meaningfully to the creation of the assessment constitutes fairness. Lastly individuals' scores which are one of the most deeply ingrained aspects of traditional assessment largely forms our social structures and norms and is based on the assumption of individual freedom to "be who you want to be", on competition, where success means beating your equals, etc. On the other hand, 21CE assumes success to be dependent on individual and collaborative effort and that leadership is not authority usurpation but rather influencing others with insight and support.

Reeves further presents five core areas – which he posits are adaptable to every academic level and subject – in the form of questions as the basis (framework) for 21st century skills assessment:

- "Explore: What did you learn beyond the limits of the lesson? What mistakes did you make and how did you learn from them?
- Create: What new ideas, knowledge and understanding can you offer?
- Learn: What do you know? What are you able to do?
- Understand: What is the evidence that you can apply learning in one domain to another
- Share: How did you use what you have learned to help a person, the class, your community or the planet?" (2010: 312-3).

Using the above three test parameters of standardisation, secrecy and individual scores and five core areas for assessment, and stating "We change from standardised to fluid assessment conditions from secret to open assessment protocols and from individual scores to a combination of individual and team scores" (2010: 319), Reeves then develops a table (matrix) of progression from the traditional assessment to the 21st century skills assessment and the accompanying implications.

After an extensive review of the literature on assessment of 21st century skills, Lai and Viering (2012) describe the common measuring tools employed: self-reports/survey, global rating scales, standard assessments and observational measures.

- Self-reports/survey: In addition to capturing skills, attitudes and dispositions of the examinee, self-reports/surveys are also used to examine personality traits, attitudes and motivation. It allows individuals to rate their own creativity-related skills, achievement, behaviours and abilities and motivation (Lai and Viering, 2012: 33). Meta-cognitive skills have also been measured using self-reports. While relatively easy and cost-effective to use, self-reports are limited in measuring for example creativity and also not very appropriate for children as it relies too heavily on verbal ability, general response and summary across a range of situations and content areas, aspects that children find difficulty in providing.
- Global Rating Scales (GRS): It is a popular method used to measure skills, particularly creativity and also motivation. An example of use of the GRS is in the form of the Personal Potential Index which is used for graduate admissions. Here, the standardised instrument "asks faculty members familiar with the candidate to rate him/her on six dimensions: knowledge, creativity, communication skills, teamwork, resilience, planning and organization, and ethics and integrity". Potential factors that can impact GRS include: "the number and types of opportunities for raters to observe the subjects, the willingness of raters to focus scores only on the criteria provided in the assessment tool, and the raters' overall understanding of the behaviours to be rated" (Lai and Viering, 2012: 35). Although less susceptible to coaching/faking as in self-reports, GRS is prone to subjectivity (Kyllonen, 2008).
- Standardised Assessments: Several of 21st century skills are assessed using standards that employ TA's multiple choice items (like critical thinking skills involving "deductive reasoning, inductive reasoning, drawing conclusions, evaluating arguments", etc. (Lai and Viering, 2012: 36)) or PBA's more open ended prompts to tap 21st century skills like the Ennis-Weir Critical Thinking Essay Test.
- Observational Measurements (OM): Dependent on location and used to catalogue students behaviour relevant to 21st century skills, OM assesses meta-cognition and motivation which are usually not measured in formal education settings. OM is also used to assess collaboration (Lai and Viering, 2012). The benefit of OM techniques is that they are based on students' actual verbal and non-verbal behaviours during engagement with tasks. Also OM tools meant for use with regular academic tasks possess the added benefit of being related to in-school learning. OM's major limitation is its heavy reliance on a professional educator

or researcher to closely observe, record and interpret student behaviours. Another limitation is that it may not be feasible for large scale testing.

Because the four types of 21st century assessment tools already discussed have their limitations, and sometimes face definitional ambiguity, Lai and Viering (2012) propose the following recommendations:

- Assessments should provide multiple measures that support triangulation of assessment approaches. To offset some of the challenges that 21st century skills assessment (e.g., limited validity for single measurements) presents, multiple measurements that represent multiple assessment modes or sample content from multiple domains so as to allow for triangulation is appropriate. Second, it is important to be clear about whether the assessment's objective is to capture for example collaborative processes (such as communication, negotiation and compromise) or products. Uncertainty and hence erring on the side of precaution is one reason it is beneficial to use multiple measures. Third, because several of the 21st century skills consist of multiple sub-components, there is the need for multiple measurements to address the issue of definitional complexity. Lastly, multiple measurements is suitable due to the indication of some degree of domain specificity which means a student may be able to exhibit a skill in one domain but fail to do so in another.
- To accurately elicit and capture 21st century skills, assessments should offer sufficient challenge and/or be of adequate complexity. The tasks should however, not be too challenging nor too easy and the level of complexity should commensurate with students age and level of cognitive ability.
- Since traditional approaches to [standardised] assessments that depend heavily on multiple-choice tests can cover and hence assess only a limited aspects of these constructs, simultaneous use of open-ended and/or ill-structured tasks – which require more meta-cognition and decision-making and provide greater examinee autonomy – will allow students to make decisions on the information they consider relevant to use to solve problems. These open-ended and/or ill-structured tasks for assessments should establish meaningful as well as genuine real-world problem contexts.
- The current implementation of most 21st century skills assessments at scale is cost ineffective and somewhat time inefficient with regard to scoring open-ended tasks. Exploring innovative approaches to address the scalability concern will hence be appropriate. Ways of addressing this challenge include use of technology enhanced formats such as tasks that use multi-media stimuli, simulations, or items that provide examinees the opportunity to record their own responses using innovative media.

Investigating Approaches for ESD Assessment
as a foundation for developing systematic M&E of ESD

ESD assessment will have to take into consideration the complexity of its nature – i.e. assessment of ESD content and learning process in addition to outcome assessment while also underpinned by the multiple dimensions of sustainable development. A number of tools described in the previous two pages currently used for assessment of 21CE could be carefully considered for use in ESD assessment either in its original or modified form especially in the formal education sector. Use of performance-based assessment techniques like interviews, journal entries, blogs, teacher observations, performance of process skills, written tests and oral and written communication have been promoted (UNESCO, 2010; UNESCO, 2012a), criterion-referenced assessments and rubrics have been suggested for use in ESD pedagogy and learning (UNESCO 2012a; UNESCO 2012c). Furthermore, self-assessment and reflective interviews were used during the global monitoring and evaluation process of the GMEF (Tilbury, 2009).

Table 9: Characteristics of Assessment within Traditional Education, 21st Century Education, and ESD Systems

	Traditional education (TES)	21st century (skills) education (21CE)	Education for sustainable development (ESD)
Content of assessment	One-size-fits-all	Personalised for individuals as well as a team/group.	Personalised, localised but with focus on overall SD context and/or team/group-work
Assessment and/or evaluation tools, frameworks, guidelines	Mainly summative assessment using traditional MCQ, pen and paper tests; Use of standards with weak formative evaluations; Promotes secrecy of information teachers nor students do not contribute to	Mainly PBA; Formative evaluations to assess knowledge and skills transparency; Self-report survey, Global Rating Scales, Standardized Assessments, Observational Measurement	Mainly formative evaluations to assess knowledge, skills and values; drawing on existing approaches and strategies that express higher order/deep learning and also those of 21CE assessments; rubrics, criterion-referenced standards
Elements for assessment/to be assessed	Knowledge (with basic skills) Mainly content (though sometimes does not necessarily meet content standards)	Knowledge and applied skills and a fair amount of values Mainly content and limited process	Knowledge, skills (basic and applied), perspectives and values (SC, WV, LR and PP) Emphasises both process and content of learning
Assessment feedback	Though faster, generally covers more breadth but weak in depth	Balanced in breadth, but more depth than TES	Strongest (makes room for needed change)
Assessors: Scorers/Raters	Not much sophistication required	More ability/sophistication needed to recognise and reward creativity, critical thinking and problem-solving	Most ability and sophistication needed

Much of the discussion done in this section covers performance assessment in the formal educational setting. Literature on assessment of education/learning in the non-formal and informal sectors of TES and 21CE is rare since by nature TES and 21CE mostly cover the formal education sector. However, since ESD covers the formal, non-formal and informal sectors it is appropriate that

assessment tools are developed for these areas as well. Table 10 compares some features of assessment in the formal, non-formal and informal learning.

Table 10: Comparing Features of Assessment in Formal, Non-formal and Informal Learning

	Formal Education	Non-formal and Informal learning
Type of assessment	Mainly dominated by quantitative assessment methods	Mainly dominated by qualitative assessment methods
Main assessor	The main assessor is the teacher (or any qualified person)	There is self-assessment and integrative assessment
Orientation of measurement or assessment	Assessment towards learning achievements that are measurable usually at the end of the topic or course	Assessment through the process, i.e., recording of individual accomplishments or development possibilities
Expression of assessment results	Expression in points or percentages in reference to standards	Assessment usually verbally expressed in a particular context
Prevalence of type of knowledge	Declarative knowledge and reproductive skills	Operative knowledge that is integrated with skills and applied in non-standard situations
Ease of assessment of learning achievement	Learning achievement easier to assess	Learning achievement involves complex assessment
Perception towards assessment	Assessment is seen as objective reality	Meaning of context is arrived at in assessment

(Modified from Table 2 of Petnuchova, 2012)

Targets and Scope of Assessment

Consideration of targets for a given M&E of ESD process, or in general any form of educational assessment, especially in relation to the desired scope, breadth and depth of the process, is an important precursor to developing an appropriate assessment/M&E system. The point of this sub-section is not to define what these targets should be as almost all options are valid depending on the needs and requirements of various educational assessment processes, rather the main aim here is to recognise that not all options can be covered from one assessment approach and thus it is necessary to consider the trade-offs between different targets in order to identify the most assessment approach. This is done by defining the *scope of the research/assessment* and by setting objectives regarding the desired usage of knowledge generation from this process.

In educational assessment, one of the key defining factors is the *desired coverage* of such assessment as this will directly impact on nature of the *breadth* and the *depth* of research. Assessment can focus on one singular case, and especially when applied at a local level or in regards to a specific course of teaching this would ideally aim for a high-level of depth on both educational implementation and performance. Many forms of educational monitoring and evaluation aim to

compare implementation across cases, and this could be between various schools or even countries. As assessment targets move towards comparison across larger units such as in multi-country comparative evaluations, there is often a natural trade-off that has to be made from research that focuses on depth to research that focuses on breadth. Elaborate, multi-method forms of cross-country assessment can still include some opportunities for inclusion of depth in the research, however as educational systems can differ considerably from country to country it is often necessary to weaken some target attributes in order to achieve adequate comparability. In ESD monitoring and evaluation, potential coverage can range from being global to local, or from looking at national implementation and curriculums to looking at individual schools or classrooms, while it is even possible to look at learning performance in regards to individual students.

A parallel defining factor is the *focus of investigation* for the research in regards to specific systems, actors or beneficiaries of educational implementation. A *systems approach*, such as reviewing progress at the level of a national education system, the national curriculum, educational policies, or the capacities of education institutions, would usually result in a focus on factors of implementation (i.e., the inputs being applied to support education or ESD). An *actor approach*, considering the education practitioners such as teachers, teacher trainers, curriculum developers, and school administrators, would provide a focus on the process of education implementation and the knowledge/expertise supporting this (i.e., the main throughputs of the system). While a *beneficiary approach*, looking at impacts of education on individual students, learning communities or even wider society, creates the strongest focus on assessing actual learning achievements and performance (i.e., the educational outputs/outcomes). It is fully possible for the focus to be multi-dimensional, but usually this remains limited to the specific bands of the overall spectrum. Strong inclusion of all three foci (system, actor or beneficiary) in regards to educational assessment generally demands a multi-method approach which may only allow for a weak synthesis of three different assessment approaches that are difficult to relate with one another.

A further important consideration to designing and structuring an M&E system or approach is the identification of the *target users* of the M&E collected data and lessons learned. Depending on the target users and the potential actions or interventions they may take to further improve ESD implementation, different types of information (and ideally correlation) are necessary. Thus, identification of the target user helps to clarify the required *types of knowledge* to be generated from the assessment process. As with other targets, it is possible to include more than one target user, however the inclusion of multiple target users may require addressing several diverse knowledge types that result in a much larger and complicated M&E process to account for the needs

of all target users. One potential option is to determine one primary target user and then a few additional secondary target users, and in this way all indicators or reporting criteria can be developed to support the needs of the primary target user while also ensuring each secondary user will be supported with valuable information for tracking general trends relevant to their main area of focus.

Two final factors that must be considered are the *process of data collection and assessment* and the desired *time scale of reporting*. These two factors are interrelated and can be addressed in parallel. It may be most appropriate to determine the desired time scale of reporting first, as this likely pre-determines several conditions about what forms of data collection and assessment will be feasible. In terms of *time scale*, there are two different aspects to be addressed: first, how regularly should reporting or M&E occur; and second, what is the desired timeliness of the reporting cycle. The first aspect, the regularity of ESD M&E refers to how often the full monitoring and evaluation process including its culmination in the evaluation and final report with recommendations for future interventions/improvements occurs. It is possible to set different time-scales for different levels of evaluation. For example some baseline information regarding system inputs could be reviewed on a quarterly basis if the information is relatively easy to both collect and evaluate (i.e. quantitative increase or decrease demonstrates likely improvements or decline); while a larger review of the process of implementation could occur on an annual basis; and finally a systematic review of learning outcomes and performance could be part of a curriculum review/evaluation process after multiple years (i.e. 3 to 5 years) of practice. The second aspect, the timeliness of the reporting cycle refers to the need for the multi-step process of data collection, compilation, evaluation, and recommendations for future interventions to occur in a timely fashion that ensures interventions are identified and implemented before problems become too serious and that the situation reflected by the collected data generally remains the same situation that is being dealt with at the time of intervention (i.e. if the full reporting cycle takes two or three years, then it would be highly likely that the data collected in year one no longer reflects the changes that have occurred in ESD implementation at the later time period when interventions are being considered two years after the original data collection).

The process of *data collection and assessment* is potentially one of the most challenging factors to truly clarify, and though it is beneficial to provide some initial outline of the desired approach it is unlikely that this will be fully detailed until the M&E system and the ESD reporting criteria (or indicators) are further clarified. Of course though, ESD Indicators are only useful if the appropriate information/data for such indicators can be collected in both a reliable and replicable manner. As

with the defining of targets, there are trade-offs that must also be considered in regards to data collection and assessment. Generally, more simple methods of data collection will of course allow for more rapid collection, but they may also result in a more superficial level of evaluation as the quality of information may be limited in providing real depth about ESD implementation. While complex systems of data collection allow for a wider, more diverse selection of reporting criteria and thus provide more depth to the quality of reporting, but they also require longer time periods for data collection, compilation, assessment and evaluation. In trying to find a middle-ground between achieving both complex reporting and timeliness, it is possible to consider a data collection process that requires many different actors in ESD implementation to respond to small, specific areas of the overall monitoring and evaluation framework. In this manner, no single actor is overburdened by the reporting process, but the total compilation of collected data provides quite a diverse amount of information (however this does not specifically lessen the burden for data compilation and assessment).

The main purpose of a *complex data collection system which includes multi-actor reporting* is to support a more holistic evaluation of ESD implementation which includes assessment across system inputs, throughputs and outputs. To effectively conduct this type of M&E though, it requires first that a stakeholder analysis is conducted to identify who are the main actors involved as administrators, implementers or beneficiaries in the ESD system and then to also consider what types of roles each actor plays in the system (which also is a useful process in regards to identifying the potential target users of the M&E outputs). It is then feasible to link specific indicators to the various roles actors play, whether this is setting appropriate policy, curriculum development, teacher training, teaching using progressive learning methodologies, or students' skill and value acquisition. For each actor or each level of reporting, different processes of reporting and data collection can be established with appropriateness to the specific types of information being collected. As already mentioned, the real burden of this process remains in terms of data compilation, assessment and evaluation. This type of multi-stakeholder data collection may be feasible to coordinate at national or sub-national levels, but it can be very tricky to establish when the objective is multi-country comparative M&E both in merely coordinating with all stakeholders across multiple countries and because the primary actors and ESD implementation system may not be completely parallel across different countries. If the international community is to gain valuable lessons from multi-country holistic M&E of ESD, then it will require substantial support from national partners in coordinating the subsequent reviews of system inputs, throughputs and outputs in their own countries.

Criteria for effective M&E methods

The core of an M&E process is defined by the approach and methods that are utilised for data collection and assessment as this determines the nature of the information that is collected and the findings that can be derived from them. There are many approaches and methods that can be used both for monitoring (reporting and data collection) and evaluation (data compilation, evaluation and recommendations) on ESD. It is not our goal to cover all possible methods for M&E of ESD, but rather to provide an overview of several of the main approaches that are representative of the wide spectrum of M&E methods. However, before discussing the individual methods, it is important to recognise a few of the key criteria that need to be met by any selected methods for M&E of ESD. These are *validity*, *reliability (*or *replicability)*, *verification*, and *comparability* (see table 11 for summary of criteria).

Validity in M&E refers to both if the indicator is providing information appropriate to the intended target of assessment and also in regards to subjective reporting criteria how accurately the information reported relates to the real world situation (or how trustworthy is the information being provided). Overall validity can be improved through a process of triangulation, ensuring that multiple indicators broadly cover the same areas of focus, or by applying both quantitative and qualitative reporting criteria in parallel to one another. Triangulation can also utilise multi-stakeholder review and assessment to eliminate natural biases of a single investigator. However, it is important to note that the concepts of both reliability and validity have historically developed in accordance to the positivist tradition and are usually postulated in regards to quantitative research methodology (Golafshani, 2003: 597-600). Although M&E of ESD usually aims to include quantitative indicators, since the nature of ESD reflects more on aspects of qualitative improvements to educational approaches, it is valuable to also understand the nature of validity and reliability in qualitative research. Validity in quantitative research is concerned with the tools or methods of measurement, and if they accurately measure what is the intended target and support generalisability (Golafshani, 2003: 599). In contrast, qualitative research intends to produce a different knowledge type based on a deeper understanding of the nuances of a given process or system, and although measurability may not be a relevant aspect it is still possible to address validity in regards to the rigor, quality, trustworthiness and transferability of the generated knowledge (Golafshani, 2003: 600-3).

Both validity and reliability are research criteria that have mainly been structured to fit the nature of quantitative research and draw heavily on the positivist tradition aimed at measuring and testing hypothesis, but as mentioned evaluating ESD performance requires a perspective more familiar to qualitative research which, "… uses a naturalistic approach that seeks to understand phenomena in

context-specific settings" (Golafshani, 2003: 600). Golafshani continues by explaining a distinction between quantitative and qualitative research in relation to their specific aims; while quantitative research aims to provide "causal determination, prediction and generalisation of findings", qualitative research aims for "illumination, understanding, and extrapolation to similar situations" (2003: 600).

> Furthermore, in conducting evaluations, quantitative methods are best suited to measuring levels and changes in impacts and to drawing inferences from observed statistical relations between those impacts and other covariates. They are less effective, however, in understanding *process*—that is, the mechanisms by which particular intervention instigates a series of events that ultimately result in the observed impact (Rao and Woolcock, 2004: 167).

This is not to argue that M&E of ESD should only focus on qualitative assessment, but rather to suggest that both quantitative and qualitative assessment serve distinct and valuable purposes. Quantitative measurements can demonstrate change and support comparability, however when the quantitative statistics demonstrates a worrisome trend or movement in the wrong direction it is often necessary to turn to the qualitative data to identify what area of the process needs to be addressed and what would be an effective intervention in countering the undesirable trend. However, if we are to consider the inclusion of qualitative assessment then it is necessary to consider how research criteria are defined when extended beyond the positivist tradition of quantitative research, especially in regards to validity and reliability.

Reliability is usually referred to as the replicability of a measurement in quantitative research, or the ability to apply the same tool or method of measurement repeatedly and receive the same finding. Again, as qualitative research may not fully depend on measurability in a statistical format, a different consideration is necessary if this concept is to be effectively applied to qualitative research. Stenbacka (2001) argues that reliability as a criteria of measurement should not apply to qualitative research, and explains that while quantitative research is evaluated on its "purpose of explaining", qualitative research should be evaluated on its ability of "generating understanding" (as cited in Golafshani, 2003: 601). In qualitative research, the reliability of inquiry depends on a strong process of *verification* and as such is quite relevant for M&E of ESD. *"Verification* is the process of checking, confirming, making sure, and being certain. In qualitative research, verification refers to the mechanisms used during the process of research to incrementally contribute to ensuring reliability and validity, and, thus, the rigor of a study" (Morse, et. al., 2002: 17). In establishing ESD indicators or reporting criteria for M&E, considerations of how verification mechanisms will be built into the system are quite crucial in establishing reliability from the collected data. Three approaches for strengthening verification are utilising data collection from multiple sources, peer review of reported

data, and multiple overlapping of reporting criteria (i.e. asking different questions in regards to same indicator/criteria).

An additional criteria that is especially important for M&E of ESD is *comparability* which allows the reported data of two or more similar cases to be compared with one another and is usually accomplished by ensuring that the information that is reported is in a standardised (and most often numerical) format where notable differences between two cases clearly represent something significant occurring (or not occurring) in practice. When conducting M&E processes that require inputs from multiple respondents, not only is it necessary to address the reported data but also the way in which research questions or ESD indicators are interpreted in the first place which influences the information the respondent considers important. Comparability must be addressed when developing the M&E approach and methods for it to be possible to achieve during evaluation. The purpose of comparability must be considered though, as comparability is often used as a precursor to ranking. However in M&E of ESD, ranking may be neither a desirable or legitimately achievable goal. A more desirable purpose of comparability is as a precursor to achieving *transferability*, i.e. the ability to recognise where an intervention that was successful in one context would also be relevant in additional cases.

Although statistical data provides easy comparability, in multi-country comparative evaluations it is also important to recognise the potential limitations such data offers in identifying the multivariate factors of the complex systems across different countries which may lead to wide discrepancies in statistical values. “While statistical control allows investigators to make broad statements with relatively little data, these broad statements are possible only because very powerful simplifying assumptions have been made” (Ragin, 1987: 64). Qualitative data generally provides more depth for understanding the nuances of ESD processes and practice, but even with a standardised system for reporting comparability can be difficult with qualitative data. A mixed-methods approach to M&E can be one of the more effective means to achieve comparability while also achieving an inductive explanation of processes and motivations. “The key is to find measurement methods that are flexible enough to capture organisational complexities and differences, yet specific enough to be calculable and comparable” (Shriberg, 2002: 256).

Table 11: Main Research Criteria relevant for systematic M&E of ESD

	Quantitative Perspective	Qualitative Perspective
General Perspective	Following the positivist tradition, aims to provide "causal determination, prediction and generalisation of findings".[a]	Following the naturalistic approach, aims to provide "illumination, understanding, and extrapolation to similar situations".[b]
Role in Educational Assessment	Statistical measurements that can demonstrate change (over time) and support comparability while also identifying positive and negative trends.	In-depth knowledge and understanding about the nuances of a given process or system which supports identification of points for effective intervention.
Main Research (or Assessment) Criteria		
Validity	Is the information measured accurate to the intended target, and does it support generalisation of findings? Main concern on the tools or methods of measurement.	Does the generated knowledge provide rigor, quality, trustworthiness and transferability? Main concern is with cross-checking and methodological triangulation.
Reliability (& replicability)	Does the repeated application of the same tool or method of measurement produce the same finding, i.e. replicability?	Reliability is closely linked to verification, and it concerns the ability of research for "generating understanding".[c]
Verification	Less significant to quantitative research, as the process of replicability serves as the verification mechanism.	Confirming research findings through incremental checking. In M&E, this is strengthened through data collection from multiple sources, peer review of reported data, and overlapping of reporting criteria.
Comparability	Producing numerically comparable data from two or more cases through standardisation of data.	Achieved less by numerical comparability and more by *transferability* of findings with value placed on inductive explanation of processes and motivations.

[a & b:] Quoted from Golafshani (2003: 600).
[c:] Quoted from Stenbacka (2001) as cited in Golafshani (2003: 601).

Specific Approaches/Methods for Monitoring and Evaluation in ESD

Summative Assessment Approaches - Standardised Testing and Performance Based Assessment

Approaches to standardised testing (as a form of traditional assessment) and performance-based assessment (PBA) have already been discussed earlier in this section, but need to be readdressed for their value in M&E of ESD. Performance-based assessment provides an ideal approach for evaluating if ESD is achieving its desired aims in terms of learning performance, although at this point it would most likely be very difficult to implement for systematic multi-country M&E of ESD. However, in individual classrooms, school systems or even at a national level, the implementation of ESD performance-based assessment would be feasible at least in regards to formal education. Summative assessment approaches represent the mainstay of assessment in the field of education because they provide the clearest way for identifying what students have actually learned. The drawback to only utilising summative assessment though is that even if learning deficiencies are identified, there is no additional information about the process of implementation to support

appropriate targeting of system interventions. Thus, for systematic M&E of ESD implementation and performance, a mixed-methods approach to assessment would provide the most useful coverage of system inputs, throughputs and outputs.

It has also already been noted that ESD ideally includes several unique learning features based around learning skills (i.e. critical problem solving, systems thinking, lifelong learning skills, etc.) and values (i.e. citizenship, stewardship, cooperation, empathy, etc.) that generally speaking have been relatively difficult to measure in summative assessment approaches. PBA provides a clear attempt to move beyond the limitations of only assessing knowledge-based learning as standard in TA approaches by including practice oriented assessment to demonstrate skill-based learning, however there are still apparent limitations in assessing value-based learning.[2] The approaches to diagnostic assessment in the education field to gain a baseline of existing knowledge, skills, and values in order to better develop lesson plans in line with learners' perceived needs provides the clearest route for assessing value-based learning. The approaches to diagnostic assessment usually include open discussions and interviews with students, but may also utilise observational studies or preferred response surveys and self-reporting inventories.

It is also worthwhile to make a distinction between summative assessment which is conducted immediately following a course education thus assessing short term learning gains and follow up assessment conducted well after completion of the course which aims to identify long term behaviour changes. As real proof of behaviour change may not materialise over a short term, some type of subsequent or follow up assessment may be required to identify if new knowledge and values have actually led to changing individuals' behaviour. Otherwise, there is a chance of duplicating the flawed approach of traditional behaviour change theories which based on a perspective of individual bounded rationality assumed that more environmental awareness would lead to an increase in pro-environmental attitudes and subsequently on to an increase in pro-environmental behaviour. Numerous studies have now demonstrated the inherent inconsistencies beyond theory and real-world practice, and in turn this has led to the identification of the attitude-behaviour or value-action gap (Blake, 1999; Kollmuss and Agyeman, 2002; Jackson, 2005).

System Inputs and Capacity Assessment

While performance-based assessment focusses on the outputs of ESD, it is also possible – at least in part – to assess ESD implementation and practice in regards to the inputs being made into the system. In fact, M&E of system inputs (and to a lesser extent, system throughputs) is one of the

[2] The ESDinds Project and the subsequent We Value online platform are working to develop an innovative approach to value-based learning and assessment for ESD; see: http://www.esdinds.eu/ and http://www.wevalue.org/index.php

more effective means for multi-country, comparative assessment of ESD. However, on its own the limitation of input assessments is that there is no clear connection or even potential understanding of what are the actual ESD outputs/outcomes and learning performance. The application of input assessments though in conjunction with other forms of assessment can be an important approach for effective M&E of ESD. While summative assessments help to identify the positive and negative trends in ESD outcomes thus highlighting the issues that need to be addressed by interventions, an input analysis is more effective at identifying the appropriate points for intervention within the system. As systematic M&E of ESD becomes more commonplace, one important area for future research and documentation is demonstrating correlations between inputs and outputs. If the impact of given intervention can be assessed during the M&E process and especially if it possible to determine which interventions work best for addressing specific issues, then such a M&E process will provide meaningful direction for ESD improvements in general. Various forms of input assessments include service availability assessments, gap analysis, system mapping, SWOT analysis and several of the tools developed for the Participatory Rural/Rapid Appraisal approach.

One form of input assessment that is quite effective is a capacity assessment which is a type of institutional assessment that aims to evaluate the functionality of an institution or organisation by assessing its available resources (including financial, human and knowledge), its organisational structure, its leadership, etc, thus covering the major institutional inputs required for effective implementation. A capacity analysis actually addresses factors relevant to system inputs, throughputs and outputs, but is discussed here as its focus is most relevant to the nature of input assessments. Several methods for capacity assessment have been prepared by various development/aid organisations to strengthen the effectiveness of the support and capacity development they provide (i.e. World Bank, UNDP, DFID, USAID, etc.). UNDP suggests that effective capacity assessment should support development to, “i) understand what constitutes a starting point (how to articulate what capacities are there to begin with); ii) uncover where the hurdles to developing capacity are and design programmatic responses that will actually address those hurdles to drive improvement; and iii) most important, measure the change in an institution’s capacity to fulfil its mandate and provide insight into where to make investments for continuing improvement” (UNDP, 2010a: 2).

It was mentioned that capacity assessments may also consider the system throughputs that underpin ESD implementation. If we define the input capacities as the institutional arrangements, policy mandates and resource capacities that frame ESD implementation, then we can also define the throughput capacities as those that facilitate effective ESD practice, i.e. the leadership,

knowledge, pedagogies and methodologies supporting ESD. These throughput capacities are equally important to account for in assessment as the input capacities as both shape and influence the overall quality of ESD achievement. There are multiple ways to assess existing throughput capacities. Quantitative assessment can be applied to things like the number of ESD related courses provided by teacher education institutes and mandatory requirements for such courses, the number of current teachers who have received in-service training on ESD, and the number of ESD related teaching materials made available to teachers. A capacity assessment may also examine if good structures for knowledge sharing and research supply-demand are in place. While qualitative assessment can be applied to examine the conceptual framing of ESD in a given system by examining the descriptive language utilised to explain ESD teaching and objectives (for example, see the characteristics of ESD identified in Table 16 in Section 3).

Case Study

As an assessment method, case studies can be a valuable way to generate in depth and highly insightful information about actual means for strengthening ESD outcomes and impact. But due to this methodology's limitations for supporting multi-subject comparison and also several misunderstanding about its benefits, the case study method is often unfairly discredited in regards to its potential application as an assessment tool. Flyvbjerg (2006, 2011) addresses five of the common misunderstandings about case studies. Drawing on his previous work on *phronetic* social science (2001), he explains that, "the case study produces the type of context-dependent knowledge that research on learning shows to be necessary to allow people to develop from rule-based beginners to virtuoso experts"(2006: 221). Based on a phenomenology of human learning (i.e. the Dreyfus model based on a five-level human learning process), Flyvbjerg demonstrates how at first three learning levels the knowledge that the learner initially relies on are based on context-independent rules, while progressively learners move towards a greater reliance on knowledge generated from rich, context-dependent practical experience and critical reflection. The qualitative jump between the first three levels and later two levels of learning, "...implies an abandonment of rule-based thinking as the most important basis for action, and its replacement by context and intuition. Logically based action is replaced by experientially based action" (Flyvbjerg, 2001: 21). Flyvbjerg continues his argument for a *phronetic* social science by explaining that social sciences attempt to emulate the natural sciences and the epistemological tradition which focuses on context-independent models and theories is misguided as social life/phenomena is inherently context relevant and these approaches only address knowledge that is relevant at beginner and early stages of learning rather than those relevant to expert practitioners (Hargreaves, 2012: 315-6).

Phronetic social science is thus proposed as a new, or reconceptualised, approach to both the philosophy and practice of the social sciences. This new approach aims to strengthen practice in daily lives by orienting social science in a way which "effectively deals with public deliberation and praxis" (Flyvbjerg, 2001: 129), and the case study provides a valuable method for establishing rich narratives of practice in context-dependent settings and the "systematic production of exemplars" (Flyvbjerg, 2001: 87). This last point is especially important for emerging fields such as ESD, for if the desire is to develop a clear understanding of how to systematically enhance ESD learning performance then it is necessary to first select for those extreme examples of best practice and then to intimately investigate the important characteristics and success factors of these cases which may only be elucidated in the minutiae of actual practice. In first selecting case studies as an assessment method for M&E of ESD, it is important that the desired M&E goal and expectations correlate with both the benefits and limitations of the case study method. While Flyvbjerg (2006. 2011) corrects the common misunderstanding that it is not possible to make generalisations and develop general propositions based on case studies, it is still important to recognise that case studies – especially in regards to the case process rather than case outcomes – can be difficult to summarise as a limited number of main results (Flyvbjerg, 2006: 238). Thus, it can also be challenging to create simple comparisons of levels of practice and implementation between cases with the application of this assessment method alone.

On the positive side, if the goals for M&E are to identify means for improving ESD practice and performance, then case studies are an appropriate means for accessing the type of expert, context-dependent knowledge that is generated through critical praxis (i.e., practical experience and critical reflection). In a mixed-methods assessment approach, case studies can fill an important knowledge niche that strongly compliments the other knowledge types generated by more comparative and quantitative methods of assessment. Through the case study narrative, it is also possible to gain a real feel for what ESD is all about in a way that is not possible from just looking at ESD indicators and in the instance of exemplar cases to be inspired by innovative practices.

Benchmarking and Goal Setting

The establishment of clear objectives and goals at the initial outset of a project is a valuable means for facilitating future M&E. In international initiatives such as the Education for All (EFA) movement or the Millennium Development Goals (MDGs), these clear goals provide achievement targets for international partners to work towards, and since these targets are in place it is then also appropriate to conduct M&E based on progress made towards meeting these goals. Once such goals are established, it is also possible to then conduct an initial diagnostic assessment to establish a

baseline of where countries or organisations are at prior to starting the initiative. Thus, it is possible to not only evaluate how close countries or organisations are to achieving the goals but also to measure the progress that is made towards these goals over time. Generally though, the types of goals appropriate for this type of benchmarking and M&E need to have clear quantitative targets that are relatively easy to measure and assess.

Such goal setting however was purposefully avoided during the establishment of the DESD agenda and programme. The original goal for DESD as stated in UN General Assembly resolutions 59/237 was for, "Governments to consider the inclusion … of measures to implement the Decade in their respective education systems and strategies and, where appropriate, national development plans". And in the International Implementation Scheme for DESD, UNESCO also defines two sub-goals:

- Provide an opportunity for refining and promoting the vision of and transition to sustainable development – through all forms of education, public awareness and training.
- Give an enhanced profile to the important role of education and learning in sustainable development (UNESCO, 2005a: 6).

Even these sub-goals though only provide very broad, conceptual approaches. This document does provide four additional objectives for DESD that helps to move closer towards actual implementation targets, but even these remain rather open-ended and are directed more to how UNESCO as the lead implementer of the decade should provide support to countries on ESD; i.e. 1) facilitate ESD networks and stakeholder interaction, 2) improve quality of ESD teaching/learning, 3) support attainment of MDGs through ESD efforts, and 4) provide opportunities for educational reform through ESD (UNESCO, 2005a: 6).

In order to understand why DESD goals and objectives remained so underdeveloped at the outset of the decade, it is important to recognise how DESD was linked to other international initiatives started around the same time. These specifically include the MDG process, the EFA movement, and the United Nations Literacy Decade (UNLD) which all provide "tangible and measurable" goals for quantitative educational improvements. ESD was positioned into these efforts for global educational improvement as a means for enhancing qualitative reforms to education systems to enable learners to better meet the needs and the challenges of the new millennium. As such, "DESD promotes a set of underlying values, relational processes and behavioural outcomes, which should characterize learning in all circumstances" (UNESCO, 2005a: 9). ESD thus was not to propose an additional set of quantitative targets on top of those already provided by the other initiatives. Furthermore, ESD was not to be presented as a dogmatic concept, but rather countries were to develop their own contextualisation of ESD as it fit with their countries' development and education needs and objectives. This however caused a serious lack of understanding and direction about ESD for many

countries during the early years of the decade, and as already pointed out it has also meant that there are no clear international targets for ESD from which progress can be monitored and evaluated.

What are Indicators?

"An indicator is something that helps you understand where you are, which way you are going and how far you are from where you want to be… A good indicator alerts you to a problem before it gets too bad and helps you recognise what needs to be done to fix the problem" (Hart, 2000). The goal of using indicators should be *action* – to understand the trends that are occurring, to improve those going in the wrong direction, and to protect what is valued (MacGillivray, 2000: 81). In conducting monitoring and evaluation, there is a natural bias towards quantitative indicators that can provide numerical measurements, which is understandable as such indicators lend themselves to statistical analysis and comparison. The following quote from the UNAIDS programme provides an example of this common understanding, "In the context of monitoring and evaluation, an indicator is a quantitative metric that provides information to monitor performance, measure achievement and determine accountability" (2010: 14). Nonetheless, the value of qualitative indicators should not be overlooked, especially in regards to M&E of ESD which needs to account for both quantitative and qualitative improvements in educational practice.

The individual value of using quantitative or qualitative indicators and assessment approaches are respectively linked to the positivistic and interpretative traditions from which they developed (Giddens, 1976). Quantitative indicators provide *objective* facts, while qualitative indicators help to explain the *subjective* nature of social phenomenon. Even when selecting and weighting quantitative indicators though, there is a reality of subjective decision-making (Dahl, 1997 and Gallopin, 1997).

While quantitative indicators may be desirable, they are only relevant for describing certain aspects of reality. Quantitative indicators address targets that can be measured in numerical terms. They are often easier to collect relevant data for than qualitative indicators, and they are also easier to evaluate. Quantitative indicators can tell you how much and how often something is being done, but if you want to know the specifics of what is being done and how it is accomplished then it is often necessary to also utilise some form of qualitative indicator. If quantitative data can accurately represent the desired targets, then for ease of measurement and analysis it makes sense that quantitative indicators would be selected over qualitative indicators. For an example of several standard quantifiable indicators or statistics on both education and development in the selected countries see Tables 13 and 14 respectively; on reviewing these statistics, also consider how well a

short list of similar quantifiable indicators on ESD would actually inform about the quality of ESD implementation, practice and achievement that is occurring in a country.

Qualitative indicators can meaningfully address aspects of ESD that are quite significant to know, i.e. value learning, knowledge frameworks shaping ESD instruction, pedagogical approaches, ESD as a stimulus of wider educational reform, etc. "[I]n 'difficult to measure' areas, such as influencing value systems, these creative qualitative indicators have the potential to provide deep insight. And, rather than holding project management to account for their ability to predict numbers we can assess whether, at the end of the project, people are expressing the type of values that it was the intention of the project to bring about" (Barnes, 2010: 4). Furthermore, qualitative indicators can produce numerical information through the use of attitude scales (such as the Likert-type scale), coding and ranking. As an example, in the country ESD status survey utilised during this research project, an initial identification of the core concepts to be addressed in ESD was made and based on this a question asked the respondents to mark all of the concepts that were being addressed. From the responses, it was possible to then create a ranking of how well or how holistically each country addressed ESD. Furthermore, from the qualitative data it was also to identify which areas had very strong or weak coverage across all of the countries.

The application of a mix of both qualitative and quantitative indicators provides a good solution for covering both the how and the how much of ESD implementation. Although, the application of both indicator types will inevitably result in a slightly longer list, this approach is also completely compatible with the mixed-methods assessment approach advocated in this work. Another division of types of indicators is relation to level of the system/process they are targeting, and just as in assessment in general this applies to system inputs, throughputs, and outputs (this type of division will be explained further in Section3). Longer term indicators may also look beyond outputs and instead at outcomes (i.e. how does learning impact on long-term behaviour and on social change).

Finally, in selecting indicators, it is important that each potential indicator is reviewed and tested for its appropriateness. Maureen Hart (2010) provides an example of criteria that can be used for judging potential indicators:

- Effective indicators are **relevant**; they show you something about the system that you need to know.
- Effective indicators are **easy to understand**, even by people who are not experts.
- Effective indicators are **reliable**; you can trust the information that the indicator is providing.
- Lastly, effective indicators are based on **accessible data**; the information is available or can be gathered while there is still time to act.

In this research project, the commonly applied SMART criteria (see Table 12) were selected as the means for testing the feasibility of potential indicators. A further criterion for selecting indicators should be policy relevance. Policy makers of course like easy to interpret indicators, or clear numbers they can highlight to show that things are improving. At the same time though, when policy makers are considering how to make improvements in a system, such quantitative data alone may not be informative enough to identify appropriate interventions.

Table 12: SMART Criteria for Indicators

S – specific	M – measurable	A – achievable & attainable	R – realistic & relevant	T – timely (or time-bound)
Is the indicator clear and well-defined? Will it provide precise information?	**What information is needed to confirm the indicator, and is it unambiguous? How would this information be evaluated?**	**Is it possible to collect the information for the indicator in a straightforward and reliable manner?**	**Is it feasible and practical to demonstrate achievement or progress in regards to this indicator?**	**Will the indicator highlight necessary actions and interventions in time to act on it?**

Table 13: General Education Statistics from Target Countries

		SCHOOL AGE POPULATION[1] (thousands)	SCHOOL LIFE EXPECTANCY[2] (years)		NET ENROLMENT RATIO[3] (% of respective school-aged children)		GENDER PARITY INDEX[4]		PUPIL-TEACHER RATIO[5]		ADULT LITERACY RATE[6] (% of population age 15 and above)	PUBLIC EXPENDITURE ON EDUCATION[7]		PUBLIC EXPENDITURE PER PUPIL[8] (2005 PPP Dollars)	
			Girls	Boys	Primary	Secondary	Primary	Secondary	Primary	Secondary		% of GDP	% of Government Expenditure	Primary	Secondary
Northeast (NE) Asia Countries	China	274,169	11.4	11.4	98.0%	n/a	n/a	n/a	17.7	16.4	93.3%	n/a	n/a	n/a	633.77
	Japan	17,057	14.8	15.1	99.8%	98.2%	n/a	1.00	18.5	12.2	99.0%*	3.5%	9.5%	6,989.45	7,052.42
	Republic of Korea	8,379	15.7	18.0	97.6%	96.9%	0.94	0.94	25.6	18.1	97.9%**	4.4%	15.3%	4,793.62	5,966.53
Southeast (SE) Asia Countries	Cambodia	4,966	9.2	10.4	89.4%	34.1%	0.96	0.88	50.9	28.9	76.3%	1.6%	12.4%	98.56	112.64
	Indonesia	62,293	12.2	12.5	94.8%	67.5%	0.96	1.01	18.8	13.0	92.0%	2.5%	17.5%	n/a	n/a
	Malaysia	8,104	13.1	12.4	99.9%	68.7%	1.00	1.10	16.9	17.0	91.9%	4.6%	25.2%	1,839.46	2,667.22
	Philippines	30,990	12.1	11.5	91.3%	61.3%	1.02	1.20	33.7	35.1	93.4%	2.5%	15.2%	278.98	295.20
	Thailand	14,825	14.5	13.4	93.9%	76.1%	n/a	1.12	17.7	21.0	94.1%	3.9%	20.9%	982.56	1,082.24
	Viet Nam	23,080	9.9	10.7	94.5%	61.0%	n/a	n/a	20.4	21.8	90.3%	5.3%***	n/a	n/a	n/a

Notes:

General: All figures in this table are based information from UN ESCAP (2010) *Statistical Yearbook for Asia and the Pacific 2009.*

1 Calculated from ESCAP (2010) *Statistical Yearbook 2009* figures [Total Population] x [Proportion of Children as percentage of total poulation]; figures come from 2008; However, children are calculated as those aged 0-14, so there is some minor error in this as the ideal calculation would be from 4-16.

2 Figures come from 2007; except for Malaysia from 2005, Philippines from 2006, and Viet Nam from 2000.

3 Figures come from 2007; except for China from 1991, Korea for primary from 2005, Malaysia from 2005, and Viet Nam from 2000.

4 Figures come from 2007; except for Korea for primary from 2000, Malaysia for primary from 2006 and for secondary from 2005.

5 Figures come from 2007; except for Malaysia from 2006.

6 Figures come from 2007; except for Indonesia from 2006, and Viet Nam from 1999.

7 Figures come from 2007; except for Japan from 2006, Malaysia from 2006 and 2004, and Philippines from 2005.

8 Calculated from ESCAP (2010) Statistical Yearbook 2009 figures [GDP per capita (2005 PPP dollars)] x [Public Expenditure per Pupil as percentage of GDP per capita]; there is a minor error in this calculation as GDP per capita is calculated from 2008 statistic, while public expenditure per pupil is calculated from earlier statistics. Public expenditure per pupil for China is from 1999, for Japan from 2005, for Korea from 2005, for Cambodia from 2004 and 2001, for Malaysia from 2004, for Philippines from 2005, and for Thailand from 2004.

Table 14: General Development Statistics from Target Countries

		Human Development Index[1]		HDI Development Status[2]	HDI Education Index[3]		Political Engagement[4] (% of people who voiced opinion to public officials)	Satisfaction with Freedom of Choice[5] (total % satisfied)	Democracy Index[6]		Political Participation category of Democracy Index[7]	Ecological Footprint[8] (hectares per capita)
		Score	Global Rank		Score	Global Rank			Score	Global Rank	Score	
Northeast (NE) Asia Countries	China	0.663	89	Medium	0.851	97	n/a	70%	3.14	136	3.89	2.2
	Japan	0.884	11	Very High	0.949	34	22%	70%	8.08	22	6.11	4.7
	Republic of Korea	0.877	12	Very High	0.988	8	22%	55%	8.11	20	7.22	4.9
Southeast (SE) Asia Countries	Cambodia	0.494	124	Medium	0.704	132	14%	93%	4.87	100	2.78	1.0
	Indonesia	0.600	108	Medium	0.840	102	11%	75%	6.53	60	5.56	1.2
	Malaysia	0.744	57	High	0.851	96	11%	83%	6.19	71	5.56	4.9
	Philippines	0.638	97	Medium	0.888	73	24%	87%	6.12	74	5.00	1.3
	Thailand	0.654	92	Medium	0.888	72	29%	84%	6.55	57	5.56	2.4
	Viet Nam	0.572	113	Medium	0.810	113	16%	73%	2.94	140	3.33	1.4

Notes:

1 Total possible score out of 1.0; Data from 2010, published in UNDP's Human Development Report 2010; pp.143-7

2 Data from 2010, published in UNDP's Human Development Report 2010; pp.143-7

3 Total possible score out of 1.0; Data from 2007, published in UNDP's *Human Development Report 2009* ; http://hdrstats.undp.org/en/indicators/93.html

4 Data from 2008, published in UNDP's *Human Development Report 2010* ; pp.164-7

5 Data from 2009, published in UNDP's Human Development Report 2010; pp.164-7

6 Total possible score out of 10.0; classifications of political system follow scoring criteria 10-8 = full democracy, 7.99-6 = flawed democracies, 5.99-4 = hybrid regimes, 3.99-0 = authoritarian regimes; Data from 2010, published in the Economist Intelligence Unit's *Democracy Index 2010*

7 Total possible score out of 10.0; Data from 2010, published in the Economist Intelligence Unit's *Democracy Index 2010*

8 Data from 2010, published in the Global Footprint Network's *Ecological Footprint Atlas 2010*

SECTION 3

Scoping Research on the Potentials for Monitoring and Evaluation of ESD in the Asia-Pacific Region:

A mixed-methods approach to elaborating key leverage points and learning performance characteristics in ESD implementation

Structure and Methodology of Scoping Research on ESD Monitoring and Evaluation

As the initial focus of this research project was oriented to provide broad scoping research, the desire was to identify both the important capacities for effective ESD implementation and the specific factors/ characteristics that support strong learning outcomes in the Asia-Pacific region. The expectation was that such findings would then lend themselves to development of an appropriate framework for the monitoring and evaluation of ESD practice and performance in the region, and this in turn would create the necessary structure for elaborating good ESD indicators. This section reviews both the main processes and major findings from the multiple research approaches utilised during this project. It then attempts to provide a synthesis of these findings and elaborate a clear monitoring and evaluation framework for future application in the region. The application of this M&E of ESD framework is also reviewed in relation to the criteria and approaches/methods discussed in the previous section. The final process of this research phase was a review of the findings presented in this section during an expert consultation that led to further refinement of the purpose and approach for future monitoring and evaluation in the region (which is presented in an addendum to this section).

The research was grounded in the traditions of pragmatism and critical praxis aiming to create clear linkages between theory and practice. To support this, a mixed-methods research strategy based on a mixed approach design was prepared to draw on multiple forms of information and knowledge types regarding ESD implementation, practice and performance. The mixed-methods research also allowed for a strong process of triangulation to further clarify and validate the research findings.

The research for this project was initially structured to investigate two broad assessment approaches. For the first research approach a quantitative (or semi-quantitative) approach was applied with the purpose to work towards the identification of specific factors, capacities and leverage points that support the effective implementation of ESD.[1] The main method of data collection under this approach was the use of a structured survey. Prior to development of the survey though, an evaluation framework was established (see appendix A) through review of relevant strategies for implementing ESD and a consultation process with international ESD experts. The initial framework was structured around both a diverse set of six sectors for ESD implementation including: national curriculum, formal education, teacher training, non-formal education, community & civil society, and private sector as well as a variety of capacities including the major inputs, throughputs and outputs of each sector for ESD implementation and practice.

[1] The main findings from this approach are published in Didham and Ofei-Manu (2012) *Education for Sustainable Development Country Status Reports.* Hayama, Japan: IGES & UNU-IAS.
Available at: http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=4140

This framework provided the basis for conducting scoping research on what are relevant factors and capacities to be addressed in ESD monitoring and evaluation, thus the survey was prepared to cover all aspects of this framework. Furthermore, multiple types of responses to the different questions were utilised to establish a greater level of understanding regarding ESD implementation in each country (i.e., quantified answers and ranking, general positive or negative responses, selection of items from lists/bullet points, and open ended responses for qualitative and/or detailed explanation). For completion of these surveys, partners in seven countries were identified to perform the role of national focal points for M&E of ESD. In most cases, the focal points then coordinated a multi-stakeholder response process with relevant agencies and officers to ensure appropriate reporting across the range of sectors covered by the survey. These surveys provided the main data, along with secondary supporting documents, to produce ESD country status reports for the seven countries. A comparative assessment of the ESD implementation status in these countries was then conducted to draw out the important factors and capacities in ESD implementation.

The second main assessment approach was qualitative in nature and was based on a series of case studies from selected "good-practice" cases.[2] While the first research approach focused on the quantity and modalities of implementation for ESD, it is equally important to this research to understand the impacts and outcomes of the types of ESD being implemented. The second approach was thus structured to provide an investigation of qualitative achievements for ESD and to consider both the educational contents and learning processes that support effective learning performance.

A case-study reporting framework was established to facilitate comparability among the collected case data, and this was based primarily on open-ended questions. Data from the case studies was coded in regards to the various education/learning processes and contents that were addressed by individual cases, and the data was assessed through theoretical sampling, comparative assessment, and analytical induction. Reflexive testing was also applied to the analysis of learning performance factors to allow for a steady interplay between theory and practice, i.e. through an action-reflection cycle.

To strengthen the ability for methodological triangulation in this research, a complementary third approach was also applied during the reporting and capacity building workshops in the form of cooperative inquiry. Data was gathered through presentations and focus group activities during these two sub-regional workshops. Two additional expert consultations following the completion of the primary research allowed for the findings from the two initial research approaches to be reviewed. Further group activities were facilitated to consider the needs and characteristics of future ESD M&E and to move towards identifying an appropriate ESD M&E reporting criteria (which could form the basis for developing actual ESD indicators).

[2] The main findings from this approach are published in Ofei-Manu and Didham (2012) *Assessment of Learning Performance in Education for Sustainable Development.* Hayama, Japan: IGES & UNU-IAS.
Available at http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=4172

Research Approach 1: Investigating ESD Implementation

The first research approach of this project aimed to identify the various system capacities and leverage points that are important in achieving effective ESD implementation, with a specific focus on how such implementation is structured at a national level. As previously mentioned, the main data collection method was through a country-level survey which was developed based on an earlier defined evaluation framework. The initial evaluation framework was divided into a total of 21 categories; i.e., for the three primary areas of focus – national curriculum, formal education, and teacher training – the framework considered the input, throughput and output capacities; while for the three secondary areas of focus – non-formal education, community and civil society, and private sector – only the input and throughput capacities were addressed. Utilising these individual categories a total of 55 specific target areas were identified for investigation. After a thorough review and consultation process on this framework, it provided the basis for establishing a total of 75 questions for the country survey to ensure broad coverage of all of the target areas.

When considering what capacities are beneficial in the implementation of ESD, it is useful to distinguish different types of capacities that should be investigated. In this work, a basic division of input, throughput and output capacities was applied. This division of capacities can also be related to the division of indicator types utilised by UNESCO Asia and Pacific Regional Bureau for Education in their publication *Asia-Pacific Guidelines for the Development of National ESD Indicators* (2007). These indicator types are respectively status indicators, facilitative indicators, and effect indicators; and defined by UNESCO with the following attributes:

1. **Status Indicators**: assess variables that determine the position or standing of ESD in a country. *Baseline* indicator types belong to this category.
2. **Facilitative Indicators**: assess variables that assist, support or encourage engagement with ESD. *Context*, *process* and *learning* indicator types belong to this category.
3. **Effect Indicators**: assess variables relating to initial, medium and long-term achievements during the DESD. *Output*, *outcome*, *impact* and *performance* indicators belong to this category (UNESCO Bangkok, 2007: 30).

The three types of capacities may also be further subdivided to generate a greater understanding of the potential or necessary components for ESD implementation. Input capacities may include institutional arrangements, policy mandates, and resource capacities (both financial and human), thus addressing the basic structural components to ensure ESD implementation. Throughput capacities should consider the important factors in framing and structuring good ESD implementation, and these may include leadership, knowledge, expertise, and educational pedagogies and methodologies. The output capacities address the quality of learning performance, the impacts ESD is having on learners, and necessary accountability mechanisms, and these may include accountability measurements such as practice standards and targets, value and behaviour change, ESD knowledge gain and assessment tools for monitoring and evaluation.

Compilation of the data and comparison: A simple selection criteria was established to determine which countries would provide the most valuable information for this study, and once the countries were selected the next step was to identify a partner from each country to perform the role of a national focal point for M&E of ESD. Seven countries, at the lead of the national focal point, participated in this study: i.e., Cambodia, China, Japan, Malaysia, the Philippines, Republic of Korea, and Thailand. The country survey was used by the national focal points to submit valuable information regarding the status of ESD implementation in their countries, and where possible the focal points engaged with other experts and professionals in their respective countries to ensure appropriate reporting across the range of sectors covered in the survey. Along with submission of the country surveys, the focal points made presentations of their countries' ESD implementation status in parallel with the case studies of ESD practice from the RCEs at two sub-regional reporting and capacity building workshops on M&E of ESD.

The collected data from the surveys, with supplemental information from the workshop presentations, policy documents and secondary literature, were utilised to prepare ESD status reports for the seven countries. Following this, a comparative assessment of the implementation status of ESD across the seven countries was conducted, drawing out common strengths and barriers in ESD implementation while also identifying several unique implementation factors. An additional capacity analysis was conducted in order to identify key system leverage points for ESD implementation. The initial evaluation framework was coupled with an ESD capacity framework to systematically review the necessary system inputs, throughputs and outputs required for ESD implementation across seven different sectors.[3]

Capacity Analysis and ESD Capacity Framework: Drawing initial insight from the capacity assessment model used by the UN Development Programme (UNDP, 2010b) in which the main levers of change – institutional arrangements, leadership, knowledge, and accountability – are identified each with a few sub-criteria, a capacity framework was developed specifically for ESD (see Table 15) which could be used to help analyse the individual country cases. The structure of input, throughput and output capacities were maintained as this proved an effective means for addressing ESD implementation, practice and performance in a smooth continuum. Each of the seven sectors were then reviewed based on the compiled information from the multi-country comparative assessment. In this analysis, effort was placed on addressing each of the sub-criteria of the ESD capacity framework and considering how they were actualised for each sector.

[3] Originally there were only 6 sectors identified in the evaluation framework, however during the process of conducting this research it was recognised that it was best to address "national policy, mandates, and budget" as a distinct category.

Table 15: Components of ESD Capacity Framework

Input Capacities *(Status Indicators)*	Throughput Capacities *(Facilitative Indicators)*	Output Capacities *(Effect Indicators)*

• Institutional Arrangements
• Policy Mandates
• Resource Capacities
  ▪ Financial
  ▪ Human
 	
Addressing the structural components to ensure ESD implementation
	• Leadership (and vision)
	• Knowledge
	• Expertise
	• Educational Pedagogies & Methodologies
	
	Addressing the framing and contents of good ESD practice
		• Accountability measurements
		  ▪ Practice standards
		  ▪ Targets
		• Value and behaviour change
		• ESD Knowledge gain
		• Assessment tools for monitoring and evaluation
		
		Addressing quality of learning performance, impacts ESD is having on learners, and necessary accountability mechanisms

Findings: Several key findings and recommendations resulted from this research process. The most significant for this work was the identification of list of capacities that proved essential for achieving effective ESD implementation (based on the 7 researched countries). The identified capacities (listed in Table 16) are 32 in total, which represents 58% of the original number of targets identified for review in the evaluation framework (55 in total).

Other important findings specifically address professional capacity, leadership capacity, integration approaches for ESD, application of ESD to different educational systems, and considering both quantitative and qualitative improvements to education. In regards to professional capacity, it was noted that the lack of adequate knowledge and skills for professionals to effectively plan and implement ESD is one of the most fundamental barriers to ESD. This holds true across numerous types of professionals, i.e., policy makers, curriculum developers, and school administrators, but is especially the case for teachers where many have received almost no training for ESD teaching. The need for greater leadership capacity is closely linked to the prior. This can be improved by the formation of a clear vision for ESD with set learning objectives, performance standards and assessment mechanisms; and it can be further support through defined structures for coordination and role sharing in multi-stakeholder implementation of ESD.

The integration of ESD into educational systems has been inconsistent at best and in many cases it has been downright tenuous. There is no one blue-print model for strengthening ESD integration as this is dependent on the context and structure of the given education system, but there are a few basic principles that can support better integration. These include the use of clear teaching strategies, learning methodologies and objectives to structure ESD integration, and the proper reflection of the progressive pedagogies, educational theories and learning methodologies elucidated under the ESD framework.

Table 16: Identified Capacities for Effective ESD Implementation *(replicated from report 1)*

Sectors	Input Capacities	Throughput Capacities	Output Capacities
National Policy, Mandates and Budget	- National SD Plan o Inclusion of education or ESD in SD Plan - Sectorial Mandates for ESD	- Budget, *but considered less significant* o *Maybe project based funding for certain sectors*	
National Curriculum	- Strategic approach for ESD integration o *Potentially addressing how it achieves pedagogical reform & second-order learning* - Authorities with clearly identified roles/ responsibilities for ESD o Also considering level of role sharing across various departments/ offices	- Structure for inter-departmental/ inter-ministerial coordination of ESD implementation - ESD curriculum development support by country's research community & good mechanisms to incorporate expert knowledge and research into curriculum - Vision for ESD identifying clear learning objectives / achievement targets	- ESD implementation leading to wider educational reform *(more as effect indicator than as a capacity for implementation)*
Formal Education	- Specific teaching strategies or educational theories guiding course content & use of progressive, per grade learning objectives o But where ESD is based on thematic inclusion in tradition subjects, the strategies and objectives may be for those subjects, not ESD - Support by school administration	- Wide coverage of important ESD themes and topics - Availability of ESD teaching materials o Supported by use of multi-media formats (electronic versions allow free distribution) - Application of innovative learning methodologies	- Identification of learning outcomes based on distinction of knowledge-based, skill-based, and value-based learning - Clear links between ESD teaching and students' behaviour change o Both provide benefit if they lead to re-setting of targets to improve outcomes
Teacher Training	- Requirement for students teachers to receive ESD training - Amount of TEIs providing ESD training - Amount of In-Service teachers with ESD training	- Wide coverage of ESD thematic topics by the TEIs - TEIs teaching innovative learning methodologies	- Mechanisms for teachers to share good practices in ESD teaching
Non-Formal Education	- Authorities for non-formal ESD both at national and local levels	- Clear vision or strategy outlining the objectives/ achievement targets for ESD in non-formal education sector - Application of good learning methodologies	
Community and Civil Society Participation	- Government support for ESD networks, partnership and relevant CSOs - Number and type of multi-stakeholder networks/ partnerships active in ESD	- Coverage of ESD thematic topics by NGOs - Government support for using media technologies to promote ESD o *Cooperation in international ESD activities (requires more research on existence across region)*	
Private Sector	- Organised cooperation between government and private sector on ESD	- Businesses provide in-service training or continuing professional development on supply chain greening, SCP, environmental management - Businesses provide consumer awareness raising on sustainable consumption options, eco or green products, efficiency issues	

In regards to the application of ESD to different educational systems, it was found that the flexibility of education systems to integrate ESD depends significantly on the state of development for a given system. "Well-developed" education systems with a long history of effective practice are often very rigid, thus ESD topics are often only interjected at the peripheries of the system. While those education systems with a low-level of development are still struggling to meet necessary quantitative improvements to education and often view the qualitative improvements suggested by ESD as a luxury to be relegated to future activities. Finally, it is those systems with a mid-level of development which have recently met target quantitative improvements and are now moving towards wider qualitative educational improvements and reform, that turn to ESD as a beneficial tool to better address the learning needs and challenges of the next generations. The final finding on quantitative and qualitative improvements to education identifies the need for continued discussion and research on how qualitative improvements driven by ESD are also very significant in supporting increased achievement across important quantitative educational statistics.

Research Approach 2: Investigating ESD Learning Performance

The objectives of the qualitative aspect of the research were to identify the important elements of ESD that support effective learning implementation and to develop an ESD Learning Performance (LP) framework based on these elemental characteristics. This was on the premise that lessons learned from the research could be used to 1) ensure better learning performance in the future, 2) help structure good projects by incorporating coverage of all the (four) elements identified and consequently, better understand ESD practices in the context of effectiveness, 3) understand how can we develop indicators for M&E of ESD with the needed contextual knowledge of how educational inputs and throughputs impact the context and processes of learning and how to achieve improvements in learning performance and ESD outputs, and 4) provide further understanding into the dynamics of ESD activities/practices implementation in the RCEs.

A reporting framework for the RCE's ESD good practice case study was developed and sent to ten RCEs in East and Southeast Asia. For East Asia, there were five RCEs: one from the People's Republic of China (RCE Beijing), one from the Republic of Korea (RCE Tongyeong), and three from Japan (RCE Chubu, RCE Okayama and RCE Kitakyushu). There were also five RCEs from Southeast Asia: one each from the Philippines (RCE Bohol), Thailand (RCE Cha-am), Malaysia (RCE Penang), Cambodia (RCE Phnom Pehn), and Indonesia (RCE Yogyakarta). The newest of the RCEs surveyed, RCE South Vietnam was unable to fill out the reporting framework, although it provided a document containing its preliminary activities. The development of the selection criteria for the RCE practice cases was tied to that of the quantitative aspect of the research involving ESD country status reporting. The selected countries were expected to meet several conditions including the following: 1) each country should have at least one RCE, 2) for a country providing more than

one RCE case study, the collection of a diversity of approaches and variations to the ESD activities in those RCEs for example the lead institution of the RCE, the date of establishment, or the geographical and cultural uniqueness were taken into consideration. The selected thematic topics were partly underpinned by the themes that emerged for collaboration at earlier and latter RCE discussions and also by emerging issues.

Compilation of the data and comparison: The questionnaire employed in the data collection from the RCEs was designed in a case-study report format and mainly consisted of open-ended questions and some coded background questions. The following areas were investigated: 1) the major objectives, focus and activities involved in the initiatives of the various RCEs; 2) the benefits of multi-stakeholder partnerships and the learning methodologies, approaches and strategies applied to the initiatives; 3) the main outcomes and achievements of the RCE initiatives; 4) the major strengths and advantages and the primary weakness and constraints, and 5) how the ESD initiative addresses the three pillars of sustainable development. Additional data was collected through presentations and focus group activities in two workshops organised as a capacity building component of the research. After data collation, comparative analysis of the good practice cases was conducted to distinguish the important aspects of ESD that lead to effective learning performance.

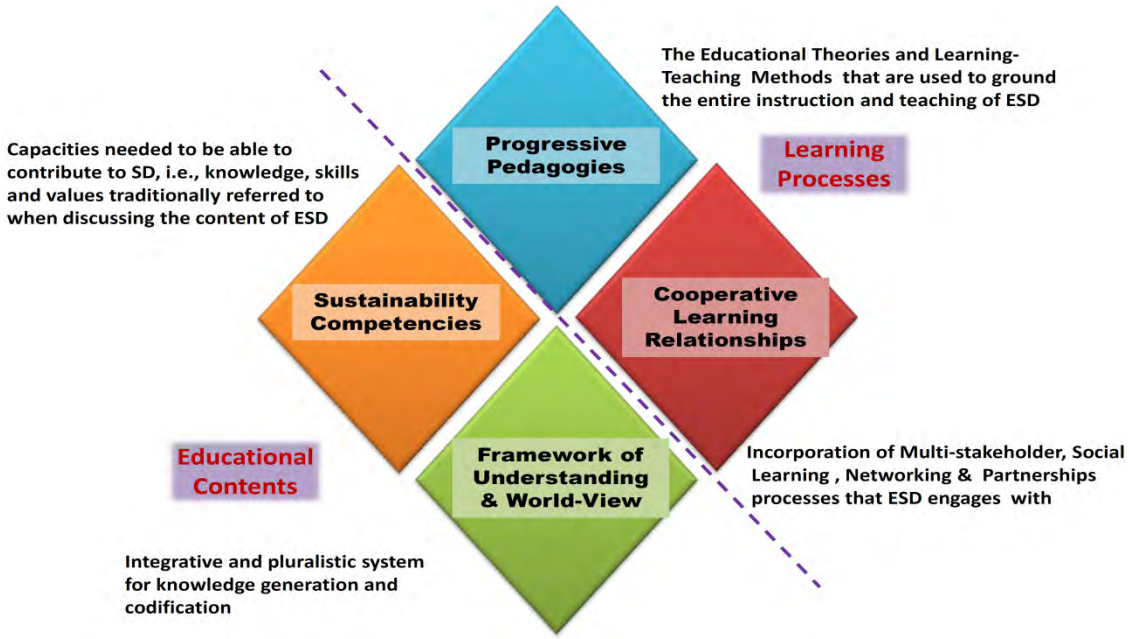
Regarding the following parameters: 1) objectives and foci, 2) learning methods, approaches and strategies, 3) ESD activities of the RCE cases and 4) main outcomes and achievements, comparative qualitative analysis based on the educational/learning process and content (ELPC) types was used to group the parameters into the following: education and knowledge transfer; knowledge acquisition, research and training, workshops/meetings etc., good practices and practical experience, platform for dialogue and community engagement, information sharing and awareness raising, multi-stakeholder partnership and networking. This was prepared with the anticipation that if successful, it would provide a means to compare 'physically different' future ESD initiatives and at different locations.

Learning Performance Framework: Because the results of the various initiatives implemented by the RCEs revealed a diversity of aspects on educational content and learning processes for ESD, evaluating these initiatives' effectiveness through monitoring their progress and contribution to sustainable development without any framework of reference would prove difficult. Consequently, an actionable conceptual framework was developed in order to understand what constitutes effective ESD learning performance (LP). This provided an initial attempt to identify the characteristics of effective ESD learning that are grounded in several educational theories, learning methodologies and approaches. This framework resulted in a definition of effective ESD practice based on the division of four Elements of ESD Learning Performance. Regarding the ESD elements as shown in Figure 2, these four elements of the framework highlight the difference between learning processes and educational contents orientations with two elements within each orientation explained. The two elements classified under learning processes are progressive pedagogies and

cooperative learning relationships, and the two elements classified under educational contents are sustainability competencies and framework of understanding and world-view.

- **Progressive Pedagogies (PP)** looks at the educational theories and learning methods that are used to ground the entire instruction and teaching of ESD; and it emphasises the psycho-social dimensions of teaching and learning relating to the use of approaches including student-centred active learning, critical reflection, problem-solving and cyclical process of inquiry;
- **Cooperative Learning Relationships (LR)** which involves the incorporation of the multi-stakeholder social learning, participatory knowledge generation, and networking processes that ESD engages with community, especially as seen occurring in the RCEs;
- **Sustainability Competencies (SC**) consists of capacities people need in order to be able to contribute to SD; this includes a diversity of knowledge and skill-sets along with values and certain ethical issues; and
- **Framework of Understanding and World-View (WV)** is the interdisciplinary/trans-disciplinary and integrative system for knowledge generation and codification that looks at the types of contextual frameworks and schemes through which individuals shape meaning from diverse knowledge.

Figure 2: The Four Elements of ESD Learning Performance



The successful development of the LP framework means reference could be made to it regarding the implementation of future ESD initiatives towards effective performance outcomes to strategically facilitate translation of the ESD agenda at the local level and into a new global educational/learning framework.

Assessment and Analysis – linking the LP framework with the practice cases: A "cyclical" process of action-reflection between investigation of the practice cases and a reflection on the elemental characteristics underpinned by existing educational theories by which critical knowledge could be acquired through

thoughtful questioning and then testing the initial propositions in comparison with the actual practices was utilised to develop a framework of the important components for ESD learning performance. To further test for validation, an attempt was made to link the elemental characteristics of the learning performance (LP) framework with the actual ELPC-based methods and approaches, activities and practices as well as the outcomes parameters of the RCE cases. Additionally, during the description of the educational/learning theory or methodology that underpin the elemental characteristics making up the ESD LP framework, examples were cited regarding actual pedagogical/learning activity that occurred in the RCEs and were found to clearly relate to the various aspects of the educational/learning theory or methodology.

Findings: The results of the analysis clarified the details (or elemental characteristics) of the four elements of the learning performance framework (as indicated in Table 17). This LP framework consists of several comprehensive, clearly distinct yet overlapping elemental components and they help to highlight the progress of the learning methods, activities and outcomes of RCE initiatives. Consequently, they also provide a better understanding of what are the important factors in delivering effective ESD learning performance.

The results also provide an opportunity to seeing how ESD activities/practices are implemented in some RCEs across East and Southeast Asia at the time it is becoming apparent that the RCEs are acting as regional networks (and also as part of a wider global network) to serve as learning spaces for ESD and platforms for the integration of ESD principles and practices. By integrating the ELPC domains that typify mainly the affective factors of knowledge, skills, values, perceptions as a way to reverberate with the fundamental competencies of sustainability, several of the ELPC-linked methods and practices of the RCE cases could be linked to the LP performance framework. The framework therefore provided a better understanding of the ESD practices in the RCEs with regard to future initiatives. Additionally, most of the learning outcomes of the RCE cases could be linked to the LP framework. We therefore assume this to be the first comprehensive LP framework that is backed fully with literature and covers both the process and content elements of the education and learning for ESD in such clearly distinct yet overlapping manner. Among the learning outcomes, increased awareness, increased knowledge and improved ESD learning were the highest while new vision for the future, ESD integration into the curriculum, values, participation and engagement with community of practice were among the lowest. The major strengths and advantages identified in the RCE cases were: networking, multi-stakeholder participation, engagement/collaborative partnership, self-efficacy, spirit of voluntarism, and high motivation.

Five good practice models namely “Teacher training” “Youth Lead” “HEI-Com” “Multistake” and “Web-based” were identified as “innovative” good practice models for future ESD implementation: 1) “Teacher training” model which was represented by RCE Beijing involves capacity development of teachers in ESD concepts and content using progressive teaching and learning methods including training and practical

Table 17: Specific Elemental Characteristics of ESD Learning Performance Framework

<table>
<tr><th colspan="3">Educational Contents</th></tr>
<tr><th colspan="2">SUSTAINABILITY COMPETENCIES (SC)</th><th>WORLD-VIEW (WV)</th></tr>
<tr><td>Knowledge</td><td>• Climate change
• Disaster risk reduction
• Sustainable consumption and production/Education for sustainable consumption
• Indigenous knowledge
• Information and communication technologies (ICT) and education/ESD
• Well-being, development & environmental quality
• Resilience and socio-ecological Systems</td><td rowspan="3">CHARACTERISTICS
• Holism and Integration
• Systems perspective or whole systems thinking
• Interdisciplinarity and Cross-boundary approaches
• Cultural relativism and Social constructivism
• Pattern recognition, Systems design from patterns to details (synergy)

SUPPORTING (EDUCATIONAL/LEARNING) THEORIES and METHODS:
Systems theory, Critical theory, Transformative learning</td></tr>
<tr><td>Skills</td><td>• Critical thinking and complex thinking
• Conflict resolution
• Seeking alternative solutions
• Adapting to change and advocating for change
• Social action, collaboration and cooperation
• Systems thinking and thinking focused on values
• Conflict resolution, negotiation, creativity and imagination
• Interdisciplinary and trans-disciplinary research skills
• Adaptive learning
• Contextualization of issues
• Personal introspection, visioning and buy-in to identify change and adapting to it</td></tr>
<tr><td>Values</td><td>• Respect, care and empathy
• Charity, social and economic justice
• Citizenship and stewardship
• Empowerment and motivation
• Commitment, cooperation, and compassion
• Self-determination and self-reliance

SUPPORTING THEORIES: Constructivism</td></tr>
<tr><th colspan="3">Learning Processes</th></tr>
<tr><th>PROGRESSIVE PEDAGOGIES (PP)</th><th colspan="2">COOPERATIVE LEARNING RELATIONSHIPS (LR)</th></tr>
<tr><td>CHARACTERISTICS
• Critical reflection & practice and problem solving
• Action/experience–oriented, student-centred learning
• Knowledge production through iterative interaction
• Life-long learning, and
• Cyclical process of collective inquiry

SUPPORTING (EDUCATIONAL/LEARNING) THEORIES and METHODS:
Experiential learning theory, Critical praxis, Critical pedagogy, Problem-based learning</td><td colspan="2">CHARACTERISTICS
• Inclusion and internal network structure for interaction (among social networks) and latitude given for democratic debate on the framing and definition of the issues at stake
• Group processing in establishing and managing systems of knowledge and making sense of information
• Participation and power sharing, shared ownership/commonality
• Clear definition and purpose of roles
• Accountability of individual/groups
• Positive interdependence and building of trust
• Opportunities for reflexive moments and discourse
• Situatedness and Social skills

SUPPORTING (EDUCATIONAL/LEARNING) THEORIES and METHODS:
Social learning (theory and process), Communities of practice, Cooperative learning (theory and model), Cooperative inquiry</td></tr>
</table>

experience; 2) “Youth Lead” model, represented by RCE Tongyeong involves self-capacity building and ESD competence acquisition using collaborative learning relationships, leadership skills acquisition and networking with international peers; 3) “HEI-Com” model as represented by RCE Penang showcases a university-community partnership for capacity building to solve relevant local problems through knowledge transfer, learning and putting knowledge into action. It also strengthens relationships between two different but interwoven communities and simultaneously serves as a research issue/pedagogical resource for the university; 4) “Multistake” model as represented by RCE Phnom Pehn involves multi-stakeholder partnership of RCE with the local farmers and local students and in cooperation with an external advisory body. It is focused on addressing the prevailing environment, economic and health problems, and concurrently using the solution to improve their livelihood including fostering social ties; Some RCEs (RCE Okayama and RCE Kitakyushu exhibited a hybrid/combination of the models; and 5) “Web-based” model represented by RCE Chubu to some extent serves as a reminder of the significant use of ICT in the near future for ESD-based learning and hence needs further investigation.

Research Approach 3: Multi-Stakeholder Participation and Cooperative Inquiry

A third approach was developed for this research project to complement the two primary tracks for data collection and to strengthen methodological triangulation and research validity. This third approach was the incorporation of a strong multi-stakeholder participation element into this research that utilised a process of cooperative inquiry to review, elaborate on, and prioritise the findings and recommendations of the research. During the course of this research process, two expert consultations and two sub-regional workshops were planned, and a third consultation was added after the full conclusion of the research process and analysis to work towards a collaborative agreement on the future application of the findings and recommendations from this research.

This type of participatory process was very important for the overall quality and legitimacy of the research project in three ways. First, it provided a valuable cross-check on both the data being collected and the findings being generated, along with being a source for supplemental information. Second, it provided an opportunity to test the concepts and approaches being developed in terms of their practical applicability, i.e., seeing if individuals familiar with ESD found the proposed frameworks as effective ways for addressing M&E of ESD. Third, as much of what this work achieved was identifying multiple valuable pathways for pursuing M&E of ESD, it was fundamentally necessary to open the decision-making process on these issues to one that included wider stakeholder participation. The research process was able to identify several potential options for supporting effective M&E of ESD, and although the various options are not mutually exclusive certain options do limit the potential for adopting other options. Thus, the trade-offs and

compromises between the various options must be fully considered to move forward with designing an M&E process, but to properly do this they must also be considered against the desired purpose and objectives for the given M&E of ESD. Such decisions though are driven by the specific selected priorities and values for the given process and are also diplomatic in nature, thus they were not ones that the research process alone could comment on in a neutral manner, and as such it was more germane to allow such decisions to be taken through a process of collaborative inquiry and communicative dialogue.

The initial consultation conducted prior to the start of the actual research work was framed to clarify the purpose, focus, and objectives of this research project. During this consultation, the participants considered 1) the primary objectives for ESD monitoring and evaluation across the Asia-Pacific region, 2) the target audience/users of the collected information and findings, 3) the desired outcomes/outputs from the M&E of ESD process, and 4) what would be a useful framework/structure for addressing M&E of ESD with consideration of which sectors, types of indicators, key actors, and leverage points to be addressed. These same questions were also returned to in the third and final consultation to reconsider if the findings from this initial research had impacted on the relevance of the previous decisions and if anything needed to be restructured to better address future activities on M&E of ESD in the Asia-Pacific region.

The two sub-regional workshops for East Asia and Southeast Asia were structured for both research reporting from the national focal points and the RCEs and also to provide capacity building/training for the assembled participants on M&E of ESD.[4] The workshops included facilitated group activities on assessing capacities for ESD implementation and on addressing the effective factors of ESD learning performance. Overall, both the workshops and consultations utilised several forms of interaction from formal presentations to open discussions. Specific sessions were usually framed around individual topics, and the appropriate approach for each session was identified in light of the purpose and goals for it. Besides more standard meeting approaches of presentations, panel discussions and open plenaries, several participatory and action methods were used to facilitate group collaboration.

The working group activities on capacity building served both as a beneficial way to strengthen the participants' abilities to effectively address issues on M&E of ESD and also to provide the research team with valuable supplemental information to that collected in the two primary research tracks. From the capacity assessment group work at the two sub-regional workshops, a long list of potential reporting criteria or factors to consider as ESD indicators were established (the individual outcomes can be seen in the two workshop proceedings, while a combined list is presented in Appendix B and an additional list from the second consultation in Appendix C). The group work on ESD learning performance also generated a large

[4] The two workshop proceedings are available at the following links for the East Asia and Southeast Asia workshops respectively:http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=3522
http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=3988

number of aspects that the participants of these workshops believed to be important for strengthening ESD learning achievements (presented in Appendix D and E).

The group activities were facilitated through different participatory techniques and they were guided by the frameworks developed during this research which specifically helped to test and reform these frameworks. The capacity assessment activity was structured based on the World Café meeting methodology and utilised a simplified version of the ESD capacity evaluation framework in order to consider what were the required capacities for ESD implementation at the level of inputs and throughputs for all six sectors. For the learning performance factors, the Adaptive Nominal Group Technique was utilised to allow groups to brainstorm, propose options, rank, and identify solutions. Utilising the four elements of the ESD learning performance framework (with individual separations for knowledge, skills, and values under the sustainability competencies), the groups were asked to identify both relevant learning goals for each element and specific examples of what would be visible outcomes demonstrating the achievement of said goal.

Benefits of a Mixed-Method Assessment Approach

The major benefit of utilising a mixed-method approach to assessment in this research is that it made possible to address both aspects of ESD implementation and outcomes – or the movement from policy to practice – in a relatively compatible manner. By looking across the entire functioning of the systems for ESD, i.e., considering the inputs, throughputs and outputs, it was possible to not only identify important leverage points for effective implementation at various points, but it was also possible to identify several significant barriers where the process of ESD is hindered. Such barriers could further be elucidated based on either limitations in implementation capacities or deficiencies in addressing the factors for learning performance.

The findings achieved from the first two research approaches also allowed for interplay between the two distinct aspects of ESD. It was thus possible to consider how implementation structures and processes of practice influence or impact on the actual achievement of learning performance for ESD. Vice versa, it was also feasible to reflect on how a clear understanding of the characteristics of ESD learning performance could actually help to improve the implementation and practice approaches of the system. Later during the discussion on the outcomes from the final expert consultation to determine the structure for proceeding forward with M&E of ESD in the Asia-Pacific region, it is possible to see where this distinction between 1) implementation structure, 2) process and practice of implementation, and 3) outcomes and achievement has been integrated into M&E of ESD framework for future application.

In this work, the first two research approaches were distinctly divided to allow for contribution from a wider diversity of stakeholders. However, such division is not a requirement for mixed-methods research, and it is

feasible to apply multiple research methods towards a single research track. Specifically for M&E of ESD, this would depend on the sectors or actors to be covered and the sample size. If the focus was on a small sample such as a school, it would be beneficial to collect some quantitative statistics on the overall implementation of ESD at the school, to conduct a coded review of the topics/themes addressed in ESD teaching materials, to engage in open interviews with both teachers and students to see how they address and understand ESD, to collect good practices on ESD teaching, and to utilise performance based testing to actually assess learning achievement. Such a holistic process would provide a thorough and exhaustive review of ESD practice and outcomes when targeted at a small scale sample. When dealing with large scale samples though, such as national education systems, this funnelling of a mixed-methods approach into one channel can become overburdening. Rather at a national level, mixed-methods approach proves beneficial by applying different research methods to different target actors and sectors with recognition of which types of knowledge/expertise can each actor or sector can best respond to. For example, government officers may find it easier to reply to quantitative questions on implementation structures, but it is still possible to ask some qualitative questions (especially by coded answers) to mix the knowledge generated from one survey.

Generally for the research process, the use a mixed-methods approach was also highly informative by offering a diversity of information and knowledge types. This is something worth keeping in mind when designing M&E processes, for while statistical indicators are usually the aim of such a process their application alone can be rather one-dimensional in the type of focus they lead to which is often a focus that is very mechanical in nature. Including both qualitative reporting and multi-stakeholder participation in the research approach brought a sense life to what otherwise could have been rather tedious material. Furthermore, this helps keep oneself oriented towards the overall goal of conducting M&E of ESD, which of course is to improve its practice and the impacts it is achieving.

Triangulating and Integrating Research Findings

It was already noted that the ideal of a systematic application of performance based testing on ESD across multiple countries is something that is unlikely to be achieved at this point in time, thus an aim was made in this work to demonstrate how it would be possible to address the inclusion of the factors for ESD learning performance across the processes of ESD implementation and practice. Table 18 (on the following page) attempts to link the learning performance elements and characteristics with the relevant capacities at the level of system inputs, throughputs, and outputs. At this point, only general questions are posed to stimulate consideration of how these system capacities can be oriented towards delivering ESD that holistically covers the aspects of ESD learning performance. These questions alone could not be indicators, but rather proxy indicators would be established to demonstrate achievement on these questions.

Table 18: Linking Learning Performance Characteristics to the Capacities for Effective ESD Implementation

	Input Capacities *(status indicators)*	Throughput Capacities *(facilitative indicators)*	Output Capacities *(effect indicators)*

Key Aspects of Capacity	***Addressing the structural components to ensure ESD implementation***
	• Institutional Arrangements
	• Policy Mandates
	• Resource Capacities
	  ▪ Financial
	  ▪ Human
		Addressing the framing and contents of good ESD practice
		• Leadership (and vision)
		• Knowledge
		• Expertise
		• Educational Pedagogies & Methodologies
			Addressing quality of learning performance, impacts ESD is having on learners, and necessary accountability mechanisms
			• Accountability measurements
			  ▪ Practice standards
			  ▪ Targets
			• Value and behaviour change
			• ESD Knowledge gain
			• Assessment tools for monitoring and evaluation

Identified Links to Learning Performance Characteristics	**Institutional Arrangements** Does the ESD system and its implementation structure address ESD in a holistic, integrated, interdisciplinary manner, etc.? *(in line with the approach towards an SD worldview)* Is there a clear definition of roles and responsibilities in ESD implementation and good mechanisms for multi-stakeholder coordination and participation in its implementation? *(establishing cooperative learning relationships in ESD institutions)* **Policy Mandates** Are there strong links between ESD and the country's vision/plan for sustainable development? *(promoting a future vision of an SD worldview)* Does the curriculum structure and ESD strategies enshrine the characteristics of progressive pedagogies *(i.e. student centred, active-experiential learning, collective inquiry, etc.)* Does the curriculum structure and ESD strategies promote skill-based and value-based learning? Does the curriculum structure and ESD strategies promote cooperative and social learning approaches? **Resource Capacity *(human resources)*** Do ESD officers/implementers/teachers hold the necessary expertise for effective Learning Performance teaching in relation to pedagogical and methodological knowledge? Do ESD officers/implementers/teachers hold the necessary expertise for effective Learning Performance teaching in relation to the knowledge and skills sets of sustainability competencies?	**Leadership (and vision)** Does the vision and objectives for ESD promote the idea of improved development, better society, increased well-being, etc.? *(in line with the approach towards an SD worldview)* Does the conceptualisation and framing of ESD include the development of skills and values? **Knowledge** Is ESD framed/practiced in a manner that supports integrative and pluralistic knowledge generation and codification? Is ESD framed around the various sustainability competencies, and does it include knowledge-based, skill-based, and value-based learning approaches? **Expertise** Is teaching informed by and framed around progressive pedagogies? Is teaching informed by and framed around cooperative learning and social learning theory/approaches? **Educational Pedagogies and Methodologies** Is teaching conducted in student centred, active-experiential learning, collective inquiry approaches? Is teaching conducted in a manner aimed at building cooperative learning relationships?	**Accountability Measurements** Is there accountability to oneself, to the group, and to society *(as framed in cooperative learning relationships)?* - Especially framed through group process, role sharing, and interdependence Are there set learning targets that account for the various sustainability competencies – *including knowledge, skills and values?* **Value and Behaviour Change** Are learners approaching the world as meaningful and empowered change agents? Are learners gaining a strong sustainability value set? *(i.e. respect, care, empathy, citizenship, stewardship, self-reliance, etc.)* Are learners demonstrating a higher appreciation towards cooperation, participation and power sharing? **ESD Knowledge gain** Does ESD knowledge and skill gain correspond with the sustainability competencies? Are learners developing the skills for active life-long learning? Are learners developing the skills systematic problem solving? Are learners gaining a holistic, integrated, interdisciplinary understanding of the world? **Assessment tools for monitoring and evaluation** Does the assessment process create a reflexive learning cycle with future improvements being generated from current observations? Does the assessment process allow for collective discourse and reflexivity? Does an effective system for knowledge management, distribution and sharing exist?

Key:

Educational Contents	
Sustainability Competencies	Framework of Understanding & World-View

Learning Processes	
Progressive Pedagogies	Cooperative Learning Relationships

This attempt to link system capacities and learning performance characteristics to develop indicators that can address both practice and performance may not be appropriate though for the initial step of indicators as it does include qualitative criteria that are hard to address for newly established ESD systems. It may be better to utilise a multi-tiered approach to M&E for ESD. The first tier of indicators could be framed solely around the capacities for effective ESD implementation (from Table 16) in a quantitative manner to answer the general question of "has a system been set in place to adequately deliver ESD implementation". Then at the second tier of indicators, it would be possible to address the question of "has the process of ESD implementation and practice been sufficiently framed to deliver the qualitative learning benefits of ESD?" based on a set of indicators that links capacities and learning performance characteristics as suggested. However, it is worth considering if such a tiered approach to M&E should be progressive in nature, and that M&E would begin with the first tier of indicators and only move onto the second tier after proving high performance/achievement in regards to the first tier. It would then also be possible to advocate a third tier of indicators to provide performance based testing of learning outcomes because if an individual country has already committed to and proved high achievement in regards to tiers one and two then the next logical step for strengthening M&E is to address the impact delivered to the beneficiaries of the ESD system.

Framework for Monitoring and Evaluation of ESD in the Asia-Pacific region

Several points must be properly considered in the development of a monitoring and evaluation system of ESD. Relevant criteria for consideration were highlighted in section 2, including:

1) Scope of the research/assessment,
2) Desired coverage, breadth and depth of the research,
3) Focus of investigation (i.e., systems approach, actor approach, or beneficiary approach),
4) Target users of information/findings from M&E process,
5) Desired types of knowledge to be generated,
6) Process of data collection/assessment and time scale of reporting,
7) Validity, Reliability (or replicability), Verification, Comparabilty, and Transferability.

Due consideration of the above criteria would result in differing suggestions for what is the appropriate framework and approach for M&E of ESD depending on the specific context of the criteria. However, in this work a general context has been established that aims to provide regionally-relevant M&E of ESD across the Asia-Pacific region and in a manner that lends itself to national status reporting and cross-country comparative evaluation. Thus, the final suggestions in this work on a possible framework for monitoring and evaluation of ESD in the Asia-Pacific region are based on meeting criteria relevant to the stated context, while other potential M&E frameworks would need to be adapted to meet their different context and criteria.

Based on the development of the evaluation framework throughout the research process, it was decided that seven distinct sectors (or areas) of ESD implementation and practice should be included in the M&E process. These are: 1) national policy, mandates and budget, 2) national curriculum, 3) formal education, 4) teacher training, 5) non-formal education, 6) community and civil society participation, and 7) private sector. However, sector one is understood as providing background context on the prioritisation for ESD in a given country; sectors two, three, and four are the primary sectors of focus; while sectors five, six, and seven are secondary sectors of focus. Input and throughput capacities (as identified in Table 15) should be addressed for all sectors, but output capacities are only to be address for the three primary sectors of focus. This provides a broad scope of coverage on ESD while also allowing for some additional depth in regards to the aspects most relevant for ESD learning performance in formal education (as highlighted in Table 19). This general structure for the M&E of ESD framework was presented at the final consultation for this research phase where it was further refined and adapted (in the addendum of this work, Table 21 is presented as an updated version of this framework based on the agreements from that consultation).

Table 19: General Coverage for M&E of ESD Framework

	Input Capacities *(Status Indicators)*	**Throughput Capacities** *(Facilitative Indicators)*	**Output Capacities** *(Effect Indicators)*
National Policy, Mandates & Budget	✓	✓	⊘
National Curriculum	✓	✓	✓
Formal Education	✓	✓	✓
Teacher Training	✓	✓	✓
Non-formal Education	✓	✓	⊘
Community & Civil Society Participation	✓	✓	⊘
Private Sector	✓	✓	⊘

Future Steps for Monitoring and Evaluation of ESD in the Asia-Pacific region

In furthering the development of effective M&E of ESD, a multi-tiered indicator and reporting process is proposed. As explained previously, three tiers are proposed with each tier having the respective targets: 1) addressing the capacities for effective ESD implementation in a quantitative manner, 2) linking the capacities and learning performance characteristics to address delivery of the qualitative learning benefits of ESD, and 3) provision of performance based testing of learning outcomes. The purpose of a multi-tiered approach is to provide the most complete, holistic version for M&E of ESD, while also recognising the likely limitations for achieving the systematic application of all aspects (related to the three tiers) of the M&E process in a timely manner across all countries in the region. Thus tier one is the easiest aspect to apply, while tier three would be the most difficult.

Tier one of the M&E process would aim to consider how well the system is established to deliver ESD implementation. For this tier, the criteria for M&E would be based on the implementation capacities identified in Table 16. Although these capacities set clear criteria for assessing the functionality of ESD systems, it is still necessary to identify specific indicators that appropriately respond to these capacities and which can be reported in a quantitative manner. Tier two would respond to how well the important learning performance characteristics have been integrated into the overall delivery of ESD. At this tier, the questions raised in Table 18 would form the basis for identifying relevant indicators. Following the final consultation process, there is greater desire to integrate both tier one and tier two into the M&E of ESD work that will be continued in the Asia-Pacific region. However, for tier three which requires the use of performance based testing at a school level, it is recognised that this is currently not feasible to achieve across the region and is rather suggested as a possibility that individual countries could implement if they so desired. For this tier's performance based testing, the assessment would aim to address achievement based on the learning performance elements and characteristics identified in Table 17. The full application of all three tiers of the approach would provide the largest scope across the ESD process from implementation to practice to impact and achievement. As such, it would also allow for the most detailed evaluation including the potential to determine correlations between inputs and outputs or interventions and impacts.

It should be kept in mind that one of key purposes of M&E of ESD is to engender a process of both individual and institutional learning by creating an action-reflection cycle that supports the continual review and improvement of ESD implementation and practice. Monitoring and evaluation is not an end in itself, rather it is used to stimulate action. The proposed framework provides a useful starting point for actualising the M&E of ESD, but further development is still required. Guided by both the framework and the criteria for selecting good indicators discussed in Section Two of this work, the

next step of this project will be to clearly identify and elaborate relevant indicators for use with the framework. Following that, a clear process for reporting and data collection will need to be established. In future efforts, the selected indicators will ideally be piloted across the region which will allow for further refinement. Finally, the hope of this project is to establish a systematic process for countries to report on their status of ESD implementation and practice during the UN Decade on ESD in time to provide valuable inputs to the discussions that will occur at the end of the decade on the future of ESD beyond 2014.

ADDENDUM

Findings from the Expert Consultation on Strengthening Monitoring and Evaluation of Education for Sustainable Development In Asia and the Pacific

Elaborating regionally relevant Indicators of ESD

3-4 December 2012
UNESCO Office, Bangkok, Thailand

Background of Consultation

The research team for this project organised in close collaboration with UNESCO Asia and the Pacific Regional Bureau for Education an expert consultation on *Strengthening Monitoring and Evaluation of Education for Sustainable Development (ESD) in Asia and the Pacific*[1] that was held on the 3-4 December, 2012 at the UNESCO office in Bangkok, Thailand. Fifteen experts in ESD came together to consider the needs for assessing the progress that has been made in implementing ESD in the region during the UN Decade of ESD (2005-2014) with the main goal of elaborating regionally relevant Indicators of ESD for future piloting in the region. This consultation provide the opportunity for the gathered experts to review the findings from the research, to consider the initial proposals made by the research team for structuring a M&E system, and to provide recommendations for moving forward with developing a regionally-relevant system for conducting M&E of ESD.

Through a series of interactive sessions and discussion, the participants considered the key objectives for conducting comparative national ESD monitoring and evaluation across the region and elaborated on the expected target users of the outcomes from this process. Reviewing the findings from the ESD Country Status reports and the ESD good practice cases from the Regional Centres of Expertise (RCEs) from the previous research, the participants also deliberated on the relevance and accuracy of the ESD Monitoring and Evaluation framework. Following on from this, the consultation focused on identifying and elaborating Indicators of ESD for application in Asia and the Pacific.

Targeted Outcomes from the Consultation

The agenda for the consultation was established to provide a logical progression of working through the various aspects of ESD monitoring and evaluation, while also presenting and reviewing the generated knowledge from the research process. During the two days, it was hoped that participants would reach agreement on three important issues:

1. M&E Objectives, Target Users, and Desired Outputs;
2. M&E structure and approach to be applied;
3. ESD Indicators and Consideration of process for piloting.

Furthermore, the agenda included a series of working sessions and activities to develop, refine and propose options for meeting the above agreements. Each session was facilitated with its own structure/approach, information pack, and supporting questions for reflection.

The consultation resulted in several important conclusions and agreements that can serve as a valuable guide for future ESD monitoring and evaluation activities. The key agreements reached during the consultation were on the purpose of regional monitoring and evaluation of ESD, the approach that should be

[1] For detailed information and outcomes from the consultation, see the *Proceedings of Consultation* (2012) at: http://enviroscope.iges.or.jp/modules/envirolib/view.php?docid=4140

taken in conducting ESD monitoring and evaluation and the target areas for investigation, and an elaboration of ESD indicators and reporting criteria in line with the agreed purpose and approach. The main objectives for ESD monitoring and evaluation were stated as: 1) to monitor progress, 2) to learn and improve, and 3) to influence future policy and practice. Three divisions of target audiences were also distinguished: first – influential actors and institution setters, second – implementers and practitioners, and third – learners and beneficiaries. A monitoring and evaluation framework was established to coordinate the strategic identification of ESD reporting criteria and indicators. The three target audiences framed the first division of sectors to be addressed by the monitoring and evaluation process. Sub-divisions were then included that identified the relevant areas of ESD implementation for investigation, including the international framework, national policy and curriculum, formal education, teacher training, non-formal education, and multi-stakeholder collaboration. Finally, the participants identified over 50 potential ESD reporting criteria that will need to be further reviewed and tested against those capacities and leverage points identified from the research as significant for effective ESD implementation.

These key outcomes and agreements from the consultation are presented in the following three tables. They are presented here not as supplemental information to main report, but rather as a continuation and evolution of the M&E approach developing throughout the course of this project. For the overall benefit of this project, on one hand our proposals and recommendations have to be grounded in our research and supported by the generated findings. However, on the other hand, it is also necessary for these proposals to gain legitimacy through the incorporation of open participation and decision making. Although the formal research process can take us the majority of way in identifying what are appropriate aspects to consider in M&E of ESD and what methods of M&E should be applied, it cannot effectively complete this mission on its own. That is because the research for this project is not dealing with a subject that has just one right answer or solution among many other misdirected proposals. As we have seen throughout this work, there are many potentially legitimate proposals for how to approach M&E of ESD, and many of these require trade-offs or comprises between one another based on what are the chosen purpose and focus of the given M&E process as whole. And it is specifically in answering these questions about the desired goals and objectives of the M&E of ESD that the research team alone felt unqualified to answer, not because there was a lack of understanding or even expertise on these issues but because these are decisions that impact on a much wider array of actors than ourselves and thus require greater stakeholder input. As such, this research project has gained cooperation, input, and support from around 50 different contributors, and to all of these colleagues we are greatly indebted.

Thus, the findings from this consultation are seen as part of the process of legitimising the proposals and recommendations initiating out of this research project.

Table 20: SESSION 1 AGREEMENTS – Defining the Purpose of the M&E of ESD Process

Primary Objectives for the M&E of ESD process	Target Audience/Beneficiaries of the findings from M&E of ESD process	Desired Outcomes/Outputs of M&E (in regards to objectives and audience)
Monitor progress • To support countries review and evaluate their progress and achievements made in ESD implementation at the national level (and also at sub-national levels). • To identify both the scale and coverage of ESD implementation. • To ensure that objectives are achieved in a timely manner. • To understand what are the improvements made through the support/stimulus of the UN DESD and the International Implementation Scheme. o Especially the impacts that have occurred at the national (and local) level. **To learn and improve** • To identify good practices in ESD implementation and performance. • To reflect on the unique context and factors of Asia-Pacific region in regards to ESD implementation. • To provide an element of capacity development/ building by strengthening national governments ability for effective M&E of ESD. • To share and disseminate good practices and lessons learned in ESD. **To influence *future* policy & practice** • To improve the support mechanisms for ESD implementation and practice. • To support both policy makers and practitioners in setting the milestones for ESD progress beyond the UN DESD.	**Influential Actors / Institution Setters** • International Organizations • Policy makers (at National level) *Ensuring appropriate allocation of resources – financial support and staffing* **Implementers** **Practitioners of ESD** • Curriculum developers (i.e., Ministries of Education) • Teachers and Students • Principals and School Administrators *Demonstrating expertise on various/diverse topics & Incorporating local context and indigenous wisdom/knowledge.* **Multi-Stakeholders** • NGO and Civil Society • Private sectors (leaders and managers) • Media **Learners and Beneficiaries** • Individuals *(i.e., Individual Learning)* • Community members *(i.e., Cooperative Learning)* • Society *(i.e., Transformative Social Learning)* *Are you convincing those who are not convinced or are you trying to convince those who are already convinced?*	**Understanding current state of ESD and Improve ESD practice, *through:*** • Finding deviations and identifying gaps. • Identifying appropriate interventions to both input side and process side. • Checking impact of ESD on individual learners, on communities, and on society. • Focusing on the quality of education and its reform, *especially education leading:* → first to attitude change, and → second to behaviour change. • Presentation of ESD progress in various manners, types and media formats. • Reducing the gap between research and policy making.

Table 21: SESSION 2 AGREEMENTS - Structure of Monitoring and Evaluation of ESD

Target Actors & Audience	Sectors	Status Indicators *(Input Capacities)*	Facilitative Indicators *(Throughput Capacities)*	Effect Indicators *(Output Capacities)*
IMPLEMENTATION STRUCTURE *Target: Institution Setters & Actors of Influence*	**International Framework & Support**			
	National Framework & Curriculum			
PROCESS AND PRACTICE OF ESD IMPLEMENTATION *Target: Implementers and Practitioners*	**Formal Education**			
	Teacher Training			
	Non-Formal Education			
	Networks & Multi-Stakeholder Collaboration *Including civil society, private sector, community, media, etc.*			
LEARNING ACHIEVEMENT *Target: Learners and Beneficiaries*	**1. Individuals (learners)** **2. Communities** **3. Country/Society**			

Table 22: SESSION 6 AGREEMENTS: Selected Reporting Criteria for M&E of ESD → as initial proposals for establishing ESD

Target Actors & Audience	Sectors	Status Indicators *(Input Capacities)*	Facilitative Indicators *(Throughput Capacities)*	Effect Indicators *(Output Capacities)*

IMPLEMENTATION STRUCTURE
 	
Target: Institution Setters & Actors of Influence
	International Framework & Support	• Does the national ESD framework or policy conform to the International framework (DESD IIS)?
		• Does the country participate in regional (e.g. ASEAN, NEA, etc.) ESD action plans?
		• What are the levels of support you receive from international sources?
		  o Resources (financial and materials)
		  o Expertise
			• Are aspects of multi-lateral international agreements integrated into ESD curriculum and programmes?
			  o Is learning provided on the targets and Agreements countries are trying to achieve?
			  o For programmes, are there activities guided by international standards highlighting good practices?
			• Are there mechanisms for sharing and showcasing good practices?
			• Does your country participate in international SD and ESD policy consultations/processes?
				• What progress has been made in terms of the steps/stages of DESD IIS? *(could be linked to other proxy indicators)*
				• Does the M&E framework address attitudinal and behavioural change? (International and Regional M&E)
				
				Global Goal: Is information from DESD informing the decision making on Post-2015 development agenda/SDGs?
				
				Achievements in regards to sub-regional coordination – ASEAN has scoring system. (existence of sub-regional review and evaluation processes on education action plans).

	National Framework & Curriculum	• Are there existing ESD mandates and policies?
		  o What is their sectoral coverage?
		• Is ESD included in the national SD plan?
		  o Does national plan call for educational response?
		  o Does national development plan incorporate SD/ESD principles?
		• Are there clear allocations of resources (financial, human, etc.) for SD/ESD?
		  o Are the roles of implementing authorities clearly defined?
			• Are there cooperation mechanisms for inter-ministerial coordination of ESD?
			• Are there clear guidelines that outline the priorities and vision for ESD?
			• Is there public and local government participation in ESD national agenda setting?
			• Is ESD literacy one of the key competencies in human resource development ?
				• Is ESD implementation leading to wider educational reform?
				• Does the curriculum allow for inclusion of local context?
				• Does M&E address attitudinal and behavioural change? (national or sub-national M&E)
				• Does the country have a National ESD Action Plan?

PROCESS AND PRACTICE OF ESD IMPLEMENTATION
 	
Target: Implementers and Practitioners
	Formal Education	• Does ESD curriculum exist for all levels?
		  o (Y/N, which levels)
		• Are there partnerships/collaboration with local community?
		  o How many/what kinds?
			• Is research being undertaken on ESD in formal education?
			  o Y/N, how many?
			• Is there compulsory community service on the ESD themes?
			• Availability of teaching/learning materials at all levels?
			  o Which themes are covered?
			• Existence of forums for teacher networking on ESD?
			  o Amount and type?
				• Are ESD themes included in evaluation and assessment?
				  o Which themes, skills, and processes?
				  o Or Knowledge, Skills and Values
				• Are the reward/recognition systems for educators and for learners (e.g. Green Schools/Eco-Schools, etc)

	Teacher Training	• Is there in-service teacher training on ESD
		  o What kind/what scale?
		• Is there a specific curriculum for pre-service teachers on ESD?
		  o What themes are covered?
			• Are there teacher training materials/modules with ESD content?
			  o Which themes are covered?
			  o Are they widely available?
			• Are there ESD resource centres for teachers?
				• Are there ESD themes included in evaluation and assessment?
				  o Which themes, skills, and processes?
				  o Or Knowledge, Skills and Values

	Non-Formal Education	• **Authorities and mandates for non-formal ESD (both at national and local level)**
		  o **Which authorities?**
		  o **How many officers?**
			• **Are non-formal ESD authorities contributing to curriculum and programme development?**
			  o **Who and how?**
			• **Availability of ESD teaching and learning materials?**
			  o **Supported by use of multi-media formats and multi-languages**
			• **Capacity building for staff in non-formal ESD?**
			• **Is there a clear vision or strategy outlining the objectives/achievement targets for ESD in non-formal education sector?**
				• **Learning outcome: is there public participation at local/community on SD projects and activities?**
				• **Is there assessment/evaluation on the achievements/impacts of non-formal ESD projects?**

	Networks & Multi-Stakeholder Collaboration *Including civil society, private sector, community, media, etc.*	• **Number and types of multi-stakeholder networks/partnerships active in ESD?**
		• **Policies and mandates on CSR (private sector)?**
		• **Do private sector/civil society/ NGOs provide support to formal/non-formal education?**
		• **Private sector support for research and learning on good practices/projects?**
		• **Extent of Local resource and knowledge base?**
			• **Quantity and Quality of Platform for SD dialogues?**
			  o **Diversity of representation**
			  o **Frequency of meetings and communication**
			• **Coverage of ESD thematic topics by private sector, media, NGOs?**
				• **Increased public knowledge, understanding and awareness on SD and ESD principles?**
				• **Acknowledgement/recognised by society/recipients of ESD awards (community and business)?**
				• **Networks contributions to formal education, policy, etc.?**

LEARNING ACHIEVEMENT* *Target: Learners and Beneficiaries*	1. **Individuals (learners)**
	2. **Communities**
	3. **Country/Society**
	
	It was noted that these types of indicators/ targets could be used to help frame assessment on interventions or for project evaluation.
		• **Individuals – Participation in community SD/ESD action projects?**
		• **Community – Acquisition of SD skills, capacities and leadership?**
		• **Community – Strong engagement by local communities in national SD agenda/projects?**
		• **Country – Does a national mandate/policy exist for Green Growth or Green Economy?**
		• **Country – Regulations for environmental and sustainability management for businesses?**
		• **Country – Percentage of companies practicing CSR?**
		• **Country – Number of policies on sustainable development, and coverage of policies?**
		• **Country – Social survey of understanding/ awareness on SD issues, policies, and plans?**
			• **Individual – Demonstrated youth leadership in sustainable development?**
			• **Individual – Amount of youth group activities dealing with sustainable development?**
			  o **Rates of participation in such youth groups?**
			• **Community – Infusion of ESD into community activities?**
			• **Community – Clear linkages between tackling global issues through local action?**
			• **Community – Existence of community forums for sustainable development?**
			• **Country – Are businesses engaging in sustainability-related continual professional development (CPD) for management and employees?**
			• **Country – Does the government provide incentives for sustainable practices (i.e., infrastructure development, rebates, grants, reward programmes, etc.)?**
			• **Country – Amount of sustainability-related initiatives linked to achieving international processes and agreements?**
				• **Individual – Improvement in rates of sustainable consumption practices?**
				• **Individual – Number of awareness campaigns and programmes focussing on more sustainable practices/livelihoods/lifestyles?**
				• **Individual – Consumption rate of natural resources and for materials with high-environmental impacts?**
				• **Individual – Improvements in daily practices that demonstrate SD behaviour?**
				• **Community – Improved participation in local governance?**
				  o **What is the level of public participation, and how does the participation occur?**
				  o **Rates of citizen engagement?**
				• **Community – Changes in well-being and livelihoods in accordance with SD?**
				• **Community – Social harmony rate?**
				• **Country – Number of jobs related to a sustainable/green economy?**
				• **Country – Improvements in quality of life indicators (i.e., health, education, emission and pollution rates, etc.)?**
				  o **Does country measure well-being as an index separate from economics (e.g., Gross National Happiness)?**
				• **Country – Good governance practices for sustainable development?**
				• **Country – Social equity?**

APPENDIX

EVALUATION FRAMEWORK AND TARGET AREAS FOR M&E OF ESD RESEARCH AND DEVELOPING ESD INDICATORS *(with coding of responses from country surveys)*

SECTORS	Sub-Sectors	**Status Indicators *(Input Capacities)*** - Institutional Arrangements - Policy Mandates - Resource Capacities	**Facilitative Indicators *(Throughput Capacities)*** - Leadership - Knowledge - Educational Pedagogies and Methodologies	**Effect Indicators *(Output Capacities)*** - Accountability - Learning Performance/Impact - Value and Behaviour Change
National Curriculum (agent: National Government)	Leverage Points	A) Political Mandate for ESD; B) Clear Authorities for ESD implementation; C) Resource provision for ESD	A) ESD mainstreaming and implementation as systemic approach; B) Inter-ministerial Coordination; C) Knowledge Sharing; D) Regional Cooperation on ESD;	A) Holistic & Interdisciplinary Approaches to education; B) Encouraging Reflective & Responsible Behaviour and Critical Thinking; C) Accountability to Promote Institutional Learning
	Barriers	Lack of continuous political support for ESD; Lack of clear authorities for ESD implementation	Coordination btw countries/sharing good practice; Inter-ministerial cooperation; & Connection w/ educational objectives	Difficulty in evaluating ESD implementation; Lack of awareness on promoting behaviour change
	Target Areas	1) Clear Policy Mandate for ESD [National SD Plan = 7+; ESD in Plan = 7+; coverage = 7+] 2) Funding/Budget for ESD (total amount or as percentage of educational expenditure) [2+ (but limited information), 2-, 3 dna] 3) How is Integration into Curriculum structured? [7+ w/ explanation] 4) Is process Streamlined? (with smooth integration and connection to previous policies) [Achievement ratio = 3 +, 4 n/a; Integration = 7+ (w/ explanation); Connection to previous policies = 2+, 4-, 1 dna (no explanation)] 5) Authority, Roles, Responsibilities; are they clearly defined? [6+, 1 n/a (w/ explanation)]	1) Inter-governmental Coordination and Intra-governmental Cooperation (good communication and management) [4+, 2-, 1 n/a] 2) Curriculum Development (expertise supply & demand) [Curriculum Development = 5+, 2- (w/explanation); Research Community = 3+, 2-, 2 n/a (w/ explanation)] Clear Objectives and Achievement Targets for ESD (i.e. Vision for ESD) [5+, 2 n/a (w/ explanation)] 3) Knowledge Sharing and Dissemination (decentralisation of curriculum to classrooms – from policy to implementation) [Decentralization = 2+, 4-, 1 n/a; Communication = 3+, 2-, 1 n/a, 1 dna]	1) ESD as a stimulus of wider educational or curriculum reform, i.e. systemic change. (Such as incorporation of interdisciplinary teaching approaches, team building activities, more action-experiential learning, etc.) [4+, 2-, 1 n/a (w/ explanation)] 2) Feedback Mechanisms & M+E Systems [1+, 5-, 1 n/a]
Formal Education (agent: School Boards, Schools, Classrooms)	Leverage Points	A) Cohesive curriculum strategies on ESD; B) Clear definitions of ESD	A) Good learning materials, B) Good learning methodologies on ESD	A) Promoting a view of synthesis, rather than just analysis;
	Barriers	Continual provision of ESD as progressive educational strategy; Lack of quality criteria/guidelines for teaching tools/materials	Lack of teaching materials and course curriculums on ESD; Weak or no linkage btw. ESD, EFA and MDGs	Difficulty in assessing student learning on ESD; Teachers difficulty in developing ESD courses and materials
	Target Areas	1) Mandate [7+] 2) Budget for ESD (managed by school boards & schools) [1+, 3 -, 3 dna] 3) Teaching Strategies & Course Content (how is ESD framed and entering teaching materials) [2+, 4-, 1 n/a (w/ explanation)] 4) Progressive Learning Objectives (educational strategy) [3+, 4-] 5) Whole School Management Approach to ESD promotion [6+, 1- (some information provided)] 6) Merit-based Appraisal (i.e. positive reinforcement by School Boards for school implementation of ESD) [3+, 4- (some information provided)]] 7) Roles/Responsibilities for implementing ESD teaching (Who coordinates, manages, teaches, assesses, etc.?) [6+, 1 n/a (w/ explanation)]	1) Content/Thematic Topics of ESD *(are these topic covered):* [6+, 1 n/a] a. Climate Change Education [5+] b. Disaster Risk Reduction [3+] c. Sustainable Consumption & Production / Education for Sustainable Consumption [4+] d. Indigenous Knowledge [6+] 2) Teaching Materials for ESD (availability & diversity) [Availability = 5+, 2 n/a; Multi-media formats = 3+, 3-, 1 n/a (some information provided)] 3) Innovative Learning Methodologies (i.e. critical reflection, problem solving, experiential learning, etc) [4+, 2-, 1 n/a (no explanation)] 4) Innovative Teaching Approach (ex. Inter-disciplinary vs. Disciplinary; integrated teaching vs. separate subject) [6+, 1 n/a]	1) Practice Standards & Auditing [1+, 5-, 1 n/a] 2) Learning Outcomes –Process & Skill set development (i.e. collaboration and dialogues, engagement of the whole system, innovation and participatory learning, etc.) [5+, 2 n/a (w/ explanation)] 3) ESD achieving change in consciousness (link between ESD and behaviour/practices, such as in-school recycling programmes and also school disciplinary issues) [2+, 2-, 3 n/a (w/ 1 explanation)]
Teacher Training (agent: Teacher Educational Institutes)	Leverage Points	A) Mandate for ESD in Teacher Education Institutes B) Teachers need to be inspired about ESD!	A) Innovative Educational Pedagogies and Theories B) Strong Professional Competency (based on progressive educational theories including holistic and systemic approach)	A) Linking highly competent ESD staff with individual schools; B) Strong ESD Teaching awareness and skill-sets
	Barriers	Disciplinary boundaries for training in ESD	Lack of trained ESD Teachers; Lack of cross disciplinary teaching ability.	Teachers receiving ESD in teacher training institutes, Teachers receiving in-service ESD training
	Target Areas	1) Mandate for all student teachers to receive ESD [2+, 4-, 1 n/a] 2) Budget (for ESD in Teacher Education Institutes) [7-] How are TEIs engaging with and implementing ESD? (question asked [How many TEIs = 6+, 1 n/a]) 3) Experience with SD experts/professionals and practical SD examples/experiences (i.e. Are future teachers being inspired about ESD?) [3+, 2-, 2 n/a]	1) Thematic SD topics *(as above)* [4+, 1-, 2 n/a] 2) Innovative Learning Methodologies & Progressive Educational Theories [3+, 1-, 3 n/a (no explanation)] 3) In-service Training & Continuing Education on ESD [2-, 5 n/a]	1) Mainstreamed process for teachers to share good practice on ESD [3+, 3-, 1 n/a (w/ explanation)] 2) Assessment of teachers qualifications on ESD [6-, 1 n/a]

KEY: + = positive answers, - = negative answers, n/a = not available, dna = does not apply to country case

= targets that were not specifically addressed via country survey

Non- Formal Education (agent: National and Local Governments, Continuing Education systems)	Leverage Points	A) Experience Based and Field Based Learning B) Practical Learning Centres	A) Outreach and Knowledge Sharing B) Learning by Doing – Good Action & Communication Process	
	Barriers	Availability of institutions specialising in ESD; Ability to provide on the ground advice for ESD implementation	Lack of Strong Coordination to Vision: often due to learning model based on good practice replication	
	Target Areas	1) Mandate [5+, 2-] 2) Budget for non-formal ESD initiatives [2+, 2-, 3 n/a] 3) Quantity and diversity of EE and ESD Learning Centres [3+, 4 n/a (only on quantity, not diversity] 4) Authority & Roles/Responsibilities for promoting ESD in non-formal education (both in the national and local governments) [6+, 1- (some information provided)]	1) Vision or Strategy outlining objectives/achievement targets for ESD in non-formal education sector [5+, 2- (w/ explanation)] 2) Public Outreach and Awareness Raising Activities/Events [3+, 4 n/a] 3) Application of good Learning Methodologies [5+, 1-, 1 n/a (w/ explanation)] • *This section is trying to address the general theme of: How to achieve an impact in non-formal ESD?*	
Community & Civil Society (agent: NGOs and Civic Participation, also role of Media)	Leverage Points	A) Networking & Partnerships on ESD B) Usage of Media Resources	A) Specialised Knowledge and Expertise; B) International Networking; C) Citizen Voice and Empowerment; D) Fostering role of media in consumer awareness raising	
	Barriers	Availability of institutions specialising in ESD; Ability to provide on the ground advice for ESD implementation; Involvement of Civil Society participation in ESD policy formation	* really depends on country context	
	Target Areas	1) Mandate [3+, 4-] 2) Budget [1+, 3-, 3 n/a] 3) NGO and Multi-Stakeholder networks/partnerships for ESD [5+, 2 n/a] 4) Quantity and quality of government support and cooperation with these networks/partnership (and what kind of legitimacy does government bring to them) [5+, 2 n/a (w/ explanation] 5) Citizen Involvement in SD Planning (Is there a mandate? To what extent does it occur?) [2+, 3-, 2 n/a] 6) Civil Society Engagement/Autonomy Rating (use Democracy Index) **not is survey*	1) Diversity of Knowledge on Sustainable Development among civil society (i.e. How many SD themes are currently covered by NGOs?) [5+, 1 n/a, 1 dna] 2) Involvement in international ESD activities. (could be government, academia, civil society, etc) [6+, 1 n/a] 3) Good usage of Media technologies in promoting ESD (or government support for media promotion of ESD) [Use = 5+, 1-, 1 n/a; Support = 4+, 1-, 2 n/a]	
Private Sectors (agent: Businesses and Corporations, Professional Organisations)	Leverage Points	A) Business engagement with government on promoting ESD; B) CSR C) Green product promotion	A) Advancing employee skill development on SCP B) Education Role to Consumer (businesses promoting consumer awareness raising on SCP activities)	
	Barriers	Lack of government cooperation with Private Sector on ESD	Lack of awareness by Business Leaders on SD/ESD	
	Target Areas	1) Mandate 2) Budget 3) Existing networks/partnerships 4) Government led training for business leaders on SD/ESD 5) Mandate on (aspects of) CSR – especially environmental areas *(*no survey question)* *This area focuses on activities led by government or in cooperation with them*	1) In-service Training and Continuing Professional Development (CPD) on supply chain greening and SCP approaches 2) Consumer Awareness Raising Programs/Initiatives *This area focuses on activities led by businesses and corporations.*	

KEY: + = positive answers, - = negative answers, n/a = not available, dna = does not apply to country case,

= targets that were not specifically addressed via country survey

Merged Input from ESD M&E and Capacity Building Workshops 1 & 2 Group Activity – Capacity Assessment on M&E of ESD Status [● East Asia: 20 Feb. 2012, Japan ○ Southeast Asia: 23-24 April 2012, Thailand]		
	Status Indicators (Input Capacities)	**Facilitative Indicators (Throughput Capacities)**
National Curriculum	• Integration (and mainstreaming) of ESD into the National Curriculum • Identifying entry points for ESD promotion and development o Secure funding arrangement specially for ESD o Legal basis for ESD o Structured competencies o Structured & coordinated arrangement for ESD governance o Clear Action Plan (short/medium/long-term) o System to support in-service/pre-service training o Integrate ESD into the national development plan (not only educational plan or environmental plan) o Self-evaluation scheme o National taskforce on ESD for curriculum development including government, public/private sectors	• Locally-relevant Interpretations and Content of ESD • Inter-Governmental and Inter-Ministerial Coordination of ESD o Continuous leadership (as institutional memory) o Develop the guideline/curriculum/materials on ESD o Locally & culturally contextualized and relevant o Dialogue platform o International platform for information-sharing and learning o Consistency in planning, but need to be flexible, adaptable and evolvable through time and needs Something in between – Linkage within different SD pillars and in institutional levels (e.g. between ministries, and among multi-stakeholders)
Formal Education	• Integrated School Management of ESD • Whole School Management of ESD • Addressing ESD in student testing and entrance examinations o Understanding of the concept of ESD (what does it mean with a view to implementation in formal education; specific courses, cross-cutting interdisciplinary approach; specific curriculum room for local knowledge/contextual learning; values/attitudes; assessment). o Formulate National Policy for ESD in line with the National Agenda (informed by PPP → informed by voters/individuals/family → community → consumers → business → government: the crux is infiltrating this knowledge into Education Sectors (often runs contrary to marketing messages/business ethics/modalities). o Embed in criteria for school evaluation and assessment o Linkages to Education Reform Processes, informed by research (practice and theoretical); which inform guideline development (curriculum, teaching pedagogy, education planning, financing as well as content specific integration) o Model cases with ESD implications (contextualized) o National ESD indicators (potentially tied to national agenda or testing measures). o Compulsory community /sustainable service o Systems evaluation o Contextualized learning (i.e. Thailand flooding, informs subject area work)	• Good Evaluation and Accountability (including reward based system) • Networking to link Formal and Non-Formal Education sectors • Allowing flexibility in interpreting ESD curriculum and educational materials to local contexts and needs o Thematic group/locals as resource persons to advise policy/consultations with experts and practitioners o ESD topics based lessons, unit plans, school reform, in-service training, co-curricular activities o Curriculum review of how ESD can be integrated (where it exists, what are the gaps, etc.) ? what levels indicate mainstreaming? (it is not creating, but incorporating) o Holistic approach (whole school, child-friendly) o Integration into formal evaluation o Synthesis of case studies/lessons learned o Policy guidelines and raising awareness o Contrast based learning (experiential education)
Teacher Training	• Funding for teacher education/training • Amount of Pre-service Teacher Education and In-service Teacher Training o Baseline Survey of ESD Knowledge capacities o Clear criteria for teacher's ESD competencies Criteria for Teacher Assessment on ESD (as Performance Indicators) o ESD performance assessment should be linked to internationals assessments such as PISA and TIMAS o Integrate ESD in pre-service teacher training curriculum o Provision of clear ESD teacher training budget o ESD Knowledge platform for sharing between teachers/educators and academe (and focusing on sharing good practices and learning methods) o Identify teachers' roles in ESD at: 1) individual level, 2) institutional level, and 3) community level o Link to research community for ESD development and utilize Action Research o Support and use Information Technologies for ESD promotion	• Trained teacher and focal points of ESD (total number) • Trained education officials on ESD (total number) • Sharing knowledge and good practice o Award scheme for innovation in ESD o ESD teaching guidelines o ESD teaching materials (including concept, exercises, games, cases, etc) o Pedagogy improvement for ESD teacher training o Synergy among interested parties (linking complimentary resources) o Mentoring/Coaching for ESD teaching as a system for teacher development o Teachers also taking role as learner – to strengthen idea of mutual learning process for ESD and to develop progressive contextualization o Genuine Passion for ESD

Merged Input from ESD M&E and Capacity Building Workshops 1 & 2 Group Activity – Capacity Assessment on M&E of ESD Status [● East Asia: 20 Feb. 2012, Japan ○ Southeast Asia: 23-24 April 2012, Thailand]		
	Status Indicators (Input Capacities)	**Facilitative Indicators (Throughput Capacities)**
Non - Formal (NF) Education	• Policy for non-formal ESD • Experiential Education o Clear concept & context of ESD in the NF sector is needed o Need analyses to identify the needs of the target group o Compilation of good (ESD) practices NF sector o Provision of training centres o Funding in priority areas o Need for public and private partnership(s) o Legality (mandate) for ESD in NF education o Need for access to the environment/green spaces for people to be confronted with the contrast o Need for some form of certification to 'protect' quality in the sector and also, to inform those (individuals/NGOs, etc.) who want to come in o Mechanisms to scale up the process & content to a larger audience and also strengthen linkages of the local groups with the regional/global NF educational institutions	• After school curriculum (total number) • Learning by doing o Experienced trainers of ESD o Appropriate knowledge transfer to the specific target group(s) o Reward incentives (as a management tool) o Need to localize & contextualise ESD to the people o Creation of platform/forum for sharing to encourage democratic debate of the issues (Including use of effective strategies for conflict resolution within the group(s)) **BARRIER** Need for a clear distinction between the Non Formal sector and Community and Civil Sector
Community & Civil Society	• Partnership and Networks for ESD (total number) • " ditto " " ditto o"Schools need to have a vision and policy for ESD o School/institution policy on should link to Quality of Life for students o Priority must be assigned to education by the community oRelationship between school and local community (with local government & community leaders and with parent) oMeetings between school and local government/community leaders oIdentification and promote the 'uniqueness of the community' (i.e. food, language, arts, etc.) oNeeds / gap analysis of all community and school/university stakeholders (i.e. what are the needs and gaps that exist) o Focus should be on community Livelihoods (i.e. sustainable livelihoods) oIdentification of strengths ('strengths model') of each stakeholder group oNeed Agreements between schools and institutions and the community oMeetings must be regularly scheduled (discipline to meet continuously and periodically) oCapacity building programs based on community and school / institution needs	• Access to Community Service opportunities • Good Communication Tools and Networks oNetwork Continuity among existing community and education networks oNeed to gauge / understand the attitudes of the different players oDecentralized communication process (2-way communication) oCooperation among stakeholders oContextualization of the process – ESD must be connected to local issues and be locally relevant supporting intergenerational communication oReciprocal Learning (cross learning) between schools/ universities and community oTransfer of Knowledge & Technology through active participation oImportance of establishing a 'safe and trusting' space for dialog and cross learning oKnowledge Management – i.e. documentation of knowledge, tools, methods and processes. oNeed for Sustainable Funding (a revolving fund) to support ESD initiatives b/w schools, universities and community – school and community cooperatives oProcess for 'match-making', linking strengths of different stakeholders to the needs of other stakeholders oCommunication guidelines developed (e.g. for students) oMaterial support based on identifying stakeholder needs, strengths and using agreements to provide 'in-kind' support, e.g. Tool library
Private Sector	• Private Funding Sources • Relationship between Education and Business Sectors	• Promotion of Public Awareness of ESD • Sustainable Reporting System • Business Sector requirements for SD and SCP training and capacity building

FINDINGS FROM GROUP ACTIVITY: REFINEMENT OF THE ESD INDICATORS – ISAP WORKSHOP ON M&E OF ESD, 23 JULY 2012

Purpose: to identify potential indicators (system leverage points) and learning performance targets)

	Indicators ***Target = 6 indicators***	**Learning Performance Targets** ***Target = 3 LPTs***
	Potential breakdown = 2 x Status, 3 x Facilitative, & 1 x Effect	*Considering: Pedagogies, Competencies, Framework of Understanding, & Cooperative Learning Relationships*
National Curriculum	• Identify the priority of the government policy • Integrate ESD into the school evaluation system, curriculum review, and the national testing content • Training for curriculum developers and policy makers • Have a clear legal basis for ESD • Integrate ESD across subjects (interdisciplinary) and label a variety of activities as "ESD"	• Soft-skills of students (collaborative learning, critical thinking, systematic thinking, problem based learning, practice-oriented, etc) • Pedagogy of teaching enhances soft skills and teacher training for learning performance on ESD • Cognitive, socio-emotional development under national curriculum
Formal Education	• Whole school management/policy • ESD Pedagogies • ESD Competencies (awareness, knowledge, skills, attitudes, action, participation, choices) • Integrated school curriculum • Progressive learning objectives • Collective action for change	• Systemic thinking • Problem solving • Action oriented
Teacher Training	• Linkage with local communities • Developing learning materials and teacher networks • Teacher training (pre-service and in-service/ number of in-service teachers having received ESD training) • Mentoring and consulting systems • Guidelines and policies for ESD teaching • Resource mobilization for teachers	• Developing relevant materials • Teacher's communication skills & sharing of best practices and experiences • Awareness of school-level issues

	Indicators *Target = 4 indicators*	Learning Performance Targets *Target = 2 LPTs*
	Potential breakdown = 2 x Status, & 2 x Facilitative	*Considering: Pedagogies, Competencies, Framework of Understanding, & Cooperative Learning Relationships*
Non-Formal Education	• Leadership of the head of the community • Government support • Number of ESD campaigns in the community • Behaviour change of the community • *Policy for non-formal ESD / Evaluation and certificate systems / Linkage with formal education and civil society activities / extra-curricular activities*	• Number of participants & demographic diversity • Multi-stakeholder programs and training programs in the community
Community & Civil Society	• Local community groups (platform for sharing) • Participation in decision-making issues • Community-based projects • Local training programmes	• Local networking and lobbying • Social learning skills • *Developing relevant materials / Managing skills of local ESD knowledge*
Private Sector	• CSR for sustainability (and participation in public campaigns) • Communication (internal systems) and sustainable management • Business to Community (B2C) communication, for example eco-product exhibitions • CSR partnerships	• Transfer of knowledge and technology • Competencies for collaboration activities for sustainability

Input from Group Activity – *Learning Performance Assessment*

From the NE Asia Reporting and Capacity Building Workshop on M&E of ESD; 20 February 2012 at Pacifico Center, Yokohama, Japan

Elements	Learning Objective/Goal *Please identify a key learning objective for this element of ESD learning?*		Educational Achievement *Please identify what would be a visible outcome that would demonstrate achievement of this goal or objective?*
Progressive Pedagogies Educational Theories & Learning	1. Lifelong learning 1. Learning to transform oneself and society 2. Develop capacities of learners in a participatory manner 3. Experimental & experiential learning		• • 2. Develop modules, documentationmaterials and participants …? 3. Combining theories and practice
Cooperative Learning Relationships Social Learning, Networking & Partnerships	1. Sharing leadership on sustainable society 1. Mutual understanding among stakeholders 2. Alignment, collaboration to bring more impact on society 2. Increased partnership 2. Learning different perspectives 3. Collaborative learning beyond schools .e.g., RCE		• • 2. Collaborative projects that emanate from learning 3. Number of networks and platforms
Sustainability Competencies *Linking* Knowledge, Skills & Values *to Action*	**Knowledge**	1. Local/ traditional knowledge, connects with environment 2. Learn, comprehend, understand [how…] 3. Core facts on ESD	1. Participation in local activities 1. System for continuing traditional knowledge and practices 2. Presentation and how students/people relate to the issues at stake 3. Environmental literacy
	Skills	1. Long-term thinking skills, (intergen'nal,future mindedness) 2. Application of knowledge to transform into action 3. Communication	• 2. Examining perspectives, understand the process then take action 3. Problem solving; application of technologies
	Values	1. Think globally, act locally and vice versa 2. To be global citizens, respecting others and tolerance 3. The 3 pillars of sustainability	• 2. Formal/informal opportunities & willingness to share with others 3. Critical thinking
Framework of Understanding & World-View Integrative & Pluralistic system for Knowledge Generation and Codification	1. Positive attitude and self-esteem; moving towards a better future 1. Diversity and multi-culturalism 2. To understand interconnectedness 3. Multiple perspectives 3. Respecting differences		• • 2. Inclusiveness 2. Tolerant society 3. Multi-cultural understanding 3. System thinking

Input from Group Activity – *Learning Performance Assessment*

From the Southeast Asia Reporting and Capacity Building Workshop on E&E of ESD; n

TOPIC: **SUSTAINABLE CONSUMPTION & PRODUCTION**

Elements		**Learning Objective/Goal** *Please identify a key learning objective for this element of ESD learning?*	**Educational Achievement** *Please identify what would be a visible outcome demonstrating achievement of this goal or objective?*
Progressive Pedagogies Educational Theories & Learning Methods		• Experiential and practice-based learning/Training on practical activity • Community immersion/field-based learning • Learning manuals on sustainability competencies	• Reducing waste • Vegetable gardening • Increasing motivation for responsible consumption • Walking/biking • Tree planting
Cooperative Learning Relationships Social Learning, Networking & Partnerships		• Building strong group dynamics • Sharing • Multi-stakeholder learning and sharing • Inter-agency groups • Engage communities	•
Sustainability Competencies *Linking* Knowledge, Skills & Values *to Action*	**Knowledge**	• Understanding the finiteness of the resources and ecological carrying capacities • Know environmental impacts • Health impacts • poverty and … human rights	•
	Skills	• Consuming in a more responsible and sustainable manner	•
	Values	• Ensuring the future and well-being of the future generation • Protection of natural resources • Fair share/equity	•
Framework of Understanding & World-View Integrative & Pluralistic system for Knowledge Generation		• Understanding the worst scenario of ”business as usual” • Knowing sustainable practices and what they can do	•

TOPIC: **CLIMATE CHANGE (CC)**

Elements		Learning Objective/Goal *Please identify a key learning objective for this element of ESD learning?*	Educational Achievement *Please identify what would be a visible outcome demonstrating achievement of this goal or objective?*
Progressive Pedagogies Educational Theories & Learning Methods		• Investigate learning method of local environment situation for CC.	• Students can understand the difference between greenhouse effects and CC • Students know reasons for regular flooding and ways to adapt • Ability to make a choice • Measure GHG using different procedures • Can use scientific knowledge daily • Can disseminate knowledge to others effectively • Can choose the right method to monitor change • Can measure environmental change holistically • Develop genuine/real understanding of the CC context • Students feel empowered to make the situation better
Cooperative Learning Relationships Social Learning, Networking & Partnerships		• To get Integrative Management Approach through consensus	• Ability to identify the different stakeholder groups and their position/perspective to CC • Ability to access and analyse data information on integrative management approach (IMA) • Ability to build consensus KPI to evaluate performance • Ability to see / build sustainable vision for the future • Ability to make policy related sustainable use among stakeholders • Build consensus on distribution of responsibility of different stakeholders • To act in the same way as towards CC (social norm).
Sustainability Competencies *Linking* Knowledge, Skills & Values *to Action*	**Knowledge**	• Ability to analyse present information to predict future situation.	• Integrate content of CC in local development policy/student projec1 • Effective mitigation plan/policy/scientific model • Data presentation and conclusion relating to the CC issue. • Involvement of diverse organizational parties on CC in the local community
	Skills	• Skills to identify the problem and solutions caused by CC	• To promote solutions for CC • Effective construction of instruments used for CC.
	Values	• Students can understand that humans are contributing to the problem of CC	• Ability to list activities that contribute to CC • All people have the responsibility to save the natural resource • Community activities such as cleaning and dredging canals
Framework of Understanding & World-View Integrative & Pluralistic system for Knowledge Generation		•	•

TOPIC: **DISASTER RISK REDUCTION**

Elements	Learning Objective/Goal *Please identify a key learning objective for this element of ESD learning?*		Educational Achievement *Please identify what would be a visible outcome demonstrating achievement of this goal or objective?*
Progressive Pedagogies Educational Theories & Learning Methods	• Critical thinking • Creativity and innovation • Open mindedness • Accountability • Understanding • Recognizing uncertainty		• Dialogical learning environment • Collective decision making • Change in student behaviour • Analytical and effective decision making
Cooperative Learning Relationships Social Learning, Networking & Partnerships	• Accountability • Resource sharing • Understanding of the same concept • Win-win towards common goal • Volunteerism • Mutual help		• Dialogical learning environment • Effective disaster learning • Systematic unity in approaches for DRR • Increase in the spirit of volunteerism and mutual help • Local contextualization
Sustainability Competencies *Linking* Knowledge, Skills & Values *to Action*	**Knowledge**	• Climate change • Disaster risk reduction • SCP/ESC • Well-being, development and environmental quality	• Uncertainty as pre-condition • Understanding of the concept of DRR • Effective procedure/practices of DRR
	Skills	• Critical thinking, systems thinking, complex thinking, real-world problem solving • Seeking alternative solutions • Social action, collaboration and cooperation	• Share common goal • React appropriately in disaster mitigation effects & during disaster • Effective procedure/practices in DRR
	Values	• Commitment, cooperation and compassion • Humanity • Self-determination and self-reliance • Care , empathy & charity	• Effective procedure/practices in DRR
Framework of Understanding & World-View Integrative & Pluralistic system for Knowledge Generation	• Social responsibility • Humanity • Accept disaster as part of the reality of life and learning		• Holism and integration

TOPIC: FOOD SECURITY

Elements	Learning Objective/Goal *Please identify a key learning objective for this element of ESD learning?*		Educational Achievement *Please identify what would be a visible outcome demonstrating achievement of this goal or objective?*
Progressive Pedagogies Educational Theories & Learning Methods	• Mapping the food chain • Food production based on local needs • Utilizing the food available in the locality • Integrated curriculum on eco-friendly food production • Lesson about agricultural production and biodiversity		• Increased awareness of food sources through competencies of the learners • Condition of population density and food availability
Cooperative Learning Relationships Social Learning, Networking & Partnerships	• Identification and coordination of stakeholders (farmers, producers and manufacturers, etc.) • Make task force and survey to connect education and food science and government departments		• Better management of production and distribution towards equal and better access to food • Recognition of diverse interest • Multi-stakeholder platform for food production, distribution and consumption
Sustainability Competencies *Linking* Knowledge, Skills & Values *to Action*	**Knowledge**	• Making curriculum, lecture and assessment for sustainable food production and practical experience in the field • Linking food with other SD issues such as traditional knowledge	• Interdisciplinarity
	Skills	• Collaboration of stakeholders for food production	• Capability of producing sustainable food
	Values	• Inter-generational concern	• Building more knowledgeable community on food sustainability
Framework of Understanding & World-View Integrative & Pluralistic system for Knowledge Generation	•		•

REFERENCES

AAC&U (Association of American Colleges and Universities) (2002) "Beyond Confusion: An Assessment Glossary"; *Peer Review,*2/3: 1-3. [http://assessment.uconn.edu/docs/resources/Andrea_Leskes_Assessment_Glossary.pdf]

ARIES (Australian Research Institute in Education for Sustainability) (2006) ''Development of a National Approach to Monitoring, Assessment and Reporting on the Decade of Education for Sustainable Development" [http://aries.mq.edu.au/projects/esdIndicators/files/ESDIndicators_Feb07.pdf]

Beckwith, J. B. (1991) "Approaches to learning, their content and relationship to assessment performance"; *Higher Education*, 22: 17-30.

Bell, S. and Morse, S. (2003) *Measuring Sustainability: Learning from doing*. London: Earthscan Publications Limited.

Billing, D. (2004) "International comparisons and trends in external quality assurance of higher education: Commonality or diversity? "; *Higher Education*, 47: 113-137.

Birenbaum, M. (1997) "Assessment preferences and their relationships to learning strategies and orientations"; *Higher Education*, 33: 71-84.

Blake, James (1999) "Overcoming the 'Value-Action Gap' in Environmental Policy: Tensions between national policy and local experience"; *Local Environment, 4(3),* 257-78.

Bondelli, K. J. (2013) "An evaluation of the ineffectiveness of the traditional educational system" [http://www.kevinbondelli.com/]

Burford, G. V. Janousˇkova, I. Zahradnik, S. Hak, M., Podger, T. Piggot, D. G. and Harder, M. K. (2013) "Field trials of a novel toolkit for evaluating 'intangible' values-related dimensions of projects"; *Evaluation and Programme Planning;* 36: 1-14.

Dahl, A. L. (1997) "The Big Picture: Comprehensive Approaches in Sustainability Indicators", in B. Moldan and S. Bilharz (eds), *Report of the Project on Indicators of Sustainable Development.* Chichester, UK: John Wiley & Sons.

Dede, C. (2010) "Comparing frameworks of 21st century skills" *In:* 21st century skills: Rethinking how students learn. Bellanca, J. and Brandt, R. (eds.). Solution Tree Press, pp. 51-76.

deHaan. G. Bormann, I. and A. Leicht, A. (2010) "Introduction: The midway point of the UN decade of education for sustainable development: current research and practice"; *International Review of Education*; 56: 199-206.

Delors, J. Mufti, I. A. Amagi, I. Carneiro, R. Chung, F. Geremek, B. Gorham, W. Kornhauser, A. Manley, M., Quero, M .P. Savane, M-A. Singh, K. Stavenhagen, R. Suhr, W. M. and Nanzhao, Z. (1996) "*Learning: The treasure within";* Paris: UNESCO. [http://www.unesco.org/delors/]

DEWHA (Department of the Environment, Water, Heritage and the Arts, Australia) (2010) "*Sustainability Curriculum Framework: A guide for curriculum developers and policy makers*"; [http://www.environment.gov.au/education/publications/pubs/curriculum-framework.pdf]

Didham, R.J. and Ofei-Manu, P. (2012) "*Education for sustainable development: Country status reports. An evaluation of national implementation during the UN Decade of Education for Sustainable Development (2014-2015) in East and Southeast Asia*"; Hayama, Japan. UNU-IAS and IGES.

EfS, (Education for Sustainability)(2012). "The Enviro schools programme in New Zealand" [http://www.education4sustainability.org/?p=234]

Elias, D. (2006) "UNESCO's approach to implementing the Decade of Education for Sustainable Development (DESD) in Asia and the Pacific"; *Australian Journal of Environmental Education*; 22:273-275.

ESDinds, 2012 “The development of indicator and assessment tools for CSO projects promoting values-based ESD” [http://www.esdinds.eu/]

Filho, W. L. (2010) “An overview of ESD in European countries: What is the role of national governments?; *Global Environmental Research; 15*: 119–124.

Flyvbjerg, B. (2001) *Making Social Science Matter: Why social inquiry fails and how it can succeed again*. Cambridge: Cambridge University Press.

Flyvbjerg, B. (2006) “Five Misunderstandings about Case-Study Research”; *Qualitative Inquiry*, *12 (2)*: 219-45.

Flyvbjerg, B. (2011) “Case Study”, Chapter 17 in Norman K. Denzin and Yvonna S. Lincoln, eds. *The Sage Handbook of Qualitative Research*. 4th Edition; Thousand Oaks, CA: Sage.

Gallopin, G. (1998) “Indicators and their Use: Information for Decision-Making”, in B. Moldan, S. Bilharz, R. Matravers (eds), *Sustainability Indicators: Report of the project on Indicators of Sustainable Tourism*. Chichester, UK: John Wiley & Sons.

Garrison, C. and Ehringhaus, M. (2012) “Formative and summative assessments in the classroom” [http://www.amle.org/portals/0/pdf/publications/Web_Exclusive/Formative_Summative_Assessment.pdf]

Giddens, A. (1976) *New Rules of the Sociological Method*. London: Hutchinson.

Golafshani, N. (2003) “Understanding Reliability and Validity in Qualitative Research”; *The Qualitative Report,* 8(4): 597-607. [Retrieved from: http://www.nova.edu/ssss/QR/QR8-4/golafshani.pdf]

Greany, V. and Kellaghan, T. (2008) “Assessing national achievement levels in higher education” The International Bank for Reconstruction and Development / The World Bank. Available at: [http://www.uis.unesco.org/Education/Documents/assessing_national_achievement_level_Edu.pdf]

Hargreaves, Tom (2012) “Questioning the Virtues of Pro-Environmental Behaviour Research: Towards a *phronetic* approach”; *Geoforum. 43*: 315-24.

Hart, M. ‘Indicators of sustainability’; *Sustainable Measures*. internet; [http://www.sustainablemeasures.com]: 2000.

Hay, P. J. and Macdonald, D. (2008) “(Mis)appropriations of criteria and standards-referenced assessment in a performance-based subject”; *Assessment in Education: Principles, Policy and Practice*, 15 (2): 153-168.

Hernandez, R. (2012) “Does continuous assessment in higher education support student learning?”; *Higher Education*, 64: 489-502.

Hill, M. F. (2011) “’Getting traction’: enablers and barriers to implementing Assessment for Learning in secondary schools”; *Assessment in Education: Principle Policy and Practice*, 18: 4 347-364.

IGES (Institute for Global Environmental Strategies) (2005) *Sustainable Asia 2005 and Beyond: In the Pursuit of Innovative Policies*; IGES White Paper: Hayama, Japan.

Jackson, T. (2005) “*Motivating Sustainable Consumption: A review of evidence on consumer behaviour and behavioural change”;* Report to the Sustainable Development Research Network; Surrey: University of Surrey.

Kadoya, S. and Goto, M. (2012) “*The past, present and future of ESD in Japan: How to develop and disseminate ESD at school with the network of the local community*”; International Symposium for Educational Reform: International Trends on Education for Sustainable Development, December 8, 2012, Tokyo.

Kay, K. (2010) “21st Century Skills: Why they matter, what they are and how we get there” *In:* 21st century skills: Rethinking how students learn. Bellanca, J. and Brandt, R. (eds.). Solution Tree Press, pp. ix-xii.

Kolb, David A (1984) *Experiential Learning: Experience as the source of learning and development.* New Jersey: Prentice-Hall.

Kollmuss, A. and Agyeman. J. (2002) “Mind the Gap: What do people act environmentally and what are the barriers to pro-environmental behavior?”, *Environmental Education Review,* 8 (3): 239-60.

Kumar, S. (2002) *Methods for Community Participation: A complete guide for practitioners.* London: ITDG Publishing.

Kyllogen, P. (2008) “The tried the true and the transpiring in creativity assessment”; [http://ec.europa.eu/education/lifelong-learning-policy/doc/creativity/kyllonen.pdf]

Lai, E. R. (2011) “Performance-based Assessment: Some New Thoughts on an Old Idea”; *Bulletin*, 20: 1-4. [http://www.pearsonassessments.com/hai/images/tmrs/Performance-based-assessment.pdf]

Lai, E. R. and Viering, M. (2012) “Assessing 21st century skills: Integrating research findings”; National Council on Measurement in Education. Vancouver, B.C. Pearson. Available at: [http://www.pearsonassessments.com/hai/images/tmrs/Assessing_21st_Century_Skills_NCME.pdf]

Lenglet, F. Fadeeva, Z. and Mochizuki, Y. (2010) “ESD promises and challenges: Increasing its relevance”; *Global Environmental Research;* 15: 95-99.

Liddy, M. (2010) “Monitoring and evaluation: Editorial”; *Policy and Practice: A Development Education Review*, 11: 1-6.

Liskin-Gasparro, J. (1997) “Comparing Traditional and Performance-Based Assessment”; Symposium on Spanish Second Language Acquisition, University of Texas, Austin. [http://www.sedl.org/loteced/comparing_assessment.html]

MacGillivray, A. “Turning the Sustainability Corner: How to indicate right”, chapter 5; Diane Warburton, ed. *Community and Sustainable Development: Participation in the future*. London: Earthscan Publications Limited, first print 1998, reprint 2000.

McKeown, R. (2002) *“Education for Sustainable Development Toolkit”* ; Waste Management Research and Education Institution: Knoxville, TN, USA.

Morse, J. M., Barrett, M., Mayan, M., Olson, K. and Spiers, J. (2002) “Verification Strategies for Establishing Reliability and Validity in Qualitative Research”; *International Journal of Qualitative Methods,* 1(2): 13-22.

Mueller, J. (2012) ‘Authentic assessment toolbox: What is authentic assessment?’ [http://jfmueller.faculty.noctrl.edu/toolbox/]

Mulà, I. and Tilbury, D. (2009) “A United Nations Decade of Education for Sustainable Development (2005-2014): what difference will it make?”; *Journal of Education for Sustainable, Development,* 3: (1): 87-97.

NZC (New Zealand Curriculum) (2013) “The New Zealand Curriculum”. [http://nzcurriculum.tki.org.nz/Curriculum-documents/The-New-Zealand-Curriculum]

OECD/CERI (Organisation for Economic Co-operation and Development/Centre for Educational Research and Innovation) (2008) “Assessment for learning: Formative assessment” OECD/CERI International Conference on Learning in the 21st Century: Research, Innovation and Policy. [http://www.oecd.org/site/educeri21st/40600533.pdf]

Ofei-Manu, P. and Shimano, S. (2010) “Ramsar Wetlands-Rice Paddies and the Local Citizens of Osaki-Tajiri Area as a Social-Ecological System in the Context of ESD and Wetland CEPA”; *Global Environmental Research*; 15: 95-106.

Ofei-Manu, P. and Shimano, S. (2012) “In Transition towards Sustainability: Bridging the Business and Education Sectors of Regional Centre of Expertise Greater Sendai Using Education for Sustainable Development-Based Social Learning”; *Sustainability*; 4: 1619-1644.

Ofei-Manu, Paul and Robert Didham (2012) *Assessment of Learning Performance in Education for Sustainable Development: Investigating the key factors in effective educational practice and outcomes for sustainable development.* Hayama, Japan: Institute for Global Environmental Strategies and United Nations University.

Okamoto, Y., Goto, M., Shirono, M. and Fukuda, O. (2012) "*The education practices utilizing the "Framework Necessary to Plan and Implement Study Guidance Processes for Education for Sustainable Development (ESD)*"; International Symposium for Educational Reform: International Trends on Education for Sustainable Development, December 8, 2012, Tokyo.

Paden, M. (2007). "Tracking progress in the UN Decade of Education for Sustainable Development"; *Journal of Education for Sustainable Development*, 1 (1): 115-116.

Petnuchova, J. (2012) Non-formal and Informal Education: Where Does It Go in the Slovak Republic?" *US-China Education Review* B, 6: 614-625.
[http://www.davidpublishing.com/davidpublishing/Upfile/8/21/2012/2012082171160777.pdf]

Postlewaite, T. N. (2004)"Monitoring educational achievement" International Institute of planning, UNESCO, Paris.
[http://unesdoc.unesco.org/images/0013/001390/139059e.pdf]

Purdue University (2012) "Assessment and evaluation"; Center for Instructional Excellence.
[http://www.purdue.edu/cie/teaching/assessment-evaluation.html]

Pusateri, T. (2009) "The Assessment cyber guide for learning goals and outcomes 2nd Edition
[http://www.apa.org/ed/governance/bea/assessment-cyberguide-v2.pdf]

QAA (Quality Assurance Agency for Higher Education) (2007) "Integrative assessment: Balancing of learning and assessment for learning"
[http://www.scieng.ed.ac.uk/LTStrategy/resources/IABalancingassessment.pdf]

Ragin, C. C. (1987) *The Comparative Methods: Moving beyond qualitative and quantitative strategies.* Berkeley: University of California Press.

Rao, V. and Woolcock, M. (2004) "Integrating Qualitative and Quantitative Approaches in Program Evaluation", Chapter 8 in Francois Bourguignon and Luiz A. Pereira da Silva (eds.), *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*. Oxford: World Bank and Oxford University Press.

Reeves, G. (2010) "A framework for assessing 21st century skills" *In:* 21st century skills: Rethinking how students learn. Bellanca, J. and Brandt, R. (eds.). Solution Tree Press, pp. 305-326.

Rink, J. and Mitchell, M. (2002). "High stakes assessment: A journey into unknown territory"; *Quest,* 54, 3: 205–223.

Scanlan, C. L. (200?) "Assessment, Evaluation, Testing and Grading"
[http://resource.mccneb.edu/edutut/Online%20Pets%20Course/Assessment_evaluation_grading.htm]

Scouller, K. (1998) "The influence of assessment method on student's learning approaches: Multiple choice question examination versus assignment essay"; *Higher Education*, 35: 453-472.

Shaeffer, S. "Beyond 'Learning to Live Together': The key to Education for Sustainable Development". Presentation at UNESCO Expert Meeting on ESD: "Reorienting Education to Address Sustainability", 1-3 May 2006, in Kanchanaburi, Thailand.

Shriberg, M. (2002) "Institutional Assessment Tools for Sustainability in Higher Education: Strengths, weaknesses, and implications for practice and theory", *International Journal of Sustainability in Higher Education. 3 (3),* 254-70.

Stanley, G. (2011) "Teachers' assessments in a standards world" Paper presented at the Oxford University Centre for Educational Assessment Seminar on Teacher Judgment, University of Oxford
[http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/07/Stanley.pdf]

Straight, S. H. (2002) "The difference between assessment and evaluation"
[http://www2.binghamton.edu]

Tilbury, D. (2007) "Monitoring and evaluation during the UN decade of education for sustainable development"; *Journal of Education for Sustainable Development*; 1 (2): 239-254.

Tilbury, D. (2009) “Tracking our progress: a global monitoring and evaluation framework for the UN DESD”; *Journal of Education for Sustainable Development*; 3(2): 189-193.

Tilbury, D. (2010) “Are we learning to change Mapping global progress in ESD in the lead up to “Rio Plus 20””; *Global Environmental Research*; 15: 101-107.

Tilbury, D. and Janousek, S. (2007) “Asia-Pacific Contributions to the UN Decade of Education for Sustainable Development”; Journal *of Education for Sustainable Development;* 1 (1): 133-141.

Tilbury, D. and Wortman, D. (2004) “*Engaging people in sustainability*” IUCN Commission on Education and Communication. IUCN.

Tognolini, J. and Stanley, G. (2007)“Standards-based assessment: A tool and means to the development of human capital and capacity building in education" Australian Journal of Education, 51 (2): 129-145.

Trilling, B. and Fadel, C. (2009) “21st century skills: Learning for life in our times” John Wiley and Sons Inc., San Francisco.

University of Exeter (2013) “Principles of assessment.Types of assessment – some definitions”; [http://as.exeter.ac.uk/support/staffdevelopment/aspectsofacademicpractice/assessmentandfeedback/principlesofassessment/typesofassessment-definitions/#footnote1]

UN (United Nations) 2005. “*United Nations General Assembly Resolutions, 59/237*”;

UNAIDS (2010) *An Introduction to Indicators*; series: UNAIDS Monitoring and Evaluation Fundamentals, pub. 2. Geneva: UNAIDS.

UNDP (2010a) *Measuring Capacity*. New York: UNDP. [http://www.undp.org/content/undp/en/home/librarypage/capacity-building/undp-paper-on-measuring-capacity/]

UNDP (2010b) *Supporting Capacity Development: The UNDP Approach.* New York: UNDP. [http://www.undp.org/content/dam/aplaws/publication/en/publications/capacity-development/support-capacity-development-the-undp-approach/CDG_Brochure_2009.pdf]

UNECE (United Nations Economic Commission for Europe) (2009) “*Learning from each other: The UNECE Strategy for Education for Sustainable Development*”, UNECE, Geneva. [http://www.unece.org/fileadmin/DAM/env/esd/01_Typo3site/LearningFromEachOther.pdf]

UNESCO (2004) “United Nations Decade of Education for Sustainable Development (2005-2014): International Implementation Scheme””; UNESCO, Paris.

UNESCO (2005a) Report by The Director-General on the United Nations Decade of Education for Sustainable Development: International Implementation Scheme and UNESCO’s Contribution to the Implementation of The Decade, UNESCO, Paris.

UNESCO (2005b) “*United Nations Decade of Education for Sustainable Development (2005-2014) – International Implementation Scheme*”; Paris: Section for Education for Sustainable Development, Division for the Promotion of Quality Education, UNESCO [http://unesdoc.unesco.org/images/0014/001486/148654E.pdf]

UNESCO (2006a) “Pacific Education for Sustainable Development Framework”; [http://unesdoc.unesco.org/images/0014/001476/147621e.pd]

UNESCO (2006b) “*Education for Sustainable Development Toolkit*– series: Education for Sustainable Development in Action – Learning & Training Tools no. 1; UNESCO Education Sector. Paris: UNESCO. [http://unesdoc.unesco.org/images/0015/001524/152453eo.pdf]

UNESCO (2006c) *Promotion of a Global Partnership for the UN Decade of Education for Sustainable Development: The International Implementation Scheme for the Decade in brief*. Paris: UNESCO. [http://unesdoc.unesco.org/images/0014/001473/147361e.pdf]

UNESCO (2007a) “*The UN Decade of Education for Sustainable Development (DESD 2005-2014). The First Two Years*”; UNESCO, Paris.

UNESCO (2007b) "*The Records of the 34th session of the General Conference: Volume 1, Resolutions*"; UNESCO, Paris.

UNESCO (2009) "*United Nations Decade of Education for Sustainable Development (2005-2014) – Review of Contexts and Structures for Education for Sustainable Development 2009*"; Prepared by Arjen Wals; Paris: UNESCO. Available at: http://unesdoc.unesco.org/images/0018/001849/184944e.pdf

UNESCO (2010) "*Education for Sustainable Development Lens: A policy and practice review tool;* series: Education for Sustainable Development in Action – Learning & Training Tools no. 2; UNESCO Education Sector. Paris: UNESCO.

UNESCO (2011) "*Education for Sustainable Development: An Expert Review on Processes and Learning for ESD*" Paris, UNESCO. http://unesdoc.unesco.org/images/0019/001914/191442e.pdf

UNESCO (2012a) "*Exploring Sustainable Development: A multiple-perspective approach;* series: Education for Sustainable Development in Action – Learning & Training Tools no. 3; UNESCO Education Sector. Paris: UNESCO.

UNESCO (2012b) "Shaping the Education of Tomorrow: 2012 Full-length Report on the UN Decade of Education for Sustainable Development DESD Monitoring & Evaluation - 2012 UNESCO Education Sector" Available at: http://unesdoc.unesco.org/images/0021/002164/216472e.pdf

UNESCO (2012c) "*Education for Sustainable Development Lens: A policy and practice review tool;* series: Education for Sustainable Development in Action – Learning & Training Tools no. 4; UNESCO Education Sector. Paris: UNESCO.

UNESCO APRBE (Asia and Pacific Regional Bureau for Education) (2005a) *"Working paper: Asia-Pacific Regional Strategy for Education for Sustainable Development",* Bangkok: UNESCO. Available at [http://www2.unescobkk.org/elib/publications/esd_working_paper/working_paper_2ed.pdf]

UNESCO APRBE (Asia and Pacific Regional Bureau for Education) (2005b) "*A Situational Analysis of Education for Sustainable Development in the Asia-Pacific Region" Bangkok*: UNESCO. Available at: [http://www.mma.gov.br/port/sdi/ea/deds/arqs/sitan_edsasia.pdf]

UNESCO APRBE (Asia and Pacific Regional Bureau for Education) (2007) *Asia-Pacific Guidelines for the Development of Education for Sustainable Development National ESD Indicators*. Bangkok: UNESCO [http://unesdoc.unesco.org/images/0015/001552/155283e.pdf]

UNESCO APRBE (Asia and Pacific Regional Bureau for Education) (2011) *Astrolabe: A guide to Education for Sustainable Development Coordination in Asia and the Pacific.* Bangkok, Thailand: UNESCO.

Volante, L. and and Jafaar, B. (2008) "Profiles of education assessment systems worldwide: Educational assessment in Canada"; *Assessment in Education: Principles Policy and Practice*, 15: 201-210.Weber, K. (2012) "The use of learner-centered assessment in US colleges and universities" *Research in Higher Education*, 53: 201-228.

Wals, A. (2010a) "DESD We Can? Some lessons learnt from two Mid-mini DESD Reviews" *Global Environmental Research;* 15: 109-118.

Webber, K. L. (2012) "The use of learner-centered assessment in US Colleges and Universities"; *Research in Higher Education*, 53: 201-228.

Wilkerson, J. R. and Lang, W. S. (2004) "A standards-driven, task-based assessment approach for teacher credentialing with potential for college accreditation"; *Practical Assessment, Research & Evaluation*, 9 (12). Retrieved February 12, 2013 from http://PAREonline.net/getvn.asp?v=9&n=12

William, D. Lee, C. Harrison, C and Black, P. (2004) "Teachers developing assessment for learning: Impact on student achievement"; *Assessment in Education: Principles, Policy and Practice,* 11 (1): 49-65.

Wilson, M. and Scalise, K. (2006) "Assessment to improve learning in higher education: The BEAR Assessment System"; *Higher Education*, 52: 635-663.

Wren, D.G (2009) “Performance Assessment: A key component of a balanced assessment system”; *Department of Research, Evolution and Assessment*, 2: 1-12] [http://www.vbschools.com/accountability/research_briefs/ResearchBriefPerfAssmtFinal.pdf]

Yorke, M. (2003) “Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice”; *Higher Education,* 45: 477-501.

ISBN: 978-4-88788-118-1

Institute for Global Environmental Strategies
公益財団法人 地球環境戦略研究機関
2108-11 Kamiyamaguchi,Hayama, Kanagawa 240-0115 Japan
神奈川県三浦郡葉山町上山口2108-11　〒240-0115
Phone:+81-46-855-3700 Facsimile:+81-46-855-3709 http://www.iges.or.jp/