History and Status of Development Evaluation: Indian Context

Submitted in partial fulfilment of the requirements
of the course
TD 694

By

Yatin RS Diwakar
174350003

Supervisor

Prof. Bakul Rao

Centre for Technology Alternatives for Rural Areas (CTARA)
Indian Institute of Technology Bombay

August, 2019

Seminar Report Approval

This is to certify that the 'TD 694 Seminar' course report titled "History and Status of Development Evaluation: Indian Context" by **Yatin RS Diwakar** (174350003) is approved.

Prof. Satish B Agnihotri   **Prof. Bakul Rao**

(Examiner)   (Guide)

Date: 24 April 2018

Place: Mumbai

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words are included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any ideas, data, facts, or sources in my submission. I understand that any violation of the above will be cause of disciplinary action by the institute and evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: 18 April 2018   **Yatin RS Diwakar**

Place: Mumbai   **174350003**

Abstract

Development evaluation is systematic assessment of community development programs aimed at analysing the level of achievement against set criteria. Evaluation has a history of over a century, while development evaluation started in the aid-era post second world war and picked up with the millennium development goals. Important international agencies involved in it, and shaping the field are United Nations Evaluation Group, Independent Evaluation Group, International Organisation for Cooperation in Evaluation, etc. and these organisations are pushing nations to commit to evaluation and assisting in building evaluation capacity. Professional-Academic bodies, such as American Evaluation Association, are at the forefront of knowledge development in evaluation, including evaluation theories, logic of evaluation and various classifications. India started evaluating planned social programs through the Program Evaluation Office under Planning Commission in the 1950's and all states also established their own evaluation offices in the early 60's. After an initial phase of great growth and focus, later evaluation function was slowly degraded and ignored, till the turn of the century when it picked up again with international stimulus.

This literature review captures the historical timeline of development evaluation globally, and in India, and discusses the current scenario, as a precursor to building in depth knowledge about the field of evaluation. It lays out the basics and terminologies of evaluation, including classification types, theories, and logic of evaluation, which are currency of the field. Important journals in evaluation are also listed. The review concludes with some future directions of research.

Table of Contents

List of Figures

List of Tables

Chapter 1   Introduction

"Measurement is the first step that leads to control and eventually to improvement. If you can't measure something, you can't understand it. If you can't understand it, you can't control it. If you can't control it, you can't improve it."

- H. James Harington (performance and quality improvement guru)

This literature review explores the field of development evaluation, its history and status. While the report focuses on evaluation, this first Chapter discusses development programmes and need of evaluation for the same and gives a background behind motive and methods of this study.

1.1 About the Review

This section provides the background to this literature review study. The primary motivation of this study is to obtain a primary understanding of evaluation as a field through review of existing literature.

1.1.1 Motivation

The current literature review is carried out to understand the field of development evaluation and what is happening in it internationally and nationally. Evaluation has a long history, though development evaluation picked up globally in the period after the second world war. India too adopted development evaluation as an integral part of its planned economy model from the beginning. In spite of this, today, about 70 years later, we see that a lot of these processes are not functional or not having intended outcome.

Today, when the world is focusing on data collection, monitoring & evaluation under the Sustainable Development Goals, India too is working towards a national evaluation policy. Due to personal interest in the subject of evaluation and its use in the development cycle, this literature review is a first step towards an in-depth understanding of evaluation as a subject, its

major theories, logic, international paradigms, current research, international and national and state-level institutions.

1.1.2 Methodology

The method followed for this report is literature review; a couple of field experts were written to, but no response was received. Initial literature was searched through use of various keywords and from initial articles obtained, their references were used for further searches. For the couple of tables in the report, extensive web-based search was carried out. For finding out state-level government evaluation agencies, each state's official website was explored wherever simple key-word searches didn't open relevant websites.

1.1.3 Scope

In current study, only first-level of exploration is done. Many further threads were opened up during this study, some of them are mentioned in future works section at the end. This study is solely based on material available (and accessible from IITB network) on internet. Being first foray into the field, this study touches upon a variety of subjects, including history, institutions, theory, current practices, etc. without delving in depth on any one.

1.1.4 Organisation of report

In Chapter 2, we discuss what is Evaluation, its types, and theories. Chapter 3 looks at growth of evaluation as a sector, especially the development sector, both internationally and in India. Chapter 4 looks at the direction development evaluation is taking today. Chapter 5 concludes this review of evaluation of development programs with some observations from the author and proposes some topics for further study.

1.2 Defining Development

While looking at development programs, the first question which often comes to mind is what development is. A great amount of literature exists on the subject; Prof. Ashok Gadgil, University of California Berkeley, and editor-in-chief of Development Engineering journal, defines it crisply in his interview at CTARA. When asked what his opinion on development was, he defines development as *freedom from being forced to do unwanted things out of economic necessity*. He mentions that development gives a person choice and allows them to say no. He concludes by saying that development is to have more control of one's destiny. (TDSAA CTARA, 2017).

Owen Barder, Vice President at Center for Global Development, defines development similarly. In his blog-post based on his Kapuscinski Development Lecture in May 2012, he uses Nobel Laureate Amartya Sen's capabilities approach and 'Development as Freedom' (Barder, 2012). He stresses that judging development by its effects on people does not mean that development is the sum of improvements in the wellbeing of people or the output of firms. He argues, borrowing from Amartya Sen's 'Development as Freedom', that we should think of

> *"Development as emergence of a system of economic, financial, legal, social, and political institutions, firms, products, and technologies, which together provide the citizens with choices, capabilities and freedoms to live happy, healthy lives with access to knowledge, a decent standard of living and allow them to participate in the life of their community."*
>
> - (Barder, 2012)

The conventional economic models of development evolved gradually to focus upon each of labour, capital infusion, technical change, policy changes and institutional changes over the century to explain development deficits between similar groups (Barder, 2012), and development programs have evolved parallelly over the years, to provide the missing link as per the incumbent paradigm.

1.2.1 Development programs and evaluation

Development programs are concerted efforts to develop a community keeping in mind the development paradigm and definition of development of the program implementors. Over

the decades, development programs and projects were planned and implemented to induce and sustain development, and United Nations has played a crucial role in shaping these programs by pushing forth Human Development Index (HDI), Multidimensional Poverty Index (MPI), Millennium Development Goals (MDGs), and now Sustainable Development Goals (SDGs). Today, the understanding of multidimensional nature of poverty and development drives most development programs, and distribution or variance is also a concern beyond the averages of measures of development (Barder, 2012).

A typical community development project has six stages: *Study of the community, Goal setting, Identification of solutions, Design, Implementation of the solution, and its sustainable operation*. For a development program to be successful, it should bring about expected change sustainably. (Honkalaskar, 2014). When defining whether a program succeeded or failed, to check whether the expected change has occurred, it has to be judged on some parameters. The expectations must be defined and progress along them needs to be measured and fed back to the process. Any course-corrections in these six stages and learnings to improve future iterations come through effective feedback mechanisms. If these feedbacks measure progress of the program, they can be termed as monitoring. If these are used to understand the processes and estimate the success of the program, either during or post completion, it can be termed as evaluation. Data is needed to efficiently monitor progress of the program and evaluate whether it succeeded or not.

Figure 1.1 shows how monitoring and evaluation is involved in the various stages of a program, including right time for developing a M&E framework and detailed evaluation.

Under the Sustainable Development Goals of the 2030 'Agenda for Sustainable Development' adopted on 25th September 2015 in UN General Assembly, countries will mobilize efforts to end all forms of poverty, fight inequalities and tackle climate change, while ensuring that no one is left behind (United Nations, 2016). The 17th SDG, partnerships for the goals, recognises data, monitoring and accountability as a way to address systemic issues. It is widely accepted that evaluation will play a crucial role in the success of the SDGs (Schwandt, Ofir, Lucks, El-Saddick, & D'Errico, 2016), (Heider, 2015) and (OECD, 2016).
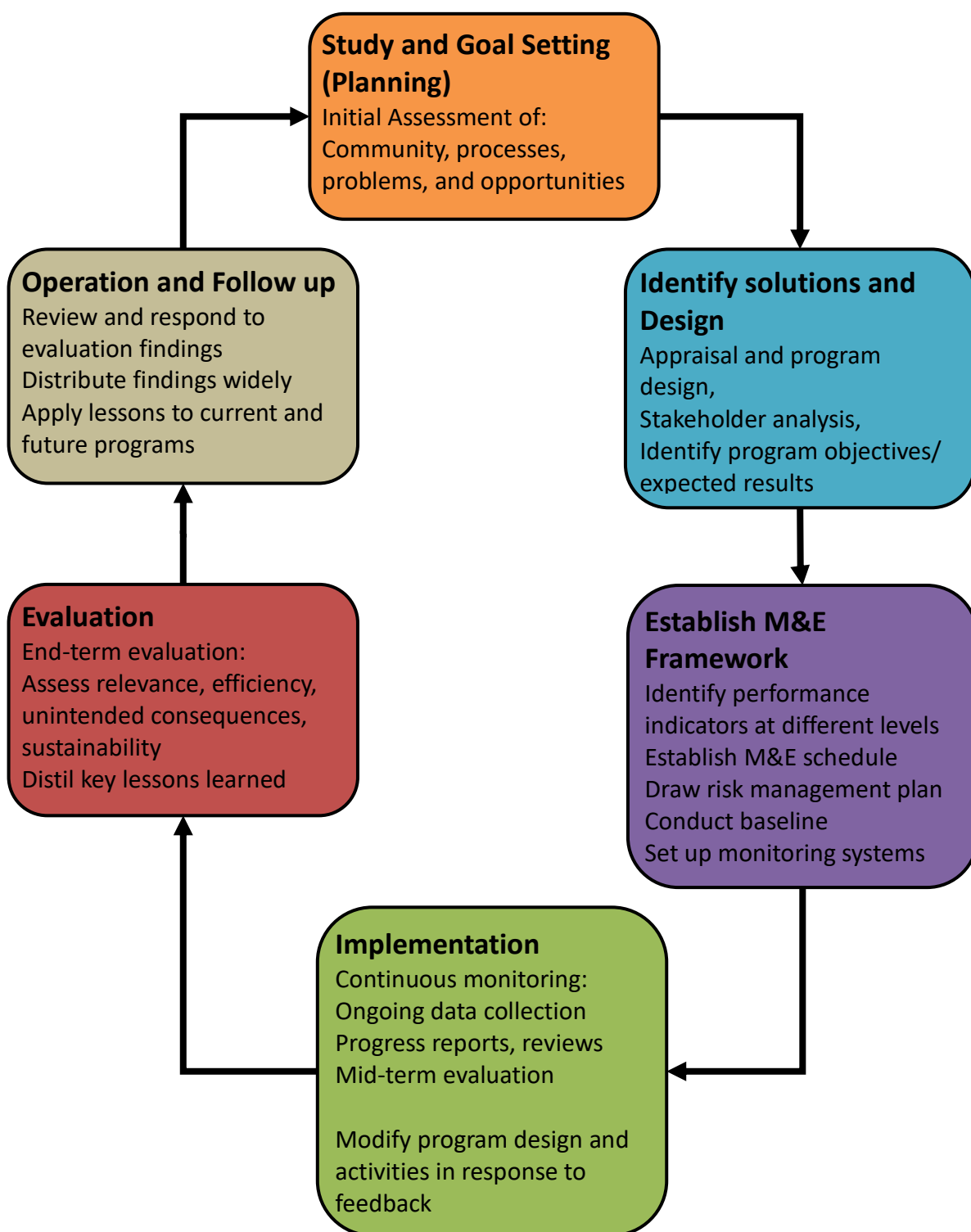
Figure 1.1: Monitoring and Evaluation in a Program Cycle

*Adapted from Virtual Knowledge Centre to End Violence Against Women and Girls, Security -
M&E plans (UN Women, 2012)*

Chapter 2   Evaluation Basics

"The discipline of evaluation is devoted to systemic determination of merit, worth, or significance."

- (Scriven, 1999).

Michael Scriven, co-director of Claremont Evaluation Center and author of many seminal texts on evaluation, including Evaluation thesaurus, defines evaluation (verb) as a process of determining merit, worth, or significance and an evaluation (noun) as a product of this process (Scriven, 2007). In this short definition, there is no mention of what is evaluated, how, and why it is evaluated. This chapter touches upon the concepts related to evaluation, including types, theories, and logic of evaluation.

2.1 Defining Evaluation

United Nations Evaluation Group, which influences how many countries think of evaluation, defines evaluation in its Norms and standards of Evaluation handbook as

"An evaluation is an assessment, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area, or institutional performance. It analyses the level of achievement of both expected and unexpected results by examining the results chain, processes, contextual factors, and causality using appropriate criteria such as relevance, effectiveness, efficiency, impact, and sustainability. An evaluation should provide credible, useful evidence-based information that enables the timely incorporation of its findings, recommendations, and lessons into the decision-making processes of organizations and stakeholders."

(UNEG, 2016)

This definition defines what is evaluation, what is evaluated, what components it has and its use. Most other definitions are variations of this, generally limiting to one aspect or the

other, and hence need not be discussed here. A few other terms used regularly along with evaluation are defined next.

2.1.1 Related concepts

Generally, monitoring and evaluation are referred to together in most development programs, while the terms assessment and audit are also regularly used. It is worthwhile to look at their official definitions.

Monitoring:

Monitoring is the "regular collection and analysis and distribution of information for the surveillance of progress of the project's implementation". Project monitoring is the collection of data prior to, and during, the project. These data, when analysed, pinpoint progress or constraints as early as possible, allowing project managers to adjust project activities as needed. Monitoring is a continuing process throughout project implementation and often extends beyond project completion (UNEP, 2008).

Assessment:

The Merriam Webster Dictionary defines Assessment as the action or an instance of making a judgment about something or the act of assessing something.

Audit:

Audit is also defined as a systemic review or assessment of something, or an official inspection of an organisation's accounts, being conducted typically by an independent agency. Audits can be conducted by bodies internal or external to the overall organisation and this gives rise to the term third-party audits, where the auditing body is external to the entire organisation. Also, audit also refers to technical assessment.

It can be seen that audit, assessment and evaluation are interchangeable, though the usual connotation attached to these terms and their use in literature means that evaluation is more preferred while referring to systemic determination of value of an activity.

Since monitoring collects data/ information about the indicators which are also used for evaluation, and provides feedback during the activity, many times monitoring and evaluation are addressed together in literature on program management. The following table gives primary differences between the two.

Table 2.1: Comparing Monitoring and Evaluation

Characteristics	Monitoring	Evaluation
Subject:	Address operational management issues	Focused on strategic aspects
Character:	Continuous, regular, systematic	Incidental, flexible subject & methods
Primary client:	Program management/ implementors	For implementers, but also for donors and beneficiaries, the broad interested public and research community, should the results be made public
Approach:	Utility	Objectivity, transparency
Aim:	Improve efficiency and effectiveness; ensure the transformation of activities into results/ outputs	Explore intended and unintended results; formulate recommendations for adjustments
Methodology:	Rapid appraisal methods	Rigorous research methodologies, sophisticated tools
Focus:	Check how planned activities are progressing, identifying operational difficulties; recommend actions for improved implementation efficiency and effectiveness	Assess achievement of results, relevance, effectiveness, and impact and sustainability of activities and their contribution to results.
Objectives:	To identify and resolve implementation problems, to assess progress towards objectives	To verify developmental hypothesis, to document successes and lessons learned

Time	Continuous process during the implementation of a project, programme, plan, or strategy	Either done during the span of a project, programme, plan, or strategy (mid-term evaluation) or upon completion (ex-post evaluation)
Content	Progress attained measured against the set indicators	Fulfilment (degree of and quality) of the indicators

Adapted from (Jaszczołt, Potkański, & Alwasiak, 2003) and (Hagen & Willems, 2010)

It can be seen that Monitoring and evaluation are both crucial parts of a program cycle and complimentary to each other.

2.2 Classifications of Evaluation

Barbara Lee, in her chapter 'Theories of Evaluation' (Lee, 2000) concedes that after decades of discussion about the discipline, well-established practitioners too are unable to define evaluation. Most people involved in or writing about evaluations are not formally educated in the sector, though programs for it exist. Many who shaped the discipline, entered it in early or mid-career, mostly from social sciences, and had a long-lived interest in evaluation.

Thus, based upon the practitioner and their interests, evaluation is grouped into various types. Here, we will look at classification based on field of evaluation, time of evaluation and purpose of evaluation.

2.2.1 Types based on evaluation fields

Michael Scriven, known for his extensive contributions to evaluation methods in public policy and social sciences, categorises by evaluation fields. He identifies over twenty fields of evaluations and gives practical and logical importance to eight of these, which he groups into the Big six and Super two. The Big Six are *program, personnel, performance, policy, proposal, and product evaluation* and the Super two are *intradisciplinary evaluation and meta-evaluation* (Scriven, 1999).

While definitions of big six are straightforward, special significance is attached to *program evaluation*. It can be defined as carefully collecting information about a program, or some aspects of it, to make value judgements about the program. It can include any combination of over 35 different types of evaluation, such as needs assessments, accreditation, cost/benefit analysis, effectiveness, efficiency, formative, summative, goal-based, process, outcomes, etc. (Authenticity Consulting LLC, 2010).

Intradisciplinary evaluation of a discipline is the evaluation of theories, hypotheses, classifications, data, research designs and results, practitioners, contributions, and journals and of the discipline itself.

Meta-evaluation is the evaluation of evaluations, in keeping with the reflexive nature of evaluation and reflective integrity of its practitioners. While the first is backbone of all disciplines, the second is backbone of evaluation as it makes it consistent by practicing what it preaches (Scriven, 1999).

Further discussions in this review will be centred around program evaluation.

2.2.2 Types based on purpose

Based on the expected outcome from an evaluation study, it is classified into 5 categories. These are formative, summative, process, outcomes, and impact evaluations (University of Minesota, 2018).

Formative evaluation is done during program formulation for refining program and discusses strategies for early phases of a program.

Summative evaluation judges program effectiveness and is used to decide upon continuing/ expanding a program, it is typically done after completion of the program.

Process evaluation focuses on the implementation of a program and addresses inefficiencies and checks whether delivery happened as planned.

Outcomes evaluation measures changes in short and long term due to the program and establishes benefits.

Impact evaluation focuses on long term, sustained changes as a result of a program, intended and unintended effects, positive and negative effects are estimated using longitudinal studies with comparison groups. This is used to influence policies.

As can be seen, these evaluations with different purposes are done at different times in program cycle. This creates a separate classification based on timing as in Section 2.2.3.

2.2.3 Types based on timing

Timing of evaluation is a relevant classification for program evaluations, which are spread out in time with a definite start and end, though it can be extended to others. Proposal evaluation, in a way, is evaluation of planned program before the beginning. Typically, program evaluations are mid-term and terminal, as explained by the terms, though longer programs may have more stages. For effective evaluations, *Baseline study* is also necessary.

United Nations Environment Program defines three types of project evaluations which examine relevance, effectiveness, efficiency, sustainability, and impact of a project based on the time in project cycle (UNEP, 2008) as below:

Mid-term evaluations are undertaken half-way through project implementation. These analyse whether the project is on-track, problems and challenges encountered, and proposes corrective actions.

Terminal Evaluations are done when a project ends to assess project performance and determine the outcomes and impacts. They offer judgments on actual and potential project impacts, their sustainability, and operational efficiency of implementation. Terminal evaluations are important feedbacks/ lessons for future projects.

Self-evaluations (or concurrent evaluations) are assessments carried out by staff managing implementation. These monitor and report achievement of results, status and challenges in implementation, budget issues, inclusivity issues, sustainability arrangements, impact, and risks.

Needs Assessments are done at the start of a program, with the community to understand the local context, institutions, stakeholders, processes, problems, and possible solutions.

Figure 2.1 presents the above three ways of classifying evaluation and the various types under them. Most of the types in the latter two classifications are part of program evaluation,

though there are certain overlaps with other fields of evaluation too. Just as important as the classification of evaluations, is the classification of evaluation theories presented next.
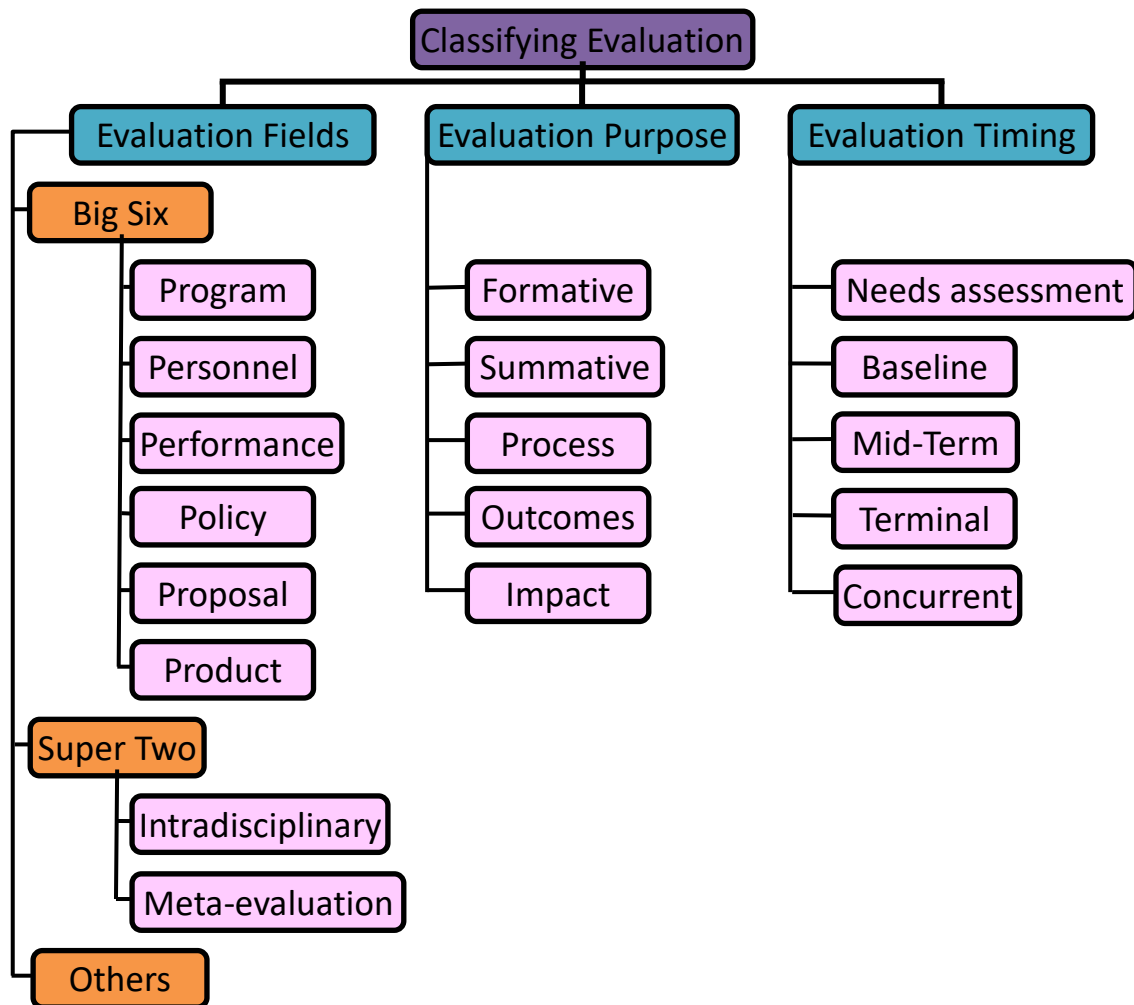
Figure 2.1: Summary of Evaluation Classifications mentioned in the text

2.3 Evaluation Theories and Their Classification

In context of evaluation, the term 'theory' is not used as defined by traditional academic standards. Though the term is conventionally used in evaluation literature, the terms models, approaches, or frameworks better convey the meaning. Evaluation 'theories' are typically prescriptive; they offer a set of rules, prescriptions, and prohibitions that define what a good evaluation study is and how it should be conducted, these aren't predictive or offer an empirical theory. As a convention in evaluation literature, the prescriptive approaches are termed as theories (Christie & Alkin, 2008), and this meaning is used here.

2.3.1 Evaluation theory classification

Christie and Alkin present an evaluation theory tree as a framework to show relation between the various prescriptive theories. The dual foundation of the field of program evaluation is accountability and social enquiry. The need and desire for accountability creates a need for evaluation. Accountability improves programs and society. Social inquiry emanates from a concern for employing a systematic and justifiable set of methods for determining accountability. While accountability provides the rationale, it is primarily from social inquiry that evaluation models are derived.

Three basic elements are discussed by evaluation theorists: use, methods, and valuing. All theorists are concerned with the *methods* employed in conducting evaluation, they recognize that evaluation involves *valuing* (distinctive feature from other researches) and accept that evaluations will be *used* to influence programs. However, emphasis placed on these dimensions, referred to as 'branches' of evaluation theory tree, differ (Alkin & Christie, 2004).

These three branches, are similar to the five-way classification given by Shadish, Cook and Levinton in their 1991 book 'Foundations of Program Evaluation: Theories of Practice' (which is currently inaccessible to me, but is widely referred to by others). These are Theory of Valuing, Theory of Knowledge, Theory of Social Programming, Theory of Use and Theory of Practise (Tarsilla, 2010). These theories and their sub-theories are debated by various proponents and have shaped up the field of evaluation.

Importance of understanding theories of evaluation

These debates are mostly theoretical, about epistemology and ontology, about assumptions made to construct knowledge, about nature of fundamental concepts like causation, generalization, and truth, and about what those theoretical concerns imply for practice of evaluation. And this evaluation theory defines who evaluators are as it provides the language used by evaluators to communicate with each other. It has produced phrases as formative and summative evaluation, instrumental and conceptual use, enlightenment, cluster evaluation, multicultural validity, realist evaluation, etc. and it gives evaluators an identity and knowledge base different from other professionals (Shadish, 1998).

2.4 Logic of Evaluation

Since evaluation is about determining values, data obtained has to be connected to value judgements through certain *logical sequence of concepts*. These can be written as four steps:

1. selecting criteria of merit, on which evaluand (thing being evaluated) will be judged,
2. setting performance standards of these criteria, comparative or absolute
3. gathering data of evaluand's performance on the criteria relative to the standards,
4. integrating the results into a final value judgment.
    (Shadish, 1998).

Although these seem simple and universal, passionate discussions happen on various nuances, enriching the subject. For example, debate on how to determine outcomes in program evaluation brought to life entire literature on causation and determines expenditure of billions of dollars. (Scriven, 2007)

> *"...it may be enough to suggest to people seriously interested in the domains of critical thinking/ informal logic, and the philosophy of the social sciences, the serious possibility that evaluation now constitutes an area in which some mapping has been done, some value has resulted, and important work remains unfinished."*

- (Scriven, 2007)

Scriven divides evaluation approaches into holistic and analytic; the latter is further divided into three: dimensional, component, and theory-driven evaluation (Scriven, 1999). *Holistic approach* integrates all factors, inclusive of environmental, economic, and social, at a systems level rather than individually (Jorjani, 1994). *Analytic evaluation* breaks down evaluation into dimensions or components and treats them separately. *Dimensional program evaluation* focuses on specific dimensions, such as effectiveness, efficiency, cost-utility, equity, sustainability, etc. Today, multi-dimensional evaluations are performed (Reinke, 1999).

A *theory-driven evaluation* uses an explicit theory of change to measure changes from an intervention and attributing the changes to the intervention. It is a conceptual analytical model and not a specific method or technique. It provides a method for structuring and analysing an evaluation (CEE, 2012). Also known as theory-based, theory-anchored, theory-oriented evaluation, it is aimed at reinforcing the explanatory power of evaluation.

The previous two sections lightly touch upon the subjects of intense discussion amongst evaluators, that produce theories, methods, and prescriptions for a good evaluation, which guide evaluation practice, central to evaluation. Evaluation theory and practice have a symbiotic relationship, while theory exists because evaluation is practiced, practice without theory would be scattered and ill-defined (Shadish, 1998).

2.5 Development Evaluation

Development evaluation refers to evaluation of development programs, which are defined as any social or economic programs aimed at community development. These are generally funded by the government, aid agencies, Non-Governmental Organisations, foundations, or Firms under Corporate Social Responsibility. Invariably, most of them require that their programs be evaluated to understand if the intervention succeeded or failed (Bamberger, 2000). As noted earlier, actions towards sustainability agenda are also to be evaluated for understanding what works and why, and if not then why not?

2.5.1 Importance of evaluation

While development evaluation may be mandated, agencies implementing the program wish to avoid evaluations, typically with following concerns, which can be addressed with arguments in parentheses (adopted from (Metz, 2007)):

- Evaluation diverts resources from core program activities (up to 20% of total, but useful to decide how to spend the rest of money)
- Evaluation is a complex process (evaluation design depends on evaluation questions and program resources)
- Evaluation burdens staff (evaluation can be part of management activities or outsourced)
- Evaluation gives negative result (knowing what doesn't work is as important as what works)
- Evaluation is just another form of program monitoring (monitoring assesses compliance while evaluation assesses if expected outcomes were achieved)

Beyond these concerns, there are other reasons which make evaluations important:

- Knowing program effectiveness and what works/ does not work allows improvement in program delivery while not knowing wastes time and resources
- Evaluation showcases effectiveness to community and funders
- Evaluation improves staff's frontline practice with community
- Evaluation allows critical self-assessment and improvement
- Evaluation can build knowledge for activities out of program ambit.

It can be seen that in spite of certain concerns, evaluations of programs, especially those targeted at community development, is essential.

Chapter 3   Evolution of Evaluation

"Evaluation is a very young discipline - although it is a very old practice."

- (Scriven, 1996)

In this chapter, an overview of the global development evaluation scenario is presented. To understand the current scenario, an idea of the historical development is necessary. Hence, it is presented next. The second section of this chapter provides a glimpse of the Indian context.

3.1 Global Picture

This section describes how evaluation evolved as a field and important organisations and journals in the field of evaluation at the international level. Most of the published literature in this field comes from United States of America and some from Europe, thus there is a clear western bias in the documentation of history and important organisations and journals.

3.1.1 History of evaluation

History of evaluation is as old as human activity, Humans (a) identify a problem/ issue, (b) devise alternatives to tackle it, (c) evaluate the alternatives, and then (d) adopt those that results suggest will reduce the problem satisfactorily (Shadish & Luellen, 2005). Shadish and Luellen give examples of earliest documented evaluations from personnel evaluation in China over 4000 years ago and evaluation of Hebrew diet in Bible.

Program evaluation is divided into 7 development periods in the western world. First, the period prior to 1900, the Age of Reform; second, 1900-1930, the Age of Efficiency; third, 1930-1945, called the Tylerian Age; fourth, 1946 -1957, called the Age of Innocence; fifth, 1958-1972, the Age of Development; sixth, 1973-1983, the Age of Professionalization; and seventh, from 1983-2000 the Age of Expansion and Integration (Hogan, 2007).

In *Age of Reforms*, earliest documented evaluations are educational and production processes. In *Age of Efficiency*, scientific management based on observation, measurement,

analysis, and efficiency became prominent, objective-based tests were used to determine quality of educational instruction. In *the Tylerian Age*, criterion-referenced testing based on internal comparison of objectives and outcomes was started. World War II was followed by a period of great growth when accountability of national expenditure was ignored, thus this period is labelled as *Age of Innocence*. Till this period, most literature on evaluation is on educational evaluation. In USA, with the Elementary & Secondary Education Act introducing supplementary programs to support education of disadvantaged students, program evaluation as we know started in the *Age of Development*. In the *Age of Professionalisation*, many journals and university courses on evaluation were started and evaluation established as a formal independent professional field. With increase in aid funding, in the *Age of Expansion and Integration*, professional associations and evaluation standards were established (Hogan, 2007).

In the new millennium, the focus is on capacity development and building institutions for evaluation where organisations like United Nations Evaluation Group and World Bank play a major role. Instead of multiple agencies following multiple standards, there is a move towards a consultative standardisation. This, I am terming as the *Age of Consolidation* (2000-current).

In past few decades, following trends emerged in program evaluation (Hogan, 2007):

- Increased priority and legitimacy of internal evaluation.
- Expanded use of qualitative methods and a shift toward mixed quantitative-qualitative methods instead of depending exclusively on either.
- Increased acceptance of and preference for multiple-method evaluations.
- Introduction and development of theory-based evaluation.
- Increased concern over ethical issues in conducting program evaluations and increased use of evaluation to empower program stakeholders.
- Increased use of program evaluation within business, industry, foundations, and other agencies in the private and non-profit sector.
- Increased options that evaluators should also be advocates for the programs they evaluate.
- Advances in technology, communication, and ethical issues.
- Modifications in evaluation strategies to accommodate increasing trends of government decentralization and delegation of responsibilities to state/provinces and localities.

3.1.2 International organisations in evaluation

In the field of development program evaluation, few organisations are widely recognised. These organisations, by their affiliation, are the leaders in the industry. These are:

- United Nations Evaluation Group, a platform for United Nations Evaluation Offices across units
- Independent Evaluations Group of World Bank
- International Organisation for Cooperation in Evaluation
- International Development Evaluation Association
- American Evaluation Association
- European Evaluation Society

The first two make the evaluating agencies for a large amount of development aid, while IOCE and IDEAS bring together different evaluation organisations. The last two are academic bodies which bring together the leading evaluation practitioners and theorists in the world.

United Nations Evaluation Group was first established in January 1984 as the 'Inter-Agency Working Group on Evaluation' (IAWG), a part of UN consultative group on programme questions (CCPOQ). This is a group of heads of UN evaluation offices to discuss system wide evaluation issues. UNEG's initial work was on designing, testing, and introducing monitoring and evaluation system for UN operations across specialised agencies, funds, programmes, and affiliated organisations. UN Development Programme (UNDP), which funded most UN operations, provided the secretariat and leadership for the Group. It was renamed to UNEG in 2003 (UNEG Secretariat, 2008). UN also has an Office of Internal Oversight Services, established in 1994 by the General Assembly. The office assists the Secretary-General in his oversight responsibilities in respect of the resources and staff of the organization through the audit, investigation, inspection, and evaluation (OIOS, 2018).

The *Independent Evaluation Group* (IEG) is independent of the Management of World Bank Group and reports directly to the Executive Board (IEG, 2018). It is charged with objectively evaluating the activities of International Bank for Reconstruction and Development (IBRD), International Development Association (IDA; together, the World Bank), work of International Finance Corporation (IFC), and Multilateral Investment Guarantee Agency's (MIGA) guarantee projects and services to provide accountability, course corrections, and avoid repetition of past mistakes in meeting the agenda of making the world poverty free.

World Bank project evaluations began in 1970 through Operations Evaluation Unit in Programming and Budgeting Department. In 1973, it was renamed the Operations Evaluation Department, and became independent from bank management. IFC established an evaluation unit in 1984, and in 1995 the unit increased its independence and was renamed as Operations Evaluation Group. MIGA created an evaluation office in 2002. In 2006 the Board of the Bank Group integrated these into a single unit, Independent Evaluation Group (Wikipedia, 2017).

The *International Organisation for Cooperation in Evaluation* is a UNEG supported moment that represents international, national, sub-national, and regional Voluntary Organizations for Professional Evaluation (VOPEs). It strengthens international evaluation through the exchange of evaluation methods and promotes good governance and recognition of the value evaluation has in improving peoples' lives (IOCE, 2018). The EvalPartners group, managed by UNICEF and IOCE, is supported by various partners, including DevInfo, IDEAS, UN Women, UNEG, UNDP, ILO, IDRC, Rockfeller Foundation, Better Evaluation, ReLAC, Preval, Agencia Brasileira de Avaliacao, SLEvA and IPEN, all working together for SDG evaluation (Eval Partners, 2017).

International Development Evaluation Association was established in 2002 as a global professional association for active development evaluators. It aims to improve and extend the practice of development evaluation by refining knowledge, strengthening capacity, and expanding networks, especially in developing countries (IDEAS, 2018).

American Evaluation Association (1986) and *European Evaluation Association* (1992) were established, to promote evaluation use and enrich its theory and practice in the two continents.

3.1.3 Global Evaluation Agenda (GEA) 2016-2020

To support monitoring and evaluation for achieving the 2030 Agenda for Sustainable Development, United Nations adopted the resolution 69/237 on 19th December 2014 for "building capacity for the evaluation of development activities at the country level". This was a step towards building global cooperation for evaluation, year 2015 being already declared as the International Year of Evaluation (EvalYear) at the 3rd International Conference on National Evaluation Capacities at São Paulo, Brazil, in September 2013. The idea behind this was to

advocate and promote evaluation and evidence-based policy making at international, regional, national, and local levels (EvalPartners, 2016).

The Global Evaluation Agenda (GEA) 2016-2020 is the first ever long-term global vision for evaluation. The GEA was developed by many global collaboration, under the EvalPartners umbrella. The discussions around evaluation capacities and capabilities intensified during the Year of Evaluation in 2015, celebrated at 92-plus events around the world. The Year of Evaluation culminated in a historic global gathering hosted by the Parliament of Nepal in Kathmandu where the GEA was launched and endorsed by various stakeholders including Governments, Parliaments, civil society, and academia, in an atmosphere of global solidarity and partnership (EvalPartners, 2016).

EvalAgenda2020 envisions to strengthen the four essential dimensions of the evaluation system, enabling environment for evaluation, institutional capacities, individual capacities for evaluation, and inter-linkages among these first three dimensions (EvalPartners, 2016).

3.1.4 Journals in evaluation

Internationally, there are many renowned journals in the field of evaluation. The following table gives a list of the few important journals:

Table 3.1: Important Journals in Evaluation

	Journal Name	**Publisher**	**Periodicity**	**Homepage link**
1	Evaluation and Program Planning – An international Journal	Elsevier	Bimonthly	https://www.journals.elsevier.com/evaluation-and-program-planning/
2	Evaluation review – A journal of applied social research	SAGE publishers	Bimonthly	http://journals.sagepub.com/home/erxb
3	American Journal of Evaluation	AEA & SAGE	Quarterly	http://journals.sagepub.com/home/ajec

4	New Directions for Evaluation	AEA & Wiley	Quarterly	https://onlinelibrary.wiley.com/journal/1534875X
5	Evaluation – The international journal of theory, research, & practice	The Tavistock Institute & SAGE	Quarterly	http://journals.sagepub.com/home/evi
6	Evaluation Journal of Australasia	Australasian Eval. Soc. Ltd. & SAGE	Quarterly	http://journals.sagepub.com/home/evja
7	Journal of Multi-disciplinary evaluation	PKP	Half-yearly	http://journals.sfu.ca/jmde/index.php/jmde_1/index
8	African Evaluation Journal	AOSIS publishing	Yearly	https://aejonline.org/index.php/aej
9	The Canadian Journal of Program Evaluation	University of Toranto Press and Canadian Evaluation Society	thrice a year	https://www.utpjournals.press/loi/cjpe
10	Journal of Development Effectiveness	I3E, Taylor & Francis Inc.	Quarterly	http://www.3ieimpact.org/en/publications/journal/

This list has been compiled by visiting websites of the major publishing houses, online resources and from references in reviewed literature. Any journals which did not primarily deal with evaluation, evaluation theories and practices or program evaluation have been excluded. Journals focusing on educational evaluation have been excluded. This is not an exhaustive list.

While this is a broad overview of the global history and scenario, of more interest to us is the situation in India. This is covered in the next section.

3.2 Development Evaluation in Independent India

System of evaluation was conceived in India simultaneously with planned economy. With the launch of first five-year plan in 1951, a need for systemic evaluation was felt, and the first plan deemed that systematic evaluation should become a normal administrative practice in all spheres of public activity and for this the Planning Commission (PC) began developing the evaluation techniques by establishing Program Evaluation Organisation (PEO) for independent evaluations of community projects and other intensive area development programmes (Chandrasekar, 2015). From there, India has come a long way over the past 67 years.

Dr S. Chandrasekar served as the Director of Regional Evaluation Office, at Chennai and then as Adviser at Directorate of Economics and Statistics, Ministry of Agriculture, New Delhi. He wrote an article about history of Development Evaluation in India, published as a web special by Yojana in November 2015, around the time when a lot of changes were happening in the Indian evaluation scenario. Most of this section is based on his article and a report by World Bank on M&E system in India (Chandrasekar, 2015) and (Mehrotra, 2013).

3.2.1 Evolution of evaluation institutions in India

The history of institutionalised development program evaluation can be divided into following phases, based on how the Government of India treated its evaluation organisations:

1. Planned economy phase 1952- 1973
2. Neglect phase 1973-1995
3. Resurgence phase 1995-2013
4. New institutions and paradigm phase 2013-current

Planned economy phase 1952-1973

The PEO was established in October 1952 as an independent organisation under the Planning Commission to evaluate development programs implemented in the first five-year plan and bring out their successes and failures through reports. Over the first four five-year plans, PEO activities expanded considerably and most states established their evaluation units in the sixties, for state level programs for cross-verification and learning in tandem with PEO. The scope of PEO extended to include plan schemes/ programmes in sectors of health,

agriculture and cooperation, rural industries, fisheries, family welfare, rural development, rural electrification, public distribution system, tribal development, social forestry etc. Later, PEO also evaluated Centrally Sponsored Schemes (CSS) (Chandrasekar, 2015).

PEO, a field-based organisation, had three-tiered structure – Headquarters in New Delhi at higher level, 3 Regional Evaluation Offices at middle level and 20 Project Evaluation Offices at lowest level. Beyond these were the state offices, taking the total offices to 40 and staff strength to over 500. PEO had relative autonomy as all its offices and the state evaluation offices reported to the Director, PEO. The evaluation reports were a major part of annual conference of State Development Commissioners, enabling follow up actions (Mehrotra, 2013).

Neglect phase 1973-1995

With the reduction in scope of planning commission activities in early seventies on the recommendations of the Administrative Reforms Commission, PEO started its phase of decline and neglect. While the extent of its work was expanded to include urban areas too, its scope of evaluations was reduced to operational, financial, and administrative aspects of schemes and programs, rather than the overall design of programs and their impacts. It was recommended that only those studies should be taken up which could be made available quickly for use by line divisions. This was accompanied by appointment of Indian Economic Service Officers, who are generalists compared to earlier subject specialist academicians, as the head of PEO.

Internal PEO functions were merged with Planning Commission in April 1973, reducing it to a division within a department (Chandrasekar, 2015). Around the same time, based on recommendations of Staff Inspection Unit of Ministry of Finance, field offices were reduced from 40 to 27 by the end of the seventies (Mehrotra, 2013).

PEO featured briefly in latter plans and received insufficient financial layouts, limiting its ability to bring out good reports on time. Its reports were delayed, didn't cover program impact & design anymore, and were given less important by the concerned ministry thus, the reducing their use. This in turn reduced the number of studies being done (Chandrasekar, 2015).

Resurgence phase (1995-2013)

The resurgence in demand for evaluation can be traced to the late nineties, when the Planning Commission got involved in design and implementation of social safety net programs

to counter the adverse effects of economic reforms initiated earlier. Unfortunately, the Fiscal Responsibility and Budget Management Act 2003 ensured that the PEO and its field offices were highly understaffed. This began the practise of outsourcing the studies to social science research institutes. The PEO involved the ministries and subject matter expert groups in ensuring some actions were taken based on its reports from the ninth plan onwards (1997-2002)

The eleventh five-year plan 2007-2012, stressed on building online MIS for all flagship programs. Development monitoring unit was setup in Prime Minister's Office in 2009, and a Performance Monitoring and Evaluation System (PMES) was created at the cabinet secretariat. The functions of monitoring and evaluation were being mixed together. A scheme named Strengthening Evaluation Capacity was launched in 2006-07, to reduce the financial problems at PEO but it did little to address the administrative and staff problems (Chandrasekar, 2015).

During this phase of resurgence in demand for evaluation activities, mixing up of monitoring and evaluation, ignoring plight of PEO, underutilisation of studies, and outsourcing to private institutions without clear policy, were a few grave mistakes made. As a result, in 2012, there were 6 regional and 8 project offices left (PEO, 2012).

New institutions and paradigms phase (2013-current)

A new Independent Evaluation Office was established in the 12th plan with a mandate to "conduct evaluation of plan programmes, especially the large flagship programmes to assess their effectiveness, relevance and impact. It also has the freedom to conduct independent evaluations on any programme which has access to public funding or implicit or explicit guarantee from the government." Instead of using regular organised services available to government, it proposes to get evaluation done by selected institutes and researchers identified through tender processes (Chandrasekar, 2015). Not much is known about how IEO was expected to function and how it was different from the PEO.

With the change in regime and dissolution of Planning Commission in 2014, PEO and IEO have been merged into Development Monitoring and Evaluation Office (DMEO) in September 2015. In 2017, most field offices were shut down and staff was attached to DMEO at New Delhi (Indian Express, 2017). Even less details are available on official websites about this office compared to PEO (and IEO). The PMES started earlier is now replaced by Pragati dashboard for direct follow-up by PMO for better implementation but this misses any

opportunity for evaluations based on the Results Framework documents prepared by the ministries (The Economic Times, 2015).

3.2.2 Concurrent evaluations

In the resurgence phase, concurrent evaluations were being regularly done by ministries themselves for their programs. For example, National Food Security Mission under Department of Agriculture and Cooperation, Ministry of Agriculture was carrying out its own concurrent evaluations in 2010 (NFSM Cell, 2010) and Ministry of Rural Development had a Concurrent Evaluation Office (CEO), set up for managing Concurrent Evaluation Network (CENET) of MoRD, in conjunction with IEO. The CEO was closed in July 2016 (PIB, 2016).

Concurrent evaluation is either a formative or process evaluation, which evaluates all the activities carried out to achieve program objectives, annually. Concurrent evaluations have been done in the past too, an example is the concurrent evaluation of Integrated Rural Development Program carried out by Department of Rural Development, Ministry of Agriculture in 36 districts of the country since October 1985 for at least a year. As ordinary evaluations in that era were usually ex post facto, they did not provide remedial measures and mid-term collections, a need for concurrent evaluation was felt. (Saxena, 1987)

The term concurrent evaluation isn't common outside India, where the term self-evaluations is used for internal, regular evaluations (UNEP, 2008).

3.3 Current Scenario

Past decade has been very eventful for the evaluation systems in India. IEO was set up and closed, PEO was closed, Results Framework Diagram based PMES was started and closed and DMEO has been started recently. This section captures the current scenario at the central and state level in India.

3.3.1 DMEO at NITI Aayog, New Delhi

While Development Monitoring and Evaluation Office (DMEO) has been established in 2015 and NITI Aayog has a very functional and updated website, very little information is available about it, in the Digital India age. The little information available is from a few newspaper articles and telephone book of NITI Aayog.

While the 2016 contacts document mentions 7 regional DME offices and 8 Project DME offices, the 2018 document mentions no regional or project offices (NITI Aayog, 2018). This change is also hinted at in news in 2017 which mentions that the 15 offices are being shut down and staff called to headquarters in Delhi (Indian Express, 2017).

In the current set up, DMEO has a Director General at helm, a Joint Secretory, two Deputy DGs, an under Secretory and staff attached to their offices. On the Technical/ specialist end, there are a few senior Research Officers, Sr. Statistical Officers, a Senior Consultant and many Economics Officers, Consultants, Research Associates and Young Professionals, a total of about 25-26 people. There is some administrative staff as well (NITI Aayog, 2018).

In 2016, DMEO called for Expression of Interest by Research Institutions, NGOs, and universities for carrying out evaluation studies. While this call for EoI is available online, the final list is not found on the NITI Aayog website. As per mandate of DMEO, it is expected to get evaluation studies done as requested by various ministries for their programs. This is similar to what PEO and IEO were doing.

3.3.2 Evaluation in Indian states

Evaluation was an integral component of every state's planning and implementation process while PEO was blooming. States have taken varied path in past few decades from there.

While Evaluation is reported just as an activity under the Directorate of Economics and Statistics in Planning Department in most states, Karnataka has an Evaluation authority, in Goa and Sikkim, Evaluation is in the name of the directorate. When we look at the official websites, we see that evaluation occupies important position in many states. Table 3.2 gives a brief overview of each state.

It is seen that across the states, evaluation is a function generally under the Planning Department, which has the Directorate of Economics & Statistics, responsible for all statistical data collection, analysis, and in most states, for monitoring and evaluation functions. Most of these functions started during the third plan period (1961-66) (PEO, 2006).

Outsourcing of evaluation studies to competent agencies has been going on for a couple of decades and the websites, developed in last 10 years mostly, show records of processes carried out by various states since 2012-13, under the 12th Five-year plan. Unlike Maharashtra though, very few states refer to the UN guidelines in their empanelment Process.

Records of how the feedback generated by these studies is used is poor. Program Evaluation Organisation had brought out one study in 2004 and another in 2006 titled Development Evaluation in PEO and Its Impact (Vol I and Vol II) which summarise the follow up actions taken based on the evaluation studies done in the preceding years (PEO, 2006). Beyond this, not much is documented.

Table 3.2: Evaluation setup in State governments

Sr	State	Name of Agency	Website	Remark
1	Andhra Pradesh	not found		
2	Arunachal Pradesh	Planning Department	http://www.arunachalplan.gov.in/html/docs/1_intro_planning.htm	evaluation fleeting mention
3	Assam	Evaluation function Transformation and Development Department	https://transdev.assam.gov.in/information-services/evaluation-and-monitoring	no studies after 2009, only monitoring of ongoing projects
4	Bihar	Directorate of Evaluation Bihar, Planning and Development Dept.	http://planning.bih.nic.in/	No details available on the website
5	Chhattisgarh	nil		
6	Goa	Evaluation Division of Directorate of Planning, Statistics and Evaluation	http://www.goadpse.gov.in/	No details available on the website
7	Gujarat	Directorate of Evaluation, General Administration Dept. (Planning)	https://gujecostat.gujarat.gov.in/director-of-evaluation1	Evaluation committees, guidelines
8	Haryana	Evaluation function Planning wing, Dept. of Economics and Statistical Analysis	http://esaharyana.gov.in/en-us/Plan-Evaluation	No details available on the website
9	Himachal Pradesh	Planning Department	http://hpplanning.nic.in/Reports.htm	few evaluation reports
10	Jammu & Kashmir	Evaluation Division, Directorate of Economics & Statistics	http://ecostatjk.nic.in/divisions/divisions.htm	No details available on the website
11	Jharkhand	nil		
12	Karnataka	Karnataka Evaluation Authority	http://kea.karnataka.gov.in/	State evaluation policy, empanelled institutes, studies
13	Kerala	Department of Economics and Statistics	http://www.ecostat.kerala.gov.in/	No details available on the website
14	Madhya Pradesh	nil		
15	Maharashtra	Evaluation Division, Directorate of Economics & Statistics, Planning Dept.	https://mahades.maharashtra.gov.in	Empanelment of institutes for evaluation studies

16	Manipur	not found		
17	Meghalaya	Program Implementation & Evaluation Dept.	http://megpied.gov.in/	Reports since '70s, tender for EOI
18	Mizoram	Research & Analysis Division, Planning & Programme Implementation Dept.	https://planning.mizoram.gov.in/page/research-analysis	No details available on the website
19	Nagaland	Directorate of Evaluation, Planning and Coordination department	http://evaluation.nagaland.gov.in/	reports from 1973 to 2013
20	Odisha	State Evaluation Unit, Planning Branch, Department of Planning and Convergence	http://pc.odisha.gov.in/DepartmentActivities.asp?lnk=2&Pl=2	Kalahandi Balangir Koraput evaluation reports
21	Punjab	Department of Planning, Govt of Punjab	http://www.pbplanning.gov.in/aboutus.html	
22	Rajasthan	Directorate of Evaluation Organisation Rajasthan	http://www.plan.rajasthan.gov.in/evaluation/	over 500 evaluation studies
23	Sikkim	Department of Economics, Statistics, Monitoring & Evaluation	http://www.desme.in/	No details available on the website
24	Tamil Nadu	Directorate of Evaluation & Applied Research TN	http://www.tn.gov.in/dear/	No details available on the website
25	Telangana	Evaluation Authority of State of Telangana (EAST)	www.telangana.gov.in/departments/planning	Evaluation of state focus programs
26	Tripura	not found		
27	Uttarakhand	not found		
28	Uttar Pradesh	Planning Department	http://planning.up.nic.in/	No details available on the website
29	West Bengal	Evaluation, Monitoring & Manpower Branch, Dept. of Planning	http://wbplan.gov.in/	No details available on the website

Note: Union territories have been left out. Delhi Government is planning to create some setup for evaluation, as details are not available beyond a news article, it has also not been included in this table. The states with 'not found' in Name of Agency column may have some evaluation office, but its details are not easily available online. Following up on status, activities, capacities, etc. of each of these agencies is a potential research topic.

Chapter 4   Building Evaluation Capacity

Previous chapter outlined the changes in development evaluation function in Government of India and its status. This chapter discusses the various aspects currently in focus to build capacity for effective evaluations. The movement for Building Evaluation Capacity is being pushed through IOCE and UNEG through EvalPartners. Primary processes stressed are creation of National Evaluation Policies (NEP) by involving parliamentarians and voluntary groups. National Evaluation Capacity building program is committed to develop and implement transparent results-based M&E framework to track the efforts and results of the implemented commitments proposed (Commitment 1 or **C1**).

4.1 National Evaluation Capacities Commitments

National Evaluation Capacity building brings together 60 countries to share solutions to challenges related to evaluation independence, credibility, and use. During the third NEC Conference in 2013, the participants developed and signed 18 commitments to enhance national evaluation capacities and to encourage accountability by calling participant countries to commit to actions and collaboration (IEO, 2015).

These 18 commitments are given as documented in the official documents:

C2: Build and strengthen **national data systems**, improve integrity of such systems to better link performance of policies and programmes.

C11: Develop/ connect national registries/ national statistical systems to M&E systems with increased frequency of data collection to support decision-making.

C3: Develop systems for transparent follow-up of evaluations, for effective monitoring of the **implementation of evaluation recommendations**.

C4: Study alternatives, assessing pro and cons of different options of **institutional set-ups**, like national evaluation legislation/ policies, accounting for local context and establishing a set of minimum requirements.

C5: Develop/ strengthen/ support/ expand joint peer-to-peer systems and mentoring programmes among professional association of evaluators and government evaluation units.

C6: Create/ strengthen Parliamentarians' Forum for Development Evaluation in various regions to **advocate for use and conduct of evaluations**.

C7: Facilitate partnership/ cooperation between governments, VOPEs, parliaments, and private sector to strengthen the understanding about evaluation and how to use it.

C8: Develop approaches to **incorporate cultural dimension**s into evaluation in different regional and national contexts.

C9: Develop **methodologies** or standards for triangulation of evidence, checks and balances, and qualitative data use.

C10: Develop standards to ensure **stakeholders involvement**.

C14: Map and analyse effectiveness of **coordination** mechanisms and practices between central evaluation units, sector/ ministry units and local government evaluation.

C16: Budgetary provision for evaluations while designing/ approving projects/ programmes/ policies or assign a percentage of the initiative cost.

C17: Use **independent evaluators** to facilitate/ moderate self-assessments and reviews.

C18: Incorporate **gender** capacities/ perspectives in M&E national systems.

Internal Commitments

These focus on how participants will follow up on NEC.

C1: Develop and implement framework to **track progress** in commitments.

C12: Online platform (NEC-COP) for sharing, connecting, and following up

C13: Translate material on evaluation into different languages.

C15: Support joint regional/ national events to take stock of progress.

It can be seen that serious push is being given to build capacity for evaluation and universalise certain aspects of it, a logical follow up of the professionalisation of evaluation phase which happened towards the turn of the century. One important starting step is National Evaluation Policy formulation.

4.2 National Evaluation Policy

EvalPartners released 'Status of National Evaluation Policies' Report for 'Parliamentarians Forum on Development Evaluation in South Asia' during EvalYear, in

February 2015. It provides status of 59 countries of 109 examined. India's National Evaluation Policy (NEP) status is also given. India has been developing a NEP since 1954 and since 2009 it was well developed. In June 2009, the President of India, announced the nation's commitment to create a Performance Management and Evaluation System for the national government, comprised of over 4 million civil servants in 84 departments and ministries (Rosenstein, 2015). PMES was established and later closed as mentioned earlier, but as the report was a work under completion since 2013, it didn't reflect this change. About the same time, at Evalweek in January 2015, Planning Minister Rao Inderjit Singh said that though India had a long history of Monitoring and Evaluation, a NEP was still not in existence and it was about time that it was developed to provide a direction to M&E activities in the country, with stress upon quality standards and sound, ethical procedures, and provide for appropriate institutional mechanisms (The Economic Times, 2015). Even today though, a search for National Evaluation Policy for India gives no relevant results. Interestingly, Karnataka has an evaluation policy.

Looking at the recent developments, it can be assumed that India, which is part of the aforementioned parliamentarians' forum, Evalpartners and organised events during the Evaluation event, is responding to the global clarion call for developing capacity for evaluation and carrying out development evaluations for Agenda 2030. Establishment of DMEO by merging PEO and IEO is a first step towards a new approach to evaluation through partner agencies and outsourcing. For outsourcing of evaluation studies, instead of carrying out complete process for each tender, agencies have been empanelled at different levels.

4.3 Volunteer Organisations for Professional Evaluations

Volunteer Organisations for Professional Evaluations (VOPEs) are evaluation non-profit membership organizations open to individuals interested in evaluation. Evaluation is a critical function for managers and policy makers in every sector, as it promotes clear and strategic thinking, learning, and accountability. VOPEs are important as professional networks or guilds where evaluators and commissioners of evaluation can meet and discuss issues of mutual interest. VOPEs help governments and civil society in their countries to understand what evaluation is, and the role evaluation can play in supporting better informed public policy and decision making. IOCE offers some support to various VOPEs, including international visibility, updates, toolkits, etc.

While various VOPEs are present globally, including American Evaluation Association, Canadian Evaluation Society, Brazilian Network, etc., from Indian Perspective, the Community of Evaluators, South Asia, with its head-office in New Delhi is most important. Development Evaluation Society of India (DESI), founded in 2002 is another VOPE registered in India, with very limited online presence today.

4.3.1 Community of Evaluators, South Asia

The Community of Evaluators is a Section 25 company registered in India and a consortium of evaluators from South Asia working together to strengthen the field of evaluation. The largest network of South-Asian evaluators across 8 countries, it provides opportunities for knowledge sharing, capacity development, networking advocacy, developing protocols, and standards for evaluation in member countries: Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka.

4.3.2 Forum on Development Evaluation

Forum on Development Evaluation is a collective of parliamentarians who are committed to the development of evaluation in South Asian Association for Regional Cooperation (SAARC) countries. The parliamentarians-forum's goal is to advance enabling environments for nationally owned, transparent, systematic, and standard development evaluation processes in line with National Evaluation policies at the country level, to ensure aid effectiveness, achievement of results, and development sustainability (IEO, 2015).

4.4 Building Evaluation Capacity in India

India has a long history of evaluation and strong institutions but unfortunately today monitoring and evaluation is focusing upon target completion only. While the government is outsourcing the evaluation function to non-government agencies, there is no clarity on how evaluation should occur. As noted earlier in this chapter, India is still finalising its National Evaluation Policy but various efforts are being made simultaneously to expand evaluations.

4.4.1 Empanelment for evaluation

Maharashtra government processed empanelment of institutions for evaluation, creating opportunity for professionals to work for the government in evaluation in 2016 triggering interest in evaluation and questions about capacity to evaluate and evaluation frameworks. This section is a short note on empanelment for evaluation.

Maharashtra Government Resolution for empanelment

Government of Maharashtra published a Request for Proposal and empanelled various private consulting firms and educational institutes, including IIT Bombay, for evaluations at state and district level (based on the size of the agency) in June 2016 through a Government Resolution (Planning Department, 2016). This GR was further amended in September 2017 to add a few new agencies and delete some. Also, Maharashtra has been conducting evaluations under this GR at district and state level for past two years. CTARA, IIT Bombay is also performing a couple of these evaluations.

The RFP document refers to UNEG guidelines and thus must have been influenced by it. It is likely that UNDP or UNICEF worked with DES, Maharashtra in this process.

Empanelment in other states

As noted before, DMEO had also called for applications for empanelment of agencies in 2017, though no updates are available. Many other states have had similar process. Karnataka Evaluation Authority has empanelled 30 agencies. Punjab, Gujarat, Meghalaya have documents on their websites which show that they have empanelled agencies. As most of these are from 2012-13, after the PEO guidelines on appointment of evaluation agencies on contract basis, it might have been a common instruction to all states from the centre.

Chapter 5   Conclusion

In this literature review of Development Evaluation, we first looked at development and then elaborated upon evaluation. As the motivation of this report is to understand the field of evaluation, Chapter two explored the basics of the field, as a theoretical perspective is necessary before entering into practise and ability to converse in the language of the field is important for effective communication. Importance development evaluation is established.

To understand a field, it is also important to gain a historic perspective. Documenting history of evaluation at the global level and at India level would be a book in itself, still a primary understanding is presented in Chapter three. Comparing the International and Indian timeline, it is interesting to note that India already had strong evaluation institutions and practise in the 1950's and 60's which were lost over the next two decades, while the global scenario was just opening up. United Nations and the World Bank Group were formalising their evaluation set up in the 80's when Indian PEO was waning. While the period of neglect in India ended with the turn of the century, the current period is also not very convincing. Development Monitoring and Evaluation Office is present now for past three years and has had Young Professionals working at it and still, in the age of Digital India, it doesn't have its own web-page. The scene is some states is reassuring though.

For better monitoring and evaluation practices, capacity development is necessary. Internationally, the focus is to get nations committed to evaluation and prepare national evaluation policies and have active VOPEs to engage the government and development agencies working in the country.

While this report is an attempt for charting the seas of evaluation for a new explorer, the first steps towards understanding navigation in this sea, i.e., learning how to conduct evaluations, develop M&E plans, etc. are yet to be taken. A vast amount of literature exists on this subject and needs to be internalised. Simultaneously, extending the metaphor, some uncharted seas beckon. There are some studies that can be carried out at India level as very little India specific analytical material on status, history, and direction of evaluation is available.

This lacuna throws up opportunities for some future work:

1. Creating database of government conducted evaluation studies in India, at central, state, and multi-lateral donor agency level.

2. Meta-evaluations based on these evaluation studies

3. Assessing status of evaluation offices in Indian states, exploring their history and current activities, capacities.

4. Assessing draft National Evaluation Policy of India, proposing a state evaluation policy for Maharashtra, if one doesn't exist already.

Beyond these activities, building up understanding of how to evaluate through various guidelines and recent researches on evaluation logic, will allow to create a generalised development evaluation framework and necessary tools for the same (needs assessment framework, theory of change, results framework diagram, SMART indicators, etc.). This review and these future studies will allow me to design and conduct evaluations in a better way.

References

Alkin, M. C., & Christie, C. A. (2004). An Evaluation Theory Tree. In M. C. Alkin (Ed.), *Evaluation Roots* (pp. 12-65). SAGE.

Authenticity Consulting LLC. (2010). *Basic Guide to Program Evaluation*. Retrieved from Free Management Library: https://managementhelp.org/evaluation/program-evaluation-guide.htm

Bamberger, M. (2000). The Evaluation of International Development Programs: A View from the Front. *American Journal of Evaluation*, 95-102.

Barder, O. (2012, May). Development & Complexity. *Kapuściński Lecture*. Retrieved from http://www.owen.org/wp-content/uploads/Development-and-Complexity-Slides.pdf

Barder, O. (2012, August 16). *What Is Development?* Retrieved from Center for Global Development: https://www.cgdev.org/blog/what-development

CEE. (2012). *Theory-Based Approaches to Evaluation: Concepts and Practices.* Ontario: Centre of Excellence for Evaluation, Government of Canada.

Chandrasekar, D. S. (2015, November). *Development Evaluation in India.* Retrieved from Yojana: yojana.gov.in/Development%20Evaluation%20in%20India.pdf

Christie, C., & Alkin, M. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, 131-135. doi:doi:10.1016/j.stueduc.2008.07.001

Eval Partners. (2017). *About Us*. Retrieved from Eval Partners: https://www.evalpartners.org/about/about-us

EvalPartners. (2016). *2015 declared as the International Year of Evaluation*. Retrieved from EvalPartners: https://www.evalpartners.org/evalyear/about

EvalPartners. (2016). *EvalAgenda 2020*. Retrieved from EvalPartners: https://www.evalpartners.org/global-evaluation-agenda

EvalPartners. (2016). *EvalAgenda 2020: Global Evaluation Agenda 2016-2020.* EvalPartners.

Hagen, M. v., & Willems, J. (2010). *Monitoring and Evaluation Handbook for National Action Plans against Trafficking in Human Beings.* Austria: International Centre for Migration Policy Development.

Heider, C. (2015, Sept 15). *Evaluation Beyond 2015: Implications of the SDGs for Evaluation*. Retrieved from IEG, World Bank Group: https://ieg.worldbankgroup.org/blog/evaluation-beyond-2015-implications-sdgs-evaluation

Hogan, R. L. (2007). The Historical Development of Program Evaluation: Exploring The Past and Present. *Online Journal of Workforce Education and Development, 2*(4). Retrieved

from
https://pdfs.semanticscholar.org/ee2f/dbbe116a30ab7a79b19e1033a7cab434feec.pdf

Honkalaskar, V. (2014). *A Participatory, Bottom-up, and Context-appropriate Inquiry for Community Development.* PhD Thesis, IIT Bombay, CTARA, Mumbai.

IDEAS. (2018). *About*. Retrieved from International Development Evaluation Association: https://ideas-global.org/about/

IEG. (2018). *About IEG*. Retrieved from Independent Evaluation Group, World Bank Group: http://ieg.worldbankgroup.org/about-us

IEO. (2015). *Towards a Baseline Study: Insights on National Evaluation Capacities in 43 Countries.* New York: Independent Evaluation Office, UNDP.

Indian Express. (2017, April 24). Modi's Think Tank cracks down on its monitoring offices. *Indian Express*.

IOCE. (2018). *Home*. Retrieved from International Organization for Cooperation in Evaluation: https://www.ioce.net/

Jaszczołt, K., Potkański, T., & Alwasiak, S. (2003). *Internal Project M&E System and Development of Evaluation Capacity – Experience of the World Bank-funded Rural Development Program.* Retrieved from http://ec.europa.eu/regional_policy/archive/sources/docconf/budapeval/work/jaszczolt.doc

Jorjani, H. (1994). The Holistic Perspective in the Evaluation of Public Programs: A Conceptual Framework. *The Canadian Journal of Program Evaluation*, 71-92. Retrieved from evaluationcanada.ca/secure/09-2-071.pdf

Lee, B. (2000). Theories of Evaluation. In R. Stockmann (Ed.), *Evaluationsforschung* (Vol. Vol 1, pp. 127-164). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:https://doi.org/10.1007/978-3-322-92229-8_6

Mehrotra, S. (2013). *The Government Monitoring and Evaluation System in India: A Work in Progress.* Washington D.C.: Independent Evaluation Group, The World Bank.

Metz, A. J. (2007). *Why Conduct A Program Evaluation? Five Reasons Why Evaluation Can Help an Out-of-School Time Program.* Child Trends.

NFSM Cell. (2010). *Handbook on Concurrent Evaluation.* New Delhi: Department of Agriculture and Cooperation. Retrieved from nfsm.gov.in/Circulars_Notifications/Handbook.pdf

NITI Aayog. (2018, April). *Contact Us.* Retrieved from NITI Aayog: http://niti.gov.in/content/contact-us#

OECD. (2016). *Sustainable Development Goals and Evaluation*. Retrieved from OECD: http://www.oecd.org/dac/evaluation/sustainabledevelopmentgoalsandevaluation.htm

OIOS. (2018). *OIOS Home*. Retrieved from OIOS UN: https://oios.un.org/

PEO. (2006). *Development Evaluatio in PEO and its Impact.* New Delhi: Program Evaluation Organisation, Planning Commission.

PEO. (2012). *Guidelines/ Procedures followed by PEO for Conducting Evaluation Studies.* New Delhi: Programme Evaluation Organisation, Planning Commission. Retrieved from planningcommission.nic.in/reports/peoreport/peo/guide_consult1902.pdf

PIB. (2016, July 27). Cabinet approves rescinding the decision of the Government to set up the Concurrent Evaluation Office in Ministry of Rural Development. *Press Information Bureau, Govt of India*. Retrieved from http://pib.nic.in/newsite/PrintRelease.aspx?relid=147856

Planning Department. (2016, July 17). Empanelment of Agencies for 'Conducting Evaluation Studies, Sample Surveys and providing expert services in statistical & economic activities'. Mumbai: Government of Maharashtra.

Reinke, W. A. (1999). A Multi-Dimensional Program Evaluation Model: Considerations of Cost-Effectiveness, Equity, Quality, and Sustainability. *The Canadian Journal of Program Evaluation*, 145-160.

Rosenstein, B. (2015). *Status of National Evaluation Policies: Global Mapping Report.* EvalPartners.

Saxena, A. P. (1987, September 26). Concurrent Evaluation of IRDP: Selected Aspects for Administrative Follow-up. *Economic & Political Weekly*, pp. A-121 - A-124.

Schwandt, T., Ofir, Z., Lucks, D., El-Saddick, K., & D'Errico, S. (2016). *Evaluation: a crucial ingredient for SDG success.* London: International Institute for Environment and Development. Retrieved from http://pubs.iied.org/17357IIED/

Scriven, M. (1996, October 1). The Theory behind Practical Evaluation. *Evaluation, 2*(4), 393-404. doi:https://doi.org/10.1177/135638909600200403

Scriven, M. (1999, September). The Nature of Evaluation Part II: Training. *ERIC/AE Digest Series*. Retrieved from http://ericae.net/digests/tm9907.pdf

Scriven, M. (1999, Sept). The Nature of Evaluation. Part I: Relation to Psychology. *ERIC/AE Digest*. Retrieved from https://files.eric.ed.gov/fulltext/ED435710.pdf

Scriven, M. (2007). The Logic of Evaluation. *OSSA Conference Archive OSSA 7* (p. 17). Ontario: University of Windsor.

Shadish, W. R. (1998). Evaluation Theory is who we are. *American journal of Evaluation, 19*(1), 1-19.

Shadish, W. R., & Luellen, J. K. (2005). History of Evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation* (pp. 184-186). Thousand Oaks: SAGE Publications, Inc. doi:http://dx.doi.org/10.4135/9781412950558.n251

Tarsilla, M. (2010, February). Theorists' Theories of Evaluation: A Conversation with Jennifer Greene. *Journal of MultiDisciplinary Evaluation, 6*(13), 209-219.

TDSAA CTARA. (2017, March 23). A chat with Prof. Ashok Gadgil. Mumbai, Maharashtra, India: YouTube. Retrieved from https://youtu.be/-IGhCt4RrE4?t=14m21s

The Economic Times. (2015, December). Government replaces RFD Model to measure performance. *The Economic Times*.

The Economic Times. (2015, January 19). NITI Aayog to monitor, evaluate government policies: Rao Inderjit Singh, planning minister. *The Economic Times*.

UN Women. (2012). *Monitoring and Evaluation Plans,* . Retrieved from End Violence Against Women and Girls: http://www.endvawnow.org/en/articles/1107-monitoring-and-evaluation-plans.html

UNEG. (2016). *Norms and Standards for Evaluation.* New york: United Nations Evaluation Group. Retrieved from http://www.unevaluation.org/2016-Norms-and-Standards

UNEG Secretariat. (2008). *About UNEG*. Retrieved from United Nations Evaluation Group: http://www.uneval.org/about

UNEP. (2008). *Evaluation Manual.* New York: United Nations Enivronment Program. Retrieved from https://wedocs.unep.org/rest/bitstreams/9758/retrieve

United Nations. (2016). *The Sustainable Development Agenda*. Retrieved from Sustainable Development Goals: http://www.un.org/sustainabledevelopment/development-agenda/

University of Minesota. (2018). *Different Types of Evaluation*. Retrieved from CYFAR: https://cyfar.org/different-types-evaluation

Wikipedia. (2017, June). *Independent Evaluation Group*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Independent_Evaluation_Group