



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Ph.D. Dissertation of Education

Evaluation for Quality Education:

Capability-based Evaluation Framework for

Basic Education Quality

Capability 기반의 교육개발협력

기초교육의 질 평가프레임워크 개발 연구

February 2022

Graduate School of Education

Seoul National University

Global Education Cooperation Major

Eunhye Lee

Evaluation for Quality Education:

Capability-based Evaluation Framework for
Basic Education Quality

Sung-Sang YOO, Ph.D.

**Submitting a Ph.D. Dissertation of
Education**

February 2022

**Graduate School of Education
Seoul National University
Global Education Cooperation Major**

Eunhye Lee

**Confirming the Ph.D. Dissertation written by
Eunhye Lee
February 2022**

Chair	Kevin Kester	(Seal)
Vice Chair	Jae-Kwang Han	(Seal)
Examiner	Jae Eun Lee	(Seal)
Examiner	Da Jung Jung	(Seal)
Examiner	Sung-Sang Yoo	(Seal)

© 2022 by Eunhye Lee

All rights reserved.

Abstract

Evaluation for Quality Education:

Capability-based Evaluation Framework for Basic Education Quality

Eunhye Renée Lee

Global Education Cooperation
Graduate School of Education
Seoul National University

International education development goals such as MDG2, EFA, and SDG4 have helped to provide educational opportunities to thousands of children around the world. And over the last 20 years, the amount of international aid invested in education sector has steadily increased, reaching around \$15.9 billion in 2019. (OECD CRS 2021). A significant quantitative result was achieved in enrolling a large number of children in elementary school through the MDG's second goal, 'Achieving universal primary education.' Since then, the international community has set a more challenging goal (SDG4) to ensure 'inclusive and equitable quality education and promote lifelong learning opportunities for all' based on these achievements (UN 2015).

However, despite steady policy and financial support from the international community, many statistical data predict that changes in international social education indicators after the MDG will stop or even worsen in some regions (UNESCO 2019; UN 2015). In particular, with the exception of the enrollment rate indicator, which was a major measure of achievement of the MDG goal, indicators such as dropout rate and graduation

rate show no room for improvement at all. Assuming that the indicator is one of the tools for capturing the essence of any state or the most appropriate moment of the certain process, the enrollment rate indicator can be considered as a first step in a long educational development process. This is why it is used as the most common and basic indicator when discussing educational development. In addition, it is also classified as Tire I indicator by the IAEG-SDGs based on its methodological development and the availability of data. So, what about other indicators such as dropout rates, graduation rates, or academic achievement rates? What I can say at a common-sense level, even at the risk of overgeneralization, is that these indicators contain a higher level of complexity than the enrollment rate indicators.

This study is not intended to blame those indicators for the slowing of educational development. Rather, it tries to highlight the shortcomings of how fragmentary and result-oriented the interpretation of the indicators of development evaluation to ensure the achievement of specific goals. And, as an alternative to overcoming it, I propose an evaluation framework that places the ‘interaction between students and instructors in the learning space’, which has not been actively discussed so far in development evaluation, in the center of evaluation when evaluating the quality of education.

Indian economic philosopher Amartya Sen, who won the Nobel Prize in Economics for his theory of welfare economy, had a truly revolutionary effect on numerous fields of development cooperation through his Capability Approach in the 1990s. In particular, he made a great contribution to providing a conceptual framework for measuring the well-being, as he extended the development of the country, which had been evaluated only by GDP, to the fields of education and health. The key contribution of the Capability approach as a theoretical framework in conceptualizing the evaluation framework for quality of education presented in this study is that it includes the 'diversity' of individuals, social and environmental contexts as a very important factor in the evaluation. In other words, it reveals that the process and results of individual choices are influenced by a various combination of personal, social, and environmental factors. Sen referred to this as the conversion factor.

A qualitative approach to assessing the quality of education requires a relatively significant amount of time and resources. Nevertheless, the proposal of such an alternative evaluation framework is for the following reasons. First, it is because many educators have already shown a correlation between the qualitative improvement of the education and ‘process’ of teaching and learning. The process here refers to the interaction between the learner and the instructor. Second, it is because the current development evaluation approach completely excludes evaluation of this learning process. In particular, the indicators used in the evaluation of development cooperation projects and programs do not reflect the most important learning process and qualitative change for economic reasons such as ease of data collection and measurability.

The evaluation framework of basic education quality proposed in this dissertation is based on the core value of Amartya Sen's capability approach, and it borrows many theories and approaches that have been used for quality evaluation in the education field. In addition, to assist more comprehensive manner of evaluation of quality education, the ‘Basic Education Quality Index’ is proposed. It is expected the Capability-based evaluation will contribute to find answers to the more fundamental question of evaluating the quality of education through ‘what’ as well as the methodological question of ‘how’ to evaluate. Finally, I wish it is also to bring a more educational perspective in defining and evaluating the quality of basic education in educational development cooperation.

Key words: education development cooperation, development evaluation, quality of education, capability approach

Student Number: 2014-31210

TABLE OF CONTENTS

CHAPTER I. INTRODUCTION	1
1.1. Research Background and Objectives	1
1.1.1 Research Questions	5
1.2 Capability Approach as a Theoretical Methodology	7
1.2.1 Contribution of the Capability Approach as a Research Method	8
1.2.2 Examined Body of Literature	11
1.3 Research Scope and Structure of Dissertation	11
CHAPTER II. Literature Review	14
2.1 Core Concepts of the Capability Approach	14
2.1.1 Functionings and Capabilities	16
2.1.2 Pluralism and Conversion Factors	18
2.1.3 Valuing as Ends and Means	20
2.2 Critiques and the Capability List	21
2.2.1 Capability List	21
2.2.2 Critiques on Capability Approach	24
2.3 Capability and Education	25
2.3.1 Expanding the Right to Education	26
2.3.2 Diversities and Heterogeneities of learners	32
2.3.3 Education as an instrumental value to expand other capabilities	34
CHAPTER III. Quality Issues in Education Development Cooperation	37
3.1 Agenda Shifting from Accessibility to Quality	37
3.1.1 Assessing the Accessibility	41
3.2 Assessing the Quality of Education	45
3.2.1 Quality Education in SDG 4	47
CHAPTER IV. Results-based Evaluation in Development Cooperation	57
4.1 Development Evaluation and Quality Measurement	58
4.1.1 Program Evaluation	58
4.1.2 Classification of Evaluation Approaches and Models	60
4.1.3 Development Evaluation	64
4.2 Results-Based Approach in Development Evaluation	69

4.2.1 What is results-based evaluation?.....	7 1
4.2.2 Logical Framework, Logic Models and Results-based management.....	7 7
4.3 Challenges of Results-based approach.....	8 1
CHAPTER V. Capability-based Evaluation Framework for Quality Education	9 0
5.1 Evaluation Framework for Basic Education Quality.....	9 4
5.1.1 Enabling Inputs	9 6
5.1.2 Learning Outcomes	9 7
5.1.3 Teaching and Learning.....	1 0 0
5.2 Discussions.....	1 0 5
5.2.1 From Education Standpoint.....	1 0 6
5.2.2 From Evaluation Standpoint.....	1 2 4
5.3 Application and Methodology	1 3 8
5.3.1 Famous Case Studies of S. Alkire	1 3 8
5.3.2 Application methods	1 4 3
CHAPTER VI. CONCLUSION	1 4 8
6.1 Evaluation of Education Quality and Capability	1 4 8
6.2 Limitation of the Study	1 5 3
6.3 Final Remarks	1 5 5
BIBLIOGRAPHY.....	1 5 8
국문 초록:	1 7 0

CHAPTER I. INTRODUCTION

1.1. Research Background and Objectives

There can be little doubt that setting global education development goals have resulted in increased opportunities for many children to be educated throughout the world. The adoption of the Millennium Development Goals (MDGs) and the Education for All (EFA) goals in 2000 was certainly a turning point in international education and development. It established a new set of goals, targets, and indicators that “provided strategic direction to educational planning and budgeting; are important to monitor progress; and have encouraged focused and sustained support from development partners” as a global discourse of progress (UNESCO & UNICEF 2013, 7). As a result, the number of out-of-school children of primary school age worldwide decreased by nearly half, from 100 million in 2000 to an estimated 57 million in 2015 and gender parity in primary school has been achieved in the majority of countries (UN, 2015).

Development goals and targets are changing over time and educational goals also have been changing. In 2015, the Sustainable Development Framework with a new set of more ambitious goals and targets have replaced the MDGs and EFA Goals. There are differences between MDGs, EFA and SDGs in terms of scope, coverage, and policy focus. While MDG2 focused on children and ‘access’ to primary education, EFA had a broader agenda aimed at ‘meeting the basic learning needs of children, youth, and adults’ (UNESCO, 2016).¹ Unfortunately, the EFA commitment to ensuring the right to basic education for all was not met by the 2015 deadline, thus SDG4 pursues this unfinished agenda, committing all countries to ensuring equal access to ‘quality learning opportunities’ at all levels of education in a lifelong perspective (UNESCO, 2017). This represents a departure from the MDG framework's narrow focus on universal primary education, and it goes far beyond the EFA that accompanied the MDGs (Unterhalter, 2019).

The SDG education targets were developed in a more inclusive manner than any other global agendas and targets, including the MDG. From 2013 to 2015, it was successful in setting targets that emphasized inclusion, quality, and equality at all levels of education through the Open Working Group and consultative processes. However, while the targets and goals have grown in scope and become more quality-related, the monitoring and evaluation framework has remained unchanged. For example, the global indicators for SDG4 has resulted in metrics that neglect many of the targets' values, particularly in terms of quality and free education, and substantive, not just distributive, meanings of equality (Unterhalter, 2019). Among these values, quality is one of the most significant focus in SDG4 targets and it is expressed explicitly in three out of seven outcome targets (UNESCO, 2017).

Though the major target value in the education development goal has shifted from access to quality, their monitoring and evaluation framework, particularly the quantitative type of indicators and evaluation approaches, has not changed. King (2017) expresses concern about focusing on simplified indicators that the flagship indicator itself, such as enrollment rate for MDG2 targets, has power to influence the value weight over targets and allocation of resources by stating:

The attraction of a single flagship lead indicator is that it hopes to dramatically simplify this complexity and offer a global measurement for 'minimum proficiency'. This flagship could become more like an education MDG, narrowing the focus to a single concern, and being more relevant to the aid community than to UN Member States as a whole. It would also become a focus for donor financing. (King 2017, 814)

Each year, tremendous amount of education ODA is committed to achieve those targets. There are 29 DAC donor countries, more than 30 multilateral organizations, and 226 recipient countries based in OECD CRS, and the number of development programs and projects are uncountable. Considering a wide array of data availability and monitoring

systems and capacities among countries, it seems reasonable to monitor and evaluate target achievements through simple and monitorable indicators that are easy to collect in development context.

Data availability is one of the most critical factors which affects the evaluation design. Depends on types and availability of data set which evaluators can attain, the framework of evaluation is determined, for example, whether to conduct experimental or not. What I found in research and practices in development evaluation is far too simple in terms of its criteria, approaches, and models . OECD DAC evaluation criteria has been dominantly accepted among evaluators with limited theoretical analysis Conlin & Stirrat, 2008). And the Results-based Management and evaluation (hereinafter, RBM) has been prevailing evaluation approach in development field since 1970s (OECD, 2019). Perhaps the difficulty of quality data collection in developing countries can be one reasonable reason for keeping development evaluation simple and monotonous in terms of approaches, criteria, and methods.

Nevertheless, there are critical limitations that current evaluation approach in development cooperation poses. OECD (2019) criticized that there is a clear tendency to focus on what can be measured easily in development evaluation. Thus, critical outcome and change that needs more complex arrangements of data analysis are easily neglected or often omitted in evaluation framework. As Conlin & Stirrat (2008) argue, the development industry has changed rapidly and demands for more sophisticated and interpretive evaluation approaches are increasing. Responding to those critiques, in recent decade, experimental evaluations are widely conducted in diverse sectors as increased demands and pressure on verifying the causality chain and aid-effectiveness in numeric form. Impact evaluations seemed successful in particular areas where economic analysis is already adopted as main evaluation approach. However, in many social sectors where experimental research settings are not easily obtainable, particularly in education, research methods to evaluate the causality of development interventions are not actively used in comparison to other sectors (Bamberger et al., 2016).

In education sector, international organizations such as UNESCO, UNICEF, Save the

Children provided their own framework for education quality. Also in academia, there have been many approaches, frameworks and methodology are developed and adopted (Allan, 1996; Carlson, 2009; Garira, 2020; Griffith, 2008; Rachel, 2017; Scheerens, 2000; Williams, 2001). Despite continuous efforts and resources devoted to achieving the target of global education development goal, to ensure equal access to quality learning opportunities for all, it remains unclear whether these interventions are effective and how they affect the learners. Though it might not be easy to unlock the 'black box' in education through classic experimental or value for money approaches, we still need ways to present the effectiveness of interventions and the changes that learners show through the education process.

In this study, one of alternative ways will be discussed and proposed to tackle the issues we face in education development cooperation. I think there are two meaningful contributions which this study can make in educational development cooperation. First, this study links discourse and practices of two disciplines, education, and development evaluation, in relation to education quality evaluation. In education, enormous number of articles and research are done within wide spectrum of topics from philosophical discussion on definition of quality in education to empirical research on what determines and how quality education looks like in different context. However, there are little studies done which links developed knowledges in two different disciplines. In development studies, most evaluation research are done by evaluators who are trained to make value judgement, collect, and analyze data, develop, and operationalize evaluation design in practice. Therefore, they tend to be empirical and focusing on methodological debate and analysis of the findings based on given evaluation framework (Bamberger, 2000). Before 2000s, when the 'project-based' modality was dominant in development cooperation, evaluations were donor-centered and implementation oriented (Conlin & Stirrat, 2008); however, expected responsibilities for evaluators are now expanded to prove effectiveness and long-term impact of development interventions through scientifically and theoretically recognized methods. This study is expected to be one of attempts to make synergies by connecting literatures of two independent area of studies and be useful

for both sides to further develop more concrete knowledge on quality education evaluation in development context.

Second, numerous studies and evaluation research which applied the capability approach in development evaluation are conducted in education sector; however, they are mostly based on small-scale project evaluation and deal with very specific context. Hart (2012), Hinchcliffe & Terzi (2009), Saito (2003), Terzi (2007), Unterhalter (2003), Walker (2005) suggested how the capability approach can be applied to social justice in education. Others provided more practical application of the capability approach in education evaluation (Camfield & Tafere, 2011; Comim et al., 2008; Kuklys, 2005; Young, 2009). However, literatures related to adopting the capability approach as epistemological framework to develop an alternative evaluation framework in sector-wide seem very limited and scarce. Few studies are found in health sector yet not in education. Therefore, this study will be first of its kind in educational research. It is unlikely to be as scientifically rigorous as other empirical; however, it will contribute to us taking a step further to get an answer on how to measure 'quality' in education.

1.1.1 Research Questions

The aim of this study is to propose an alternative evaluation framework for basic education quality in development context. On the journey to the end, to make my argument more scientific and convincing, policies, discourses and theories that are frequently used to explain and assess the quality of education in both fields of education and international development, are critically reviewed through the lens of the capability approach by Amartya Sen.

	Research Questions	Research Objectives
1	What are the critical issues that affect the evaluation of quality education in development cooperation and what limitations remain to be overcome?	To identify the issues affecting educational evaluation of quality in basic education;
1.1	How has the quality of education been evaluation in education?	

1.2	How has the quality of education been evaluated in development cooperation?	
2	What are transformative alternatives to manage identified issues affecting the evaluation of quality in basic education?	To identify the measures for educational evaluation of basic education and; To propose an alternative evaluation framework for assessing quality of basic education in development cooperation.

First, in order to identify issues and limitations to be overcome in terms of education quality evaluation in development cooperation, prevailing evaluation approach in development evaluation as well as several quality education frameworks in education are critically examined in terms of quality evaluation. And after that this study proposes an evaluation framework by applying core concepts of the capability approach suggested by Amartya Sen. It examines the dominant evaluation approach in international development cooperation, known as Results-based Management, from the perspectives of education and evaluation, and offers both scholarly evidence and real-world examples to demonstrate why the new proposed framework, as well as the capability approach, is a better way to deal with quality issues in education development.

Some see the capability approach as an alternative 'paradigm' that can fill all the gaps left by utilitarian welfare economics that is solely focused on economic growth. And, while I agree that the capability approach provides compelling insight, as Alkire argued, in order to be called a "paradigm," it must visit all of the possible workshops of discipline, including development, poverty, evaluation, or whatever it takes, and figure out how the insight leads to better work, what changes, and what remains the same.

According to Thomas Kuhn (1970), scientific revolutions occur when a theory emerges that appears "better than its competitors" and has the "promise of success" – that is, it appears more likely to assemble facts in a meaningful way than the predecessor. This study is one of those attempts to see how the capability approach can be operationalized differently in the study of development evaluation and provide a better explanation for quality evaluation in basic education sector.

1.2 Capability Approach as a Theoretical Methodology

As critical theoretical research, this study sees the capability approach as a theoretical methodology. Though many scholars including Butin (2010) recognize the importance and contribution of theoretical studies in education, educational research often reduces theory to a framework for analyzing empirically collected data and not considered methods (Matias, 2021). Attention to theoretical research in education which applies theory as research method is increasing and its academic demand is officially recognized by the American Educational Research Association (AERA) (Allen, 2021). The most recent work of Matias (2021), a University of Kentucky professor and Director of Secondary Education in the Department of Curriculum and Instruction, provides a great guide explicating how theories can be applied as methods in educational research.

Capability approach has been employed in various fields of study to the development of conceptual and normative theories within political philosophy, development ethics, environmental ethics, philosophy of education and so on. Most associated capability scholars such as Martha Nussbaum, Ingrid Robeyns, Severine Deneulin, David Crocker, Elaine Unterhalter, Melanie Walker sophisticated theoretical foundation of the capability approach in each devoted field of study. There is another group of capability studies which is more inclined to empirical research that shows how the capability approach can be applied in practice such as a critique on existing social practices or an evaluation exercise. Most of them require empirical research techniques to measure functionings from the data, analyze the determinants of the functionings achievements and the choice of the relative weights among functionings.

This study applies the capability approach as an evaluative framework to critique existing social practices of how education quality has been evaluated in international development. Based on the critical analysis, an alternative evaluation framework for education quality is proposed which also reflects the core concepts of what capability approach offer to the field of education, namely a normative commitment to conceptualize distinction between capabilities and functionings.

The integrative review, also known as the critical review approach, of literatures on approaches to development evaluation and education quality assessment is conducted for the theoretical examination part. This method is similar to the semi-structured review approach (Snyder, 2019), but it usually serves a different purpose: assessing, criticizing, and synthesizing the literature on a research topic in a way that allows new theoretical frameworks and perspectives to emerge (Torraco, 2005).

1.2.1 Contribution of the Capability Approach as a Research Method

As this study not only aim to provide philosophical elaboration on how the capability approach contribute to evaluate the quality of education in development evaluation but also to propose an alternative evaluation framework to be used in practice, the application of the capability approach does is underspecified and open-ended. Following are three important contributions what the capability approach proposes in this study:

Epistemological reason

Robeyn (2000) emphasizes that the capability approach is primarily and mainly a framework of thought, a mode of thinking. Evaluation practices has been considered more of a professional work using specific methods of studies. And because the work itself is conducted on the ground and practically utilized, many think that evaluation is not a theory-based discipline. Lewin (1951) said there is nothing so practical as good theory, on the other hand Fullan (2001) said there is nothing so theoretical as good practice. However, as Shadish (1998) claimed “evaluation theory is who we are”, what we say about what we do (our theory) is just as important as what we do (our practice) and who we are (our profession) as evaluators. Carden and Alkin (2012:103) said “perhaps the complexity of program contexts and evaluator actions are so vast that we will never be able to create a descriptive evaluation theory.”

Likewise, many of evaluators and theorists in the field of evaluation argue differently about the evaluation theories and what comes first: the theories or the practices. However,

more important fact is that, regardless of having refined theories or not, evaluators cannot avoid from containing their own epistemological reflects in their practices. Therefore, the capability approach has its meaning to provide alternative framework of thought and a mode of thinking in the evaluation studies. Because it affects to not only valuing, but also methods and use of the evaluation, epistemological reason of the capability approach to be considered is very crucial.

Methodological reason

As Alkire (2002) empirically proved in her work, applying the capability approach in the process of collecting and analyzing data greatly affects the results of the evaluation. After she compared three development projects evaluation conducted through both a standard economic cost-benefit analysis and capability analysis, she concludes that only one project has the similar result but two have not. She used participatory method to identify relevant capabilities, and she claimed that a methodology is needed by both local and international institutions with heterogeneous considerations – such as how participatory an activity was, how much it benefited the poor, how much it empowered women, built capacity, strengthened institutions, improved the environment, catalyzed local government, and mobilized communities to take collective action, and so on (Alkire, 2002).

Hatakka and Lagsten(2012) argue that by adopting Sen’s capability approach when evaluating the outcome of students use of internet resources, they could understand why and how development outcomes are achieved. To understand it in more Sen’s explanation, Sen(1997) differentiates comprehension outcomes from culmination outcomes. Culmination outcomes, which most of evaluation and research focus on, are a list of capabilities, whereas comprehension outcomes include the identity of the chooser (agency) and the choice. “Most of the studies are based on culmination outcomes, but if we want to see the process – why, or why not, individuals choose to do or be and how – we should rather focus on comprehension outcomes. (Hatakka and Lagsten 2012:36)” Those two different concepts of outcomes will not be incorporated into the evaluation

framework in this study; however, operationalizing capability approach in evaluating education quality will certainly influence evaluation design, methods, and data collection and analysis.

Evaluative reason

As explained above, because of its pluralism, the capability approach can capture individual diversity as well as external conversion factors which affect people's well-being. "It is problematic because it may be hard to generalize the findings due to the individualistic nature of Sen's capability Approach. It is an advantage because we do not limit the findings to organizational or cultural common perceptions. Values of each individual are taken into consideration in the analysis. (Hatakka and Lagsten 2012)"

As UNDP's Human Development Report enlarged the dimensions of well-being from economic outcome to education and health, the capability approach enables evaluators to delimit themselves from opening the possible list of valued capabilities of different opinions of people. In Alkire's work (2002), there is a story of a woman named Dadi Taja, a widow, who are the beneficiaries of the rose cultivation project. After one and a half year, Dadi Taja said she is able to "walk about without shame" and, "people in the village now respect me". She explained that having income from rose cultivation is valuable; however, the benefit is not only constraint to it. She said it is very much delightful to have the fragrance of roses with her and she is satisfied from working in a group, and proud of her job because the rose garland are used sacredly in saint's shrines.

If evaluative practices only focus on the income generation effects, those benefits including empowerment, knowledge, and meaningful work which are highly valued by participants will be excluded (Alkire, 2002). Although it is difficult and sometimes impossible to measure those values in economic terms, there are good reasons to capture them in evaluating well-being and already many have applied the capability approach in their practices.

1.2.2 Examined Body of Literature

This study examines three groups of literatures. First, in Chapter II, capability literatures are reviewed with specific focus on the education sector. As previously stated, the capability approach is applied in a wide variety of studies with different purposes, including theoretical discourse and empirical application. Literatures referenced in Chapter II are the work of major capability scholars to present the overall landscape of the capability studies and narrow down to mostly empirical application studies in education as well as development evaluation.

Second, to examine how quality is defined and evaluated in education development cooperation through the lens of the capability, scholarly works of policy analysis related to the global education agendas such as MDG, EFA and SDG as well as reports and documentations produced by international organizations are reviewed in Chapter III.

Lastly, to link the theoretical analysis of this study to the alternative evaluation application, the most prevailing evaluation approach in current development context so-called results-based evaluation is critically examined in terms of its theoretical foundation, structure, and application methods.

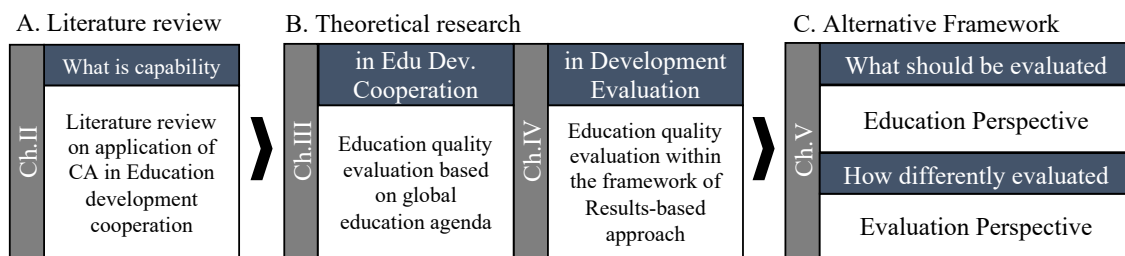
1.3 Research Scope and Structure of Dissertation

Though this topic offers numerous research opportunities, it is critical and proper to define the research scope in order for it to be feasible. First of all, the subject of evaluation is narrowed to outcomes of development intervention in basic education sector. According to OECD CRS data, it is the most invested sub sectors in education sector in ODA and the most significant education level for learners for development and expand their learning path. Second, this study mainly examines literatures and practices in relation to development evaluation because this study is focusing on the assessment of the quality of basic education in development context, not education in general. Third, the majority of the literatures examined in this study were published after 1990, which coincides with the

adoption of major global education agendas and the introduction of the capability approach in the foreign aid community.

This dissertation focuses on comparing two conceptual framework of development evaluation: results-based approach and capability approach. As opposed to a dominantly used approach, the results-based, Amartya Sen’s capability approach has been partially reviewed as an evaluation framework by various field of studies. There are some research and evaluation cases that adopted the approach; however, the number is still too small, and the array of sectors, levels and scope of practices are quite wide to make a meaningful integrative analysis. Therefore, this dissertation will be one of the pioneer works that discloses the special linkage between the meaning of capability and education in development studies with specific focus on evaluation.

It consists of six chapters and divided into three parts of discussion, first part is the review of capability literatures, and second part is theoretical examination on how education quality is evaluated in development evaluation. Lastly, an alternative evaluation framework is proposed.



In Chapter II Amartya Sen’s general viewpoint on education is briefly discussed followed by work of scholars who applied Sen’s capability approach in educational research. Due to its open-ended and underspecified nature, capability-based discourses are developed in numerous fields of study and it is necessary to focus on literatures that are related to education. In Chapter III and IV, quality related global education agenda and the representative evaluation framework are critically reviewed through the lens of capability approach respectively. It aims to show how education quality is conceived in policy level and how it is connected to the real-world evaluation. Based on critiques and

limitations discussed in Chapter III and IV, an alternative framework for quality evaluation in education sector is proposed in Chapter V. It tries to apply the core conceptual value of the capability approach which discussed in Chapter II.

CHAPTER II. Literature Review

2.1 Core Concepts of the Capability Approach

The 'Capability Approach' is introduced by an economist-philosopher, the Nobel laureate Amartya Sen, in 1990s. As Sen indicates (Sen, 1993; Sen, 1999), the concept of capability is influenced by works of Aristotle, Adam Smith, John Rawls, and Karl Marx. As he explicitly states in his famous book 'Development as Freedom', the capability approach defines development as a process of expanding human freedoms:

Development can be seen, it is argued here, as a process of expanding the real freedoms that people enjoy... Growth of GNP or of individual incomes can, of course, be very important as means to expanding the freedoms enjoyed by the members of the society. But freedoms depend also on other determinants, such as social and economic arrangements as well as political and civic rights. (Sen, 1999:3)

One of the most prominent contribution of the capability approach is to the development of the notion of human development (Fukuda-Parr, 2003, 2011; Fukuda-Parr & Cid-Martinez, 2009; Ul Haq, 2003). It had such a political impact that the United Nations Development Programme (UNDP) has published the Human Development Report every year since 1990, which is based in part on the capability approach. It has emerged as the leading alternative to traditional economic frameworks for considering poverty, inequality, and human development in general (Clark, 2005). Since the Tanner Lecture 'Equality of What?' delivered at Stanford University in 1979, framework that is directly concerned with human capability and freedom are developed throughout many of his articles and book.

It has been applied in various disciplines including public health, environmental ethics, and educational justice. For health sector, Prah Ruger (2004) and Venkatapuram (2013) have discussed comprehensive theories of health justice and Venkatapuram's research

influenced works of the World Health Organization. Nielsen (2015) makes the case for why capability theorists should be concerned about health, arguing that it plays a critical role in human well-being and agency. The capability approach also applied to healthcare ethics that notably Entwistle & Watt (2013a, 2013b) and Millar (2013) argued the importance of person-centered care and capability applied evaluation of healthcare justice.

It also applied to climate justice in various ways. Gutwald et al. (2014), Lessmann and Rauschmayer (2014) examined the concept of sustainable development from a perspective of capability. Holland (2008a, 2008b, 2014) argues that capabilities that people have affects the environmental justice. Therefore, environmental justice as well as climate adaptation planning and process should concern development of people's capabilities (Kronlid, 2014; Schlosberg, 2012; Schlosberg & Carruthers, 2010).

In addition, in education sector the capability approach has been influential in expanding the notion of right to education as well as education justice. Unterhalter (2003, 2009a, 2009b, 2014, 2017) and Walker (2005, 2008) are the key capability scholars in education together with Hart (2012b, 2012a), Hinchcliffe & Terzi (2009), Saito (2003) and Terzi (2007).

Lastly, there are numerous studies and evaluations applied the capability approach for the assessment of human development and well-being. Robeyns (2006) grouped capability applications into nine different types and three of them including general assessment of the human development of countries; assessing small-scale development projects; and identifying the poor in development countries are closely related the capability application in development context.

In terms of methodology of capability measurement, Robeyns (2006) used a wide variety of empirical methods. Descriptive statistics of single indicators, scaling, factor analysis, and structural equation modelling are found from the literatures of Brandolini & D'Alessio (1998), Kuklys (2005), Qizilbash & Clark (2005). Alkire (2005), Hatakka & Lagsten (2012), and Wolff & De-Shalit (2013) used qualitative empirical techniques.

2.1.1 Functionings and Capabilities

According to Sen, when development means a process of expanding human freedom, a person's capability means the alternative combinations of functionings that are feasible to achieve, i.e. a person's substantive freedom to achieve those alternative combinations of functionings (Sen, 1999). Understanding difference between the concept of functionings and capabilities is one of the most important steps to understand the conceptual foundations of the Capability Approach as an evaluation framework. They are found in Sen's criticisms of traditional welfare economics, which typically associate well-being with opulence such as income and commodities or utilities such as happiness and desire fulfillment. According to Sen, capability and functionings are distinguished that "A functioning is an achievement, whereas a capability is the ability to achieve. Functionings are, in a sense, more directly related to living conditions. Capabilities, in contrast, are notions of freedom, in the positive sense: what real opportunities you have regarding the life you may lead" (Sen & Hawthorn, 1988, p. 36)

According to the Stanford Encyclopedia, functionings are 'beings and doings', that is, various states of human beings and activities that a person can undertake. Examples of the 'beings' are being well-nourished, being undernourished, being educated, being illiterate, being part of a supportive social network and so on. Examples of the 'doings' are travelling, caring for a child, voting in an election, taking part in a debate, taking drugs, killing animals, donating money to charity and so on. These examples indicate that many features of a person could be described either as a being or a doing. The notion of functioning is a conceptual category that is in itself morally neutral. Functioning can be univocally good (e.g., being in good health) or univocally bad (e.g., being raped). Sometimes badness and goodness of certain functionings may depend on the context. On the other hand, capabilities are the real freedom and opportunities that a person has to achieve functionings. Therefore, if being able to read is a functioning, the real opportunity to learn can be the corresponding capability. The distinction between functionings and capabilities is between achievements, on the one hand, and freedoms or opportunities that

on can choose, on the other (Yoo et al., 2019). Gore (1997) notes that while ‘functionings’ refers to achievement, ‘capabilities’ refers to the opportunity set.

Not only the capability approach provides new definition of development, but it also suggests fundamental perspective, way of considering what development means and how to be measured, of evaluating development. As Robeyns (2005) indicates, the capability approach is used and has been used across multiple academic disciplines with different epistemological goals and methodologies. Because the purpose and the methods of applying the capability approach are different, its interpretation also has been very diverse.

When capability is defined as freedom, it is also important to understand what kind of freedom it refers to. As Kaufman (2006) argued, capability does not only contain opportunity concept, but it should be conceived as ‘real freedom’ as Sen stressed (A. Sen, 1985, 2003). What does ‘real freedom’ mean in capability approach? It means that one has every means to achieve doings and beings that one has reason to value. In other words, it indicates not only the formal freedom to do or be something, but also the substantial freedom and opportunity to achieve it. To take an example in education opportunity, a child is living in a country which ensures free compulsory basic education as a national system. The child has right to be educated and no one will stop her from being so; however, the child might lack substantial opportunity to be educated if there is no school in walking distance from home and no means of transportation is provided. In other occasion, a child may live near school but still may not have real freedom to be educated because her community and parents do not permit females to go to school.

In traditional development evaluation, we only measure the ‘state of being’ instead of ‘real freedom to achieve those state of beings.’ We evaluate whether or not a child is enrolled at school without asking whether the child had real opportunity to go to school. In case of being literate, we only have information how many children are able to read and write in certain grade without asking why some are not able to complete the primary education and why not able to reach minimum proficiency level of literacy after several years of formal schooling. Therefore, applying the capability approach as an alternative

framework for evaluating outcomes of basic education in development context expands the boundary or information space what to be evaluated.

2.1.2 Pluralism and Conversion Factors

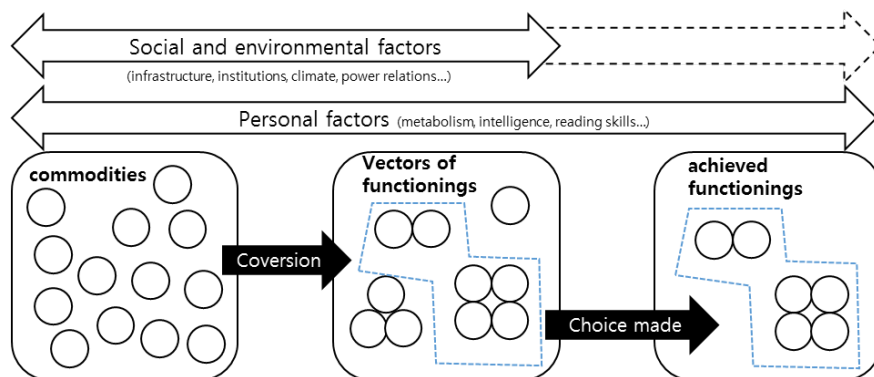
One of the main strengths of the capability approach is its pluralism that it could account for interpersonal variations (Robeyns, 2000). This is not a side-effect or by-product of the capability approach, but is of central importance to Sen: “Human diversity is no secondary complication (to be ignored, or to be introduced ‘later on’); it is a fundamental aspect of our interest in equality” (Sen, 1992:xi). The capability approach accounts for diversity in two ways: by its focus on functionings and capabilities as the evaluative space, and by the explicit role it assigns to individual and social conversion factors of commodities into functionings (Robeyns, 2000).

Because income reveals different extent of people’s well-being, for example, for a person who is healthy physically and psychologically and employed; but for an unemployed person, or a person who is suffering from emotional or psychological stress, Sen (1992, p. 101) said “these standard measures are all basically parasitic on the traditional concentration on the income space and ultimately ignoring the fundamental fact of human diversity and the foundational important of human freedom.”

In capability approach, whether a person has the capability, i.e. the real opportunity to achieve a certain functionings depends on so-called ‘conversion factors’ (A. Sen, 1992, 1993). It refers to the degree in which a person can transform a resource into a functioning. Recognizing pluralism and diversity of each individual, the capability approach thus stresses that “different people need different types and different amount of capability inputs to reach the same wellbeing. In the terminology of the capability approach, this is highlighted by pointing out that there are factors which influence how well a person can convert capability inputs into capabilities” (Robeyns, 2005, 6). When a person converts the commodities into functionings, the conversion factors differ the results. In conversion factors, there are individual, social and environmental differences.

There are several different types of conversion factors, and the ones that are discussed are usually divided into three categories. All conversion factors influence how a person can be or is free to convert resource characteristics into functioning, but the sources of these factors vary. Otto & Ziegler (2006) differentiated conversion factors into three: personal, socio-structural and cultural, and institutional. They are slightly different from what Robeyns perceived that personal conversion factors are physical condition, literacy, competences, etc while socio-structural and cultural conversion factors are social or religious norms, gender roles, power relations and any forms of discrimination. And institutional conversion factors include welfare and educational arrangements, and collective provisions.

In this study, I divided conversion factors into three: personal, social, and environmental and the list of information space for evaluation of each dimension will be provided in Chapter V. In relation to education, personal conversion factors include physical and emotional health, sex and functional literacy. If a child is disabled or malnourished, attending school every day and concentrating on learning will be of limited functionings to achieve. For social factors there are education policies and laws, social and religious norms, discriminations or hierarchies to gender, class or race that directly and indirectly affect children to choose to be educated. Lastly, environmental conversion factors are related to physical or built environment where children require for their education for example, safe and accessible school and learning space in both community and home.



Source: (Yoo et al., 2019)

Sometimes not only the personal conversion factors matter but also social characteristics such as social norms, power relations, culture and tradition can also affect the conversion of the commodities to functionings. The three types of conversion factors all emphasize that knowing what resources or commodities a child owns or can use is insufficient to assess the level of well-being he or she has achieved or possibly achieve in the future; instead, we need much more information about the child and the circumstances surrounding. What we can learn from capability approaches' conversion concept is that we need to also evaluate both internal and external conversion factors which make children's real freedom to be feasible or constrained by.

Other scholars depending on the purpose and scope of capability application provide diverse lists on the conversion factors. However, the concept of conversion factor has important implications and makes a significant contribution to education and development evaluation because it recognizes the diversity of each child and the various conversion factors that influence the choice made.

2.1.3 Valuing as Ends and Means

For the capability approach, people's real freedom is the ultimate ends to achieve and be evaluated. One of potential implication of the capability approach to evaluation research is that it stresses the importance of distinction between means and ends when evaluate something as development goals. When one says that the ultimate ends of what capability approach value are people's capabilities, it implies that the evaluation of policies and development interventions should be conducted focusing on their impact on people's substantial opportunities and their actual functionings.

Therefore, when setting targets, goals or valuing something, the capability framework asks us to distinct whether they are ends in itself, or as means to achieve ends. For example, it asks us whether children are able to read, and whether the means or required conditions for this capability, such as safe and accessible learning space, qualified teachers, and enough learning materials, are present. Sometimes it is difficult to

distinguish the end and the mean because being able to read can be the end in itself, but someone can argue that it can also be means to read the Bible, street sign, or vote.

Of course, the capability approach not only focus on normative ends to evaluate but also means such as financial resources and resources. When Jean Drèze & Sen (2002) evaluated development of India, an analysis of resources such as income and food took a critical part of their work. Therefore, means of development such as availability of resources, legal entitlements, and other social institutions, are also taken critically in capability evaluation; however, the point that capability approach makes is that they are not the ends of development, only their means.

If it causes controversies to distinct ends and means and if it is more complicated and more informationally demanding to focus both on functionings and capabilities, why the capability approach still valuable? There are several possible answers. The main reason is provided by Sen (1992) that focusing on the ends rather than the means is better because people are different in their ability to convert means into valuable functionings or capabilities. In other words, when the means to achieve development become the subject of evaluation, it neglects the important part of how those means actually contribute to the realization of the ends. Another important reason is that by evaluating the ends, we can avoid the false premise that there is only one overarchingly important means to that end such as income, but instead ask explicitly what means are taken into account to foster a particular capability.

2.2 Critiques and the Capability List

2.2.1 Capability List

The capability approach is deliberately incomplete (Alkire, 2002). In *Inequality Re-examined* (Sen, 1992) Sen identifies two grounds for allowing incompleteness: fundamental and pragmatic by saying that: “the ideas of well-being and inequality may have enough ambiguity and fuzziness to make it a mistake to look for a complete ordering

of either... The ‘pragmatic reason for incompleteness’ is to use whatever parts of the ranking we manage to sort out unambiguously, rather than maintaining complete silence until everything has been sorted out and the world shines in dazzling clarity... ‘Wait for toto’ may not be a cunning strategy in a practical exercise (Sen, 1992: 49, 1999: 153-4).”

This incompleteness seems evasive and has been attacked by other theorists; however, it is in fact one of the most important advantages of the capability approach (Alkire, 2002). Because of this reason, we can always welcome other alternatives and not taking this framework as the only way to interpret the capability approach. Sen’s capability approach received criticisms regarding its incompleteness as a theory. Nussbaum (2003) proposed a list of ten “central human capabilities” and claimed that Amartya Sen has to endorse one specific and well-defined list of capabilities. In response to those critiques, Sen refused to make a specific list because he argued that first, it is not a theory to make the list of capabilities and second, it should be made among people through democratic process. Because of this incompleteness, arguments and discussions, misunderstandings and clarification process has been on and off in academia. Following table shows different lists of capabilities from various studies.

Authors	Nussbaum (1995, 2000, 2003)	Alkire and Black (1997)	Narayan and Petesch (2002)	Robeyns (2003)
Aim/scope of the list	Universal	Universal	Universal	Gender inequality in Western societies
Level of abstraction	High	High	High	Low
Dimensions	<ol style="list-style-type: none"> 1. Life 2. Bodily health 3. Bodily integrity 4. Senses, imagination, and thought 5. Emotions 6. Practical reason 7. Affiliation 8. Other species 9. Play 10. Control over one's environment 	<ol style="list-style-type: none"> 1. Life 2. Knowledge and appreciation of beauty 3. Work and play 4. Friendship 5. Self-integration 6. Coherent self-determination 7. Transcendence 8. Other species 	<ol style="list-style-type: none"> 1. Material assets 2. Bodily health 3. Bodily integrity 4. Emotional integrity 5. Respect and dignity 6. Cultural identity 7. Imagination, information, and education 8. Organizational capacity 9. Political representation and accountability 	<ol style="list-style-type: none"> 1. Life and physical health 2. Mental well-being 3. Bodily integrity and safety 4. Social relations 5. Political empowerment 6. Education and Knowledge 7. Domestic work and nonmarket care 8. Paid work and other projects 9. Shelter and environment 10. Mobility 11. Leisure activities 12. Time-autonomy 13. Respect 14. Religion

Source: author

2.2.2 Critiques on Capability Approach

There are two major critiques on Sen's refusal to endorse a list of capabilities. Robeyns (2005) has labeled them the strong and the weak critique. The strong one is notably raised by Nussbaum, who has proposed a list of ten "central human capabilities", by arguing that if Amartya Sen wants his version of the capability approach to have any bite for addressing issues of social justice, he has to endorse one specific and well-defined list of capabilities. The weaker critique admits that Sen should not develop a specific list of capabilities; however, there needs to be some methodological reasoning on how such a selection could be done.

Sen does not accept the strong critique. There are reasons. Firstly, Sen does not intend to develop the capability approach as a well-defined theory. In many of his work, it is easy to find that the capability approach is developed as a general framework to the evaluation of individual development and social arrangements. Therefore, the capability approach as such is deliberately too underspecified to endorse one single list of capabilities (A. Sen, 1993, 2003). Instead, he argues that we must leave it to democratic processes and social choice procedures to define.

Secondly, Sen stresses the importance of the role of agency by choosing their own relevant capabilities with the freedom to reason and its process of choice. Because the processes matter, a pure theory could not provide a specific list of capabilities and should not be universalized. It could be explained better by borrowing Sen's distinction between *culmination outcomes* and *comprehensive outcomes*. The *culmination outcomes* present the functionings (doing and being) which one achieved as an end. This is what most of evaluations have focused because it is easy to measure. In contrast, the *comprehensive outcomes* include

aspects of the choice process, including the identity of the chooser as well as information on the legitimacy, fairness, or democratic content of the process that led to the outcome (Sen, 1997). Therefore, when Sen's capability approach is applied, it automatically implies that theories of deliberative democracy and public participation are also engaged. For Sen, a list of capabilities must be context dependent. Here is what Sen responds by arguing:

The problem is not with listing important capabilities, but with insisting on one predetermined canonical list of capabilities, chosen by theorists without any general social discussion or public reasoning. To have such a fixed list, emanating entirely from pure theory, is to deny the possibility of fruitful public participation on what should be included and why. (2005:158)

Many studies are conducted to fill this gap between theory and practice, especially the methodological attempts from economic formula to evaluation methods. Although Sen identified the capability approach as an evaluation framework, it has not been actively reviewed in the field of evaluation. There can be two possible reasons. First, because of its incompleteness as a theory, studies found clear limitation to apply it to practice. Second, the evaluation discipline tends to stress more of practical research rather than developing a theory.

2.3 Capability and Education

Capabilities have been widely applied in many sectors including development studies since Sen's remarkable contribution on Human Development Index which shifts the perspective of development from income based to the expansion of

substantial freedom of people. It also has been applied to education by various researchers and theorists including Saito (2003), Terzi (2008), Unterhalter (2007), Vaughan (2007), Walker and Unterhalter (2007) and their work have provided distinctive perspectives on social justice in education, particularly in relation to gender.

In this chapter, after reviewing related literatures, I would like to provide three reasons why capability approach matters in education. First, as a theoretical framework, the Capability Approach expands the notion of ‘right to education’. Second, respecting and considering diversities and heterogeneity of individual learner is the core aspect in terms of evaluative space both in education and the Capability Approach. Third, education plays both intrinsic and instrumental role in the Capability Approach. Being educated itself could be the capability that one should pursue and at the same time it could be valuable to expand one’s other capabilities.

2.3.1 Expanding the Right to Education

Robeyns (2006) pointed out, the human rights approach to education has been criticized for being reduced to legal rights only, which are operationalized at a high level international and state-led level, while moral rights are ignored. Tikly & Barrett (2011) argue that this understanding of a rights-based approach has predominated in the context of the MDGs and Dakar Framework, as well as through the responses of rights-driven multilateral agencies like UNESCO. However, rights-based quality frameworks are frequently concerned with moral rights, that is, with an understanding of rights that extends beyond the boundaries

of international and national agreements, laws, and policies to consider the underlying moral and ethnical dimensions of education.

There could be another chapter needed to discuss about the right to education in terms of its philosophical meaning and historical review. In this chapter, as Sen (2004) made useful distinction *between recognition, agitation and legislation* to promoting rights, upholding of right to education with legislation practices are introduced. Since emergence of modern nation states, compulsory state-funded education had become part of national law in many countries by the nineteenth century. Nowadays with numerous international declaration and agreement, free compulsory education as a basic human right began to appear at the international level of discussion.

The first explicit statement of the international right to education is appeared in the UDHR of 1948 which implied by the Declaration of the Rights of the Child of 1924. At the Article 26 of UDHR (1948), there are three elements emphasized. The first concerns the levels of education, with a stress for elementary education which is understood to be basic literacy, numeracy, and other essential skills. Second, it stipulates elementary education to be compulsory. Third, there is somewhat puzzling statement about parental prerogative to choose forms of education for their children.

As a 'declaration', the UDHR (1948) did not have binding power over nation states. However, it paved the road to concrete legal binding documents followed by the International Covenant on Economic, Social and Cultural Rights (ICESCR,1966) and the Convention on the Rights of the Child (CRC,1989). The ICESCR follows the UDHR to a large extent; however, it provided much more details and specific demands on states. Given its legally binding nature, it is

considered by many theorists to be the most authoritative statements of the right to education (Beiter 2006; Hodgson 1998). The covenant at the end were ratified by 160 states-parties.

Speaking of wide ratification, the CRC is the most ratified among all the UN rights treaties. The CRC have consisted of three 'P's: protection of children from harm; to provide services and opportunities, and to allow children participation in decisions that affect them. One of the most important factor in the CRC is the mention of 'equal opportunity' in terms of stratification of schools. Additionally, it drew attention to attendance and dropout from the enrollment. In following year in 1990, the Education for All initiative (EFA) launched with close links to the CRC.

Through those international movement of considering basic education as a fundamental right for all human beings, many nation states as duty-bearers made a legal guarantee for all children to be educated for free after adaptation of EFA. However, the focus of EFA on legal right to education was primarily about institutional education i.e., formal schooling. When the conceptualization of the right to education is limited to access to school, it contains substantial limitations in the opportunities for many children. As Robeyns (2006:70) describes:

Policymakers... being contented when they have strictly followed the rules that a limited interpretation of the rights imposes on them, even when additional efforts are necessary to meet the goal that underlies the right.

For example, in Brazil, free compulsory education from the ages of 4 to 17 is made in constitutional right to all people. Thanks to this strong legal guarantee,

Brazil made near universal coverage in terms of primary education and in a privileged position to EFA. Nevertheless, out of school ratio of primary school ages has been increased steadily since 2005 of 0.5% to 3.9% in 2011 (UNESCO), many children drop out before completion. In relation to the general quality issues, with the latest PISA results in 2018 estimating that only 32% of students in Brazil attained Level 2 or higher in mathematics while OECD average is 76%. This is not only the case of Brazil, but many children in third-world countries are facing quality issues in their basic education.

Even though it has been the strongest and most effective way of providing minimum education to children in the world, there are still clear limitations of the rights-based approach to education in practice to overcome. Robeyns (2006) analyzed four critiques including: for being overly rhetorical; for overemphasizing the legal aspect; for inducing policy makers to be contented with a limited interpretation of the right; and for being too government focused. Nussbaum (2000) also agree with what Robeyns's reservations about rights framework as deficient in their emphasis on formal education rather than effective entitlement. While these limitations of actualization in practice are justified, they should not be considered as faults inherent in the concept of 'human right'.

Why is rights framework not enough in practice? As Robeyns (2006) states, 'rights always need a prior moral criterion' at the theoretical level. In other words, rights are always rights *to something*. Rights are not ends in themselves but rather it may be used rhetorically instrumentally to achieve particular results. And this is where the capability approach could contribute to overcome such limitations of rights framework. Because capabilities, unlike rights, always give focus on people's substantial freedom, expanding capabilities become the real end of all

development endeavors (Robeyns, 2006).

How do rights framework focusing on formal education perceive a child who is in situation with no adequate education? Unterhalter demonstrates one of its aspects in her research:

At a school in Durban, South Africa, in 2005, children described hunger and social isolation as aspects of poverty. One girl graphically recounted the months in which she had no money for soap or water, could not wash herself on her clothes, and was unable to come to school because of shame. (Unterhalter 2007: 64)

As what Unterhalter states, there are multiple obstacles to uphold of the right to education. Regardless of the number of schools available in town, real opportunity for children to access and get educated could be hampered in real world. This gives good reason why capability approach is important to be considered in education in development as well as rights framework. When Nussbaum extended Sen's idea of capabilities for proposing ten central capabilities, she enumerates that "if people are systematically falling below the threshold in any of these core areas, this should be seen as a situation both unjust and tragic, in need of urgent attention – even if in other respects things are going well (Nussbaum 2000:71)". Perceiving certain condition as a threshold for all human beings is significant that it implies essential requirement, not an objective to be achieved.

Right to education does not just involve formal entitlement to schooling, but other economic, social, cultural, and so forth conditions which necessary for being

able to access to education and to learn. As Unterhalter describes, providing resources for schooling are not a guarantee of real opportunities for learning. When conventional right to education primarily focus on 'availability' of education, capability approach provides much fuller upholding of the right such as Tomasevski's '4 As' such as availability, accessibility, acceptability and adaptability.

As outlined by Sen (2005), the Capability Approach suggests broader evaluative space compared to the rights-based approach by focusing on the real opportunity aspect of freedom. While the right to education discourse is mostly concerned with securing political process of transforming choices into achievement, the Capability Approach also incorporates relations of people's choices and abilities.

Therefore, capability approach provides us a new evaluative framework that assessing whether the stated objectives are achieved does not always mean the successful contribution to results of development. Especially in basic education, which nowadays considered as legally guaranteed public service and of basic human right, it makes easier to be provider-centered which could be government in nation states and aid donors in development.

Additional important aspect of capability approach is recognition of education as conversion factor for other capabilities. In order to convert functionings and resources to real opportunity, education is essential not only as accumulating functionings but also in broadening human capability, the freedom. According to Sen(1999, p.294) a person can benefit from education "in reading, communicating, arguing, in being able to choose in a more informed way, in being taken more seriously by others and so on".

2.3.2 Diversities and Heterogeneities of learners

Human diversity and the pluralism has been the central concept of capability approach. Without understanding and respecting how different each individual is in terms of personal, social and environmental factors, it is more likely to be superficial to evaluate their attainment through education. The sensitivity of the capability approach to human diversity and its attention to look beyond extent of inputs and resources is another important contribution to education (Unterhalter, 2009a).

Why the capability approach stresses the importance of pluralism of individuals? It is because the concept of conversion is crucial. Sen (1992) argues that equalizing the ownership of resources “need not equalize the substantive freedoms enjoyed by different persons, since there can be significant variations in the conversion of resources and primary goods into freedoms” (p.33). Learners differ in various dimensions including personal interests for academic areas; artistic ability; environmental differences such as family income level or societies with different degree of education inequalities among gender, ethnic, race differences. There are no such groups consisted of individuals with no differences. As Terzi (2005) said, there is nothing inherently unequal about differences; however, some differences can become inequalities. For example, if the language taught in school is not what learner is using with their family and community members and not having enough chance to get used to it, it become a critical hinderance to convert accessing school to learning opportunity.

Generally, in development cooperation, the approach that we take to enhance

education accessibility is to provide equal resources such as school facilities, teachers, textbooks, and free tuition, so called the 'equal treatment'. However, the assumption that providing equal treatment will bring equal result, better access to education, has severe flaw. It sounds reasonable and logical to assume that when enough classrooms and textbooks are available, the enrollment may increase. Of course, having adequate educational environment is prerequisite for learning. The importance is that equal treatment is necessary but not sufficient in terms of expanding real freedom to learn.

For example, 'equally' poor toilet facilities in school does not equally affect boys and girls. It affects girls more seriously that it leads to absences during menstruation. Equally, it is necessary to be sensitive to evaluate the same academic achievement of students with different learning conditions. As Unterhalter (2007) argues, it is not fair to judge two students equally who failed the exams if one has chosen to spend most of time to go out with friends instead of studying, and the other has tried her best but has been unable to succeed because she had to take care of her little siblings every day after school.

According to Unterhalter (2007), 'evaluation should look at the condition of being educated, the negative and positive freedoms that sustain this condition and the ways in which being educated support what each and every individual has reason to value'. As mentioned, concept of rights *within* education is responding to the need to meet the prerequisites of learning. And the capability approach provides cogent reasoning why it is important to look at and evaluate more closely about those prerequisites, the condition of being educated. Because individuals and groups are situated in different circumstances, providing equal treatment and resources does not always guarantee the same results of learning opportunity.

2.3.3 Education as an instrumental value to expand other capabilities

In capability approach, its main perspective towards development is the process of expanding real freedom that people enjoy. Therefore, in this approach, expansion of freedom is viewed as both the primary end and the principal means of development (Sen, 1999). To what extent education can contribute to expand the real freedom? So far, education as core capability to be achieved as an end and its intrinsic value has been discussed earlier. In this part, its instrumental value as means to expand other capabilities will be focused.

When Sen (1999) explains about the concept of freedom, he emphasizes the ‘constitutive’ role of freedom. When development is defined as the process of expanding freedom, it automatically pays attention on what kinds of deprivation are there which hold people’s freedom. Because it is considering development as enriching process of expanding real freedom, defining and tackling down the obstacles and difficulties that impede the progress becomes primary concerns. As Sen points out that “the effectiveness of freedom as an instrument lies in the fact that different kinds of freedom interrelate with one another, and freedom of one type may greatly help in advancing freedom of other types (Sen 1999, p.37)”, education, as *means* to development, plays critical role to empower people to promote and expand other capabilities.

Educators and scholars have discussed for centuries to answer what should be the aim of education. In Chinese letter, education is written as 教育, the shape of ‘教’ represents a teacher holding a stick towards learners and ‘育’ means nursing

and parenting; therefore, in Asian cultures education generally means life-long process of forming oneself as a human since her birth with or without help of elders. In western culture, on the other hand, the word education is derived from the Latin word 'ēducātiō' which means dragging something out from, and brining up, in other words, to discover inside potentials and talents to outside. In other languages, education has different names such as 'paideia', 'pedagogy', and 'bildung'. Each has unique stress point; however, they share more or less the same idea and concept of education that 1) aims to develop potentiality, possibility and disposition that person already possess with or without help of others, 2) indicates the process of turning people from imperfection to perfection and immature to mature.

From what education means in different cultures and languages, it is found that the aim of education contains doings and beings of what development means. As the derivation of word development represents 'a gradual unfolding' and 'internal process of expanding and growing', the meaning of development and education share lots in common. In this connection, Sen's capability approach shows close linkage between education and development by adding a concept of freedom. According to the approach, the purpose of development should be expansion of real freedom to choose what people reason to value and it is not the final end to achieve but the process of gradual advancement. Therefore, in capability approach, education could serve both as ends and means of development (Saito, 2003) and the role of education 'for development and 'in' development could be distinguished (Yoo et al., 2019). Andresen et al. (2010) suggest that the capability approach might basically be considered as an educational approach.

Why is education important to expand other capabilities? Philosophers like

Sturma (2000) or Garrett (2001) argues that most of capabilities are formed by experiences and education, both formal and informal, and that, vice versa. Robeyns (2005) and Unterhalter (2003) also insist that “having access to an education that allows a person to flourish is generally argued to be a valuable capability” in their work. Sabina Alkire (2004) argues that education conduces “to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms”.

As conversion factor is one of the core concept of expansion of capabilities, the abilities of individuals to transform resources into valuable functionings could be obtained through education; therefore, person who is well-educated has advantages over not educated in terms of conversion ability to expand functional choices that they reason to value.

CHAPTER III. Quality Issues in Education Development Cooperation

3.1 Agenda Shifting from Accessibility to Quality

Global education agenda has changed, and its focus has changed. Though education is a stand-alone goal (SDG 4) and it still is considered as a central goal in SDG to achieve other goals including health and well-being (Goal 3), gender equality (Goal 5), decent work and economic growth (Goal 8), responsible consumption and production (Goal 12), and climate change mitigation (Goal 13) (UNESCO, 2017).

Understanding how the education development agenda has evolved and responded to countries' needs is necessary for assessing what is new for the education sector under the 2030 Agenda. As the first international education agenda, The World Conference on Education for All was held in Jomtien, Thailand, in 1990. According to UNESCO (2015), the conference was prompted by stagnating schooling rates in many parts of the world, a belief that human development should be at the heart of all development, and optimism generated by the end of the Cold War, all of which led to an ambitious call to support EFA.

In Dakar in 2000, the World Education Forum reaffirmed the EFA's six goals while also attempting to improve global education through resource mobilization, coordination, and monitoring mechanisms. The Dakar Framework for Action argued that strong national strategies would benefit from increased effective development cooperation support, despite the fact that the "heart of EFA activity lies at the country level." The partners agreed to make the Dakar Framework for Action known to "every international and regional organization, every national

legislature, and every local decision-making forum." Regular oversight and accountability would be ensured by coordination and monitoring mechanisms; countries would be aware that they were being watched and that actions found to be lacking could raise questions at home and abroad.

Five months later, the MDGs were approved. Their poverty-focused development agenda had a much narrower education focus: universal primary education (MDG 2), youth literacy (MDG 2), and gender parity at all levels of education (MDG 3). (UNESCO, 2019). As the dominant development narrative, the MDGs diluted the EFA message. The emphasis on universal primary education received support from the least developed countries, who were the furthest away from meeting it, as well as wealthy countries willing to finance it, but not from many countries that had already achieved it or were close to doing so.

Many of these flaws are addressed in the 2030 Agenda, which brings together the poverty arm (MDGs) and the environment arm (Rio process) of the previous development agenda. A key difference is that education is one of the goals (SDG 4) and is explicitly or implicitly incorporated into most other goals: education is linked to other development outcomes. Furthermore, the SDGs' universality recognizes that education systems all over the world face similar challenges in meeting the demands of sustainable development.

The SDG 4 targets are based on the EFA vision, but they also include additional commitments, such as universal secondary school completion, equal access to tertiary education, inclusive learning environments, and qualified teachers (UNESCO, 2019). Furthermore, the SDG 4 monitoring framework places a strong emphasis on outcomes – in other words, the effects of education on children, youth, and adults. Minimum reading and math proficiency levels, school

readiness, and digital literacy skills are among the outcomes (UNESCO, 2016).

Another significant shift is the analysis of educational participation and outcome indicators by individual demographic and socioeconomic characteristics. The gender parity index in enrolment rates by level of education was the only comparable MDG indicator. Other characteristics, such as location and wealth, have been included as a result of the increased availability of data sources from school and household surveys in the last two decades, allowing comparisons within and across countries.

A focus on educational content, particularly the mainstreaming of education for sustainable development and global citizenship through curricula, textbooks, and teacher training, is another difference from the MDG era. In addition, the inclusion of lifelong learning opportunities in the SDG 4 formulation indicates that education is broadly defined to include both formal schooling and its non-formal and informal aspects.

According to UNESCO (2017) SDG 4 differs from both the education-related MDGs and EFA as a global education agenda. While MDG2 focused on children and primary education, EFA had a broader agenda aimed at "meeting the basic learning needs of children, youth, and adults."

Global Education Agendas			
	MDG 2	EFA	SDG 4
Timeframe	2000-2015	2000-2015	2015-2030
Scope	Primary Education	Basic Education	Basic Education; Post Basic Education/Training; Lifelong learning
Coverage	Low-income countries	Universal but focus on lower-income countries	Universal regardless of development status of countries

Policy focus	Access to and completion of primary education for all	Access to quality basic education for all	Access to quality basic education for all; Equitable access to post-basic education & training; Relevance of learning for both work and ‘global citizenship’
--------------	---	---	--

This table is initially created by UNESCO (2017) and modified by the author

The MDGs' geographical coverage was limited to low-income and conflict-affected countries in the South, due to their narrower focus on ensuring children's access to and completion of primary education. In the case of EFA, this was not the case. While EFA was intended to be a universal agenda to ensure quality basic education for all in all countries around the world, in practice, it became focused on countries with the highest proportion of children out of school. SDG4 is, by definition, a universal agenda that applies to all countries in both the global North and the global South.

EFA was concerned with access to quality basic education for all children, youth, and adults, as opposed to the narrower MDG focus on access to and completion of primary education. Early childhood care and education, primary and secondary education, as well as youth and adult literacy and life skills, were all included. The SDG4 builds on the EFA's focus on quality basic education for all by expanding the agenda to include concerns about equitable access to post-basic education and training for youth and adults through equitable access to appropriate learning opportunities. The SDG4 also includes a new emphasis on the relevance of learning outcomes for both the workplace and citizenship in a global and interconnected world.

Though MDG 2, EFA and SDG 4 differ from each other in terms of policy focus, scope, and coverage, two main themes of ‘accessibility’ and ‘quality’ has

been the center of each global agenda. And it is worth asking through what kind of evaluation framework was used to assess the achievement of three goals and targets. As previously stated, the concrete theoretical basis or philosophical discussion in the evaluation of development cooperation is not explicitly developed. Nevertheless, it can be said that the perspective and approach of evaluation are reflected throughout the framework, indicators, and methodology of evaluation. In this regard, this chapter will look at the theoretical frameworks that underpin the evaluation of improving education 'accessibility' and 'quality'.

3.1.1 Assessing the Accessibility

The Millennium Development Goals and targets were signed in September 2000 by 189 countries. Among eight goals, the educational goal number 2 aimed 'to achieve universal primary education' and the progress has been monitored through four indicators: Net enrollment ratio in primary education; proportion of pupils starting grade 1 who reach grade 5; primary completion rate; literacy rate of 15-24 year-olds (UN, 2015).

The significance of indicators in evaluation will be discussed in detail in the following chapter. Nonetheless, the importance of education indicators during the MDG era cannot be overstated because they reflect the perception of education accessibility as well as impact on policy priorities. In the final year of MDGs, UN(2015) celebrated educational achievement referring to those indicators. The primary school net enrolment rate in developing regions increased to 91% in 2015, up from 83% in 2000. Globally, the number of out-of-school children of primary school age has decreased from 1000 million in 2000 to 57 million in 2015.

Between 1990 and 2015, the global literacy rate of youth aged 15 to 24 increased from 83% to 91%. According to the report (UN, 2015) “Sub-Saharan Africa has had the best record of improvement in primary education of any region since the MDGs were established. The region achieved a 20% increase in the net enrolment rate from 2000 to 2015, compared to a gain of 8% between 1990 and 2000. The literacy rate among youth aged 15 to 24 has increased globally from 83% to 91% between 1990 and 2015 and the gap between women and men has narrowed.”

As noticed, the achievement of accessibility has been evaluated as well as discussed through headcount of children enrolled in school throughout decades. In the MDG era, a lot of resources and efforts were put into increasing the enrolment rate. Schools have been built, and in the international community, each government has made efforts to make basic education free and compulsory (Tomaševski, 2001). As a result, we have achieved a lot and will continue to enjoy some more results. This is because some areas still lack schools, and if education is not legally mandated, children are being put into economic activities other than schools.

However, this approach of evaluating education accessibility has a number of problems that have consequences in achieving education targets because it neglects many other important aspects. First, accessibility to education should not be taken simply as school enrolment. Access to education means and requests more than just enrolling in school. Ensuring a certain level and minimum period of education should be covered. The mandatory basic education period legally protected by each country varies slightly, but many countries require more than nine years of education. Therefore, the scope of evaluation should be expanded beyond the indicators of whether school-age children have enrolled in school to

whether they fully enjoy the period of basic education protected as rights.

Through the 4-As scheme, Katarina Tomasevsky, the first United Nations Special Rapporteur on the right to education, explains how to define accessibility to education as a more comprehensive concept. It was created to track the extent to which 'education is available, accessible, acceptable, and adaptable,' as she put it. Because her focus was on how to measure the right to education, the government was given primary responsibility for ensuring the right; as a result, her 4-As scheme places a greater emphasis on institutional and country-level achievement.

According to Klees & Thapliyal (2007), the availability of primary education is dependent upon the government's duty to offer free and compulsory primary education. Availability does not imply that “people have an entitlement to all the education they may want, throughout their life, at government expense” (Human Rights, 15). Nonetheless, post-compulsory education may become mandatory as a result of human rights law's "progressive realization." Additionally, availability does not require only government-run education. “Human rights safeguards are orientated towards balancing the right of the state to compel children to be educated and the right of their parents to decide where and how” (Human Rights, 29). Following is what Klees & Thapliyal (2007) explained about Tomasevsky's understanding of 4-As:

Even if schooling is nominally available, accessibility may be affected by a variety of factors. Cost is a major barrier to access, prompting Tomasevsky to propose a fifth A: affordability. She examined other factors as well: gender, citizenship, migrant status, disability, race, ethnicity, language, religion, and imprisonment. Public education systems continue to be embedded with

inequalities, discrimination, and other obstacles that contribute to the underperformance of disadvantaged learners and/or their inability to complete compulsory education. In making this critique, Tomasevsky emphasized that the challenge of access is not as it is often characterized by mainstream education policies seeking to reach and integrate the unreached; rather, from a rights-based policy, “tackling exclusion requires halting and reversing exclusionary policies and practices, not only countering their effects” (Human Rights, 44).

Acceptability “requires a guaranteed quality of education, minimum standards of health and safety, or professional requirements for teachers which have to be set, monitored, and enforced by the government” (Education Denied, 51). Moreover, “the yardstick of acceptability would necessitate ascertaining what is—and is not—acceptable to people (including small people whom we call children) and changing the contents of teaching and learning accordingly” (Human Rights, 69). Acceptability also involves taking into account diverse issues such as indigenous and minority rights, language of instruction, textbook censorship, unregulated privatization, inadequate spending, and teachers’ rights (Human Rights, 69–99).

Adaptability “requires schools to adapt to children following the yardstick of the best interests of each child in the Convention on the Rights of the Child” (Education Denied, 52). Tomasevsky sees the previous three As as the outcomes of progressive stages in applying the right to education and the fourth A as the most recent stage and perhaps “utopian” (Human Rights, 103).

It moves education beyond the assimilationist stance, as access is improved and unreached populations are integrated. “The requirement upon children to adapt themselves to whatever education is made available to them is replaced by adapting education” to the child (Human Rights, 103). Adaptability recognizes that “children do not start school as equals” and that each child has “the right to be regarded as different” (Human Rights, 12)- a reality and an entitlement that have been obscured by the move toward standardized criteria for assessing learning. From this perspective, rights-based education critiques the very structure of schooling as a structure through which most children must fail and are branded as failures (Human Rights, 105–7).

Likewise, measuring net enrolment rate indicates partially. Putting this indicator up front and interpreting the whole picture of educational development is even more dangerous to distort the reality. It only focuses on whether children in school age entered primary school or not. In order to track how many years those children entered primary school continue their learning or how many of them complete the primary education, more indicators and data is required.

3.2 Assessing the Quality of Education

Agreeing with what Abby Riddell, senior lecturer at Harvard Institute for International Development, states that observing and measuring learning outcome is extremely difficult, unlike measuring mortality rates in health sector, because of its abstractness. It may took much longer than the project duration, even decades, to provide scientific and methodologically accurate, if ever possible, evidence to

prove what works and what not.

Therefore, most aid agencies have focused on achieving MDGs by providing school supplies, improving literacy rates, and increasing school attendance in order to shift public opinion, secure funding, and achieve results in a short timeframe. All of this appears to be great on paper, but it ignores the people and frequently fails to have a lasting impact on the community (Riddell & Niño-Zarazúa, 2016).

As previously discussed, the focus of global education agenda from MDG to SDG has been shifted to the quality issue. It is not because the goal of universal primary education is already achieved but because its progress has been nearly stopped. In 2019, four years after the adoption of SDG 4 and the promise to provide universal primary and secondary education, UNESCO(2019) provided the least UIS data showing that there has been no progress in reducing the global number of out-of-school children, adolescents and youth. In 2019, 258.4 million children, adolescents and youth were out of school, representing one-sixth of the global population of this age group.

What about children who are enrolled in schools? In the UN report (2015), it stresses that the progress of the completion rate in primary education has not been made as great as the enrolment rate has. It says “the proportion of adolescents aged about 14 to 16 years who had finished primary school increased from 70 per cent in the early 1990s to 81 per cent in 2008. This means that in 2015 one in every six adolescents in those countries, almost 100 million adolescents, are still not completing primary school.”

Is the approach to building a school and ratifying compulsory education policy no longer working? Was this effective in increasing the enrollment rate, but

not in keeping children in school? Some might agree that this approach was necessary and pre-requisite to make next level of educational development. In order to promote the right to education for children, it is certainly necessary to create an educational environment and establish a necessary institutional system. However, although the results to be pursued is the same, the process of reaching it would have been different.

3.2.1 Quality Education in SDG 4

SDG 4 aims to ‘Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all,’ (UN 2015). The idea of ‘quality’ education has changed throughout years of discussions and consultations. According to Unterhalter (2019), in 2014, at the UNESCO EFA steering Committee, the idea of quality had been linked with a narrow perspective of learning outcomes. However, by the later meetings taken place in Muscat in 2014 and Incheon in 2015, quality have been associated with equalities and broader values, such as sustainable development, global citizenship, skills for decent work (Unterhalter, 2019).

Among two meetings between Muscat and Incheon, the focus on quality education are slightly different. The summit in Muscat emphasized the need of free education as an important part of inclusion in the early years and at the basic levels. A significant statement on combating exclusion and striving for gender equality was included in the Incheon Declaration. It discussed the need for “transformative public policies to respond to learners’ diversity and needs, and to address the multiple forms of discrimination and of situations, including emergencies, which impede the fulfilment of the right to education” (World

Education Forum, 2015, 6).

Considering SDG 4 as the results of the most extensive consultations and discussions among major actors and interest groups in Education sector including UN agencies, international organizations, civil society organizations, academic networks, trade unions, and some interest groups such as disability, gender, or indigenous rights (Sayed and Ahmed, 2018), it is worth looking how the SDG 4 is formulated, as well as how the idea of quality education has evolved through time. Unterhalter (2019) highlights eight occasions when certain constituencies made significant contributions to formulating SDG4 in an abbreviated history of these talks. Following table summarizes the different versions of SDG 4 visions, definition of quality and adopting levels of education.

Target Focus	High Level Panel (May 2013)	EFA Steering Committee (April 2014)	SDSN Report (May 2014)	Muscat agreement (May 2014)	OWG SDG Proposal (July 2014)	UN SG Synthesis Report (Dec 2014)	Incheon Declaration (May 2015)	Education 2030 Framework for Action (Nov 2015)
Visions of the SDG education goal	Provide quality education and lifelong learning	Ensure equitable and inclusive quality education and learning for all by 2030	Ensure effective learning for all children and youth for life and livelihood	Ensure equitable and inclusive quality education and lifelong learning for all by 2030	Ensure inclusive and equitable education and promote lifelong learning opportunities for all	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
Definition of Quality	Meeting 'Minimum standards' of reading, writing and counting at primary level; meeting measurable learning outcomes at junior secondary	Developing capacity to learn, meeting global benchmarks of learning outcomes, fully participating in society, access decent work; express values(GCE & SD)	Wide range learning outcomes, reduce dropout in higher education develop knowledge economies	No definition for early years, primary and secondary. For youth and adults' increasing participation in society, enhancing peace, sustainable development and global citizenship	No definition of quality but associated with free & equitable schooling, affordable vocational educated and relevant skills	No definition of quality but discussion links education with dignity, sustainability, gender equality, and enhancing knowhow	Very full definition of quality looking at practice, includes rights and gender equality	Links quality with the humanistic inclusive vision of SDGs and with learning processes as well as outcomes

Levels of Education	Early years, primary, lower secondary & youth skills	All levels except higher education	Early years, primary & secondary, youth skills adult literacy. Academic environments	All levels	All levels	All levels implied; stress on people, planet and prosperity	All levels	All levels
Quantity and/ or quality	Quantity & Quality for primary and secondary	Quantity and Quality	Quantity and quality for early years, primary and secondary.	Quantity and quality linked for all levels	Quantity & quality for all levels	Relevant knowledge, skills, quality and access linked all levels	Quantity and quality linked all levels	Quantity and quality linked all levels

Source: Author reproduced the work of Unterhalter (2019)

As shown, goals and targets for SDG4 were agreed upon in 2015, but disagreements over the definitions of quality and equality persisted, hidden in part by a consensus-building politics. Tikly (2017) identifies areas of disagreement around EFA, which can be seen in several of the conversations around the SDG4 target setting at the eight events listed in the following table. Unterhalter (2019) identifies five disagreements between the narrow and broad approaches to defining quality and equality, as well as educational target levels: 1) Debates concerning education sub-sectors and whether to include early childhood, postsecondary, adult, and vocational education in the SDG targets; 2) Whether to focus on enrolment, attendance, and progression or quality; 3) How to define quality education, and whether it included free education, inclusion, and disputed values like sustainability, rights, and gender equality; 4) Whether equitable education meant extending formal educational rights to excluded groups or addressing intersectional inequities involving redressing disadvantage and transforming injustice within and beyond school; 5) How to engage with the significant growth of the private sector in education, an issue expressed partly as a dispute surrounding the nature of accountability.

SDG 4 consists of seven targets, three means of implementations and 43 indicators. There are three main themes in SDG 4 including quality, equity and inclusiveness as it aims to ‘ensure inclusive and equitable quality education and promote lifelong learning opportunities for all’ (UN 2015). According to UNESCO (2017) 43 indicators are categorized under concepts of learning, completion, participation, provision, readiness, skills, equity, policy, knowledge, resources, environment, number, qualified, trained, motivated and supported. And

among twelve global indicators presented in the table, quality related indicators are four including 4.1.1; 4.4.1; 4.6.1; and 4.c.1. There are two important problems to be dealt. First, considering that indicators are shaping how we view the world (Mair et al., 2017), those SDG 4 global indicators are clearly presenting that SDG framework is focusing on test results as quality indicator. Three of four indicators are about test results and one is about teacher qualification. Secondly, based on the tier classification, only one indicator 4.1.1 among quality related indicators is under tier I group. As tier classification is based on the level of methodological development and the availability of data at the global level, out of many SDG 4 indicators, only few quality related indicators that are provided by countries for at least 50 per cent of countries and of the population in every region.

Target	Indicator	Custodian Agency	Tier Classification
4.1 By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes	4.1.1 Proportion of children and young people (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex	UNESCO-UIS	Tier I
	4.1.2 Completion rate (primary education, lower secondary education, upper secondary education)	UNESCO-UIS	Tier I
4.2 By 2030, ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education	4.2.1 Proportion of children aged 24–59 months who are developmentally on track in health, learning and psychosocial wellbeing, by sex	UNICEF	Tier II
	4.2.2 Participation rate in organized learning (one year	UNESCO-UIS	Tier I

	before the official primary entry age), by sex		
4.3 By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university	4.3.1 Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex	UNESCO-UIS	Tier II
4.4 By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship	4.4.1 Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill	UNESCO-UIS, ITU	Tier II
4.5 By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations	4.5.1 Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated	UNESCO-UIS	Tier I/II depending on indice
4.6 By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy	4.6.1 Proportion of population in a given age group achieving UIS at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex	UNESCO-UIS	Tier II
4.7 By 2030, ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and	4.7.1 Extent to which (i) global citizenship education and (ii) education for sustainable development are mainstreamed in (a) national education policies; (b) curricula; (c) teacher education; and (d) student assessment	UNESCO-UIS	Tier II

nonviolence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development			
4.a Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all	4.a.1 Proportion of schools offering basic services, by type of service	UNESCO-UIS	Tier I
4.b By 2020, substantially expand globally the number of scholarships available to developing countries, in particular least developed countries, small island developing States and African countries, for enrolment in higher education, including vocational training and information and communications technology, technical, engineering and scientific programmes, in developed countries and other developing countries	4.b.1 Volume of official development assistance flows for scholarships by sector and type of study	UNESCO-UIS	Tier I
4.c By 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States	4.c.1 Proportion of teachers with the minimum required qualifications, by education level	UNESCO-UIS	Tier II

Continuing with the quality related SDG indicators, as shown in SDG agenda setting process which is one of the most comprehensive and largest number of experts are involved through the history of international development, defining

and contextualizing what quality education means in development context has been and still is in process. The problem is whether or not the definition of quality is agreed, the quality of education has to be evaluated.

According to Mingat (2003), when asked how to improve educational quality, the vast majority of educational professionals will likely respond, "increase the inputs into the system and improve the processes that take place." However, it appears that the dominant method for quality improvement is the establishment of predefined targets that are expected to produce an improvement that can then be measured by testing the learning outcomes. In World Bank (2002) reports, the term 'quality' also appears to be closely linked to learning outcomes that quality of education is measured by how far student achievement meets the defined minimum standards.

A common misconception among economists working within a human capital framework is that the concept of "quality" is synonymous with performance on standardized examinations. This helps to explain, in part, why participation of low-income countries in international evaluations such as the PISA test is so important (Tikly, 2017). They say that the quality of education, that is, the results measured by students' academic achievement, are more closely related to economic growth than simply how many years they attended school.

When academic achievement become the focus of targets, schools, teachers, and learners are pressured to achieve the highest possible test scores, regardless of whether this improves useful learning outcomes. This washback effect is most visible in low-income countries, where examinations that select learners for the next educational level are high-stakes for learners and their families as stepping

stones to eventual formal sector employment. In many SSA countries, schools are ranked based on their students' performance in end-of-cycle exams, children are over-tested, and teaching and learning experiences throughout upper primary and secondary school revolve around testing. Curriculum areas that are not tested or given less weight in selection procedures are undervalued, under-resourced, and receive less teaching time. The phenomenon of 'cramming' and exam anxiety, combined with low levels of professionalism among teachers, drives large numbers of children and youth to extra-tuition, eroding time for leisure and play. (Barrett, 2009)

Readily measurable cognitive outcomes shift from being privileged indicators of quality to defining quality. When this happens, qualitative indicators and scrutiny of processes can be overlooked (Alexander, 2008). UNICEF (2000) working document also highlights that quality has previously been focused on inputs and infrastructure, with the process only recently coming into focus. And the SDG indicator framework shows clearly that the quality of education is very much inclined to be determined through test results.

In part because indicators influence how we think about and approach the concept itself (Espeland and Sauder, 2007), they shape the way donors and partner countries think about quality education as being synonymous with academic achievement and away from more essential aspects of quality education that they do not monitor.

CHAPTER IV. Results-based Evaluation in Development Cooperation

In this chapter, evaluation theories, approaches and models which frequently adopted in development evaluation are introduced as compared to the Capability Approach in next chapter. As Osgood et al. (1957) said, evaluation is perhaps society's most fundamental discipline; it is an essential characteristic of the human condition; and it is the single most important and sophisticated cognitive process in the repertoire of human reasoning and logic.

Development theories and approaches have evolved along with dynamics of international development agendas as well as international economic politics. The emphasis of aid discourse has been moved from economic growth to poverty to well-being and the players are becoming more diverse. Considering the 70 year-length of development records of practices, discourses on development evaluation have relatively shorter history. The importance of evaluation in the development field was emphasized when discussions on aid effectiveness began. Criteria and frameworks for evaluation were presented by some international agencies, including the OECD DAC. Currently, most donor countries are conducting evaluations based on the six evaluation criteria proposed by the OECD DAC: Relevance, Coherence, Efficiency, Effectiveness, Impact and Sustainability.

As pointed out in several papers, it is difficult to find a solid theory, standardized evaluation model, or method in the evaluation of development cooperation. Instead, the principle of aid effectiveness mentioned earlier, the six criteria for evaluation proposed by the OECD DAC, are used as fundamental

principles for establishing an evaluation framework, and key evaluation questions are also made based on those criteria.

In this chapter, assuming that the development evaluation is part of the program evaluation, its general theories, models, and approaches are introduced with keen consideration of its relation to development perspective. Additionally, as a central theme of the development evaluation, the ‘results-based approach’ is introduced with its underlying logic, practices, and limitation.

4.1 Development Evaluation and Quality Measurement

4.1.1 Program Evaluation

Objects, so-called evaluands, of evaluations varies. There are evaluations of program, personnel, commercial products and services, organizations, manufacturing, governance, policies etc. depends on what is evaluated. The scope of evaluation application could broadens even greater when one considers the wide array of disciplines, activities and endeavors to which evaluation applies (Stufflebeam & Coryn, 2014).

In this study, theories, approaches, and models are reviewed focusing on program evaluation because it is what development evaluation has rooted and heavily influenced by. Some may argue that development evaluation has become an independent discipline and a field of study; however, for me, its theoretical foundation seems still controversial, and the empirical evidence are not yet fully proven. I agree with many scholars and researches addressing that program evaluation tend to be pragmatists (Scriven & Scriven, 1998) and “have not

systematically generated and tested propositions from their conceptualizations of program evaluation nor used such findings to improve those conceptualizations” (Smith, 1993). Therefore, I take development evaluation as one of many sub-categories of program evaluation in this study.

There are main historical milestones in the evaluation field and key theorists contributed to deepen the theoretical foundation. Evaluation is perhaps society’s most fundamental discipline (Osgood et al., 1957); therefore, the history of evaluation might go back to further than we think. In this chapter, because we give emphasis on development evaluation, its first historical milestones come along with the history of international development.

Time	Features	Evaluation examples or Key theorists
1800-1900	The Age of Reform Industrial Revolution; Increasing concern for improvement of educational and social program in the UK and the US	Behavior recording; Mental tests; Written examinations; Spelling tests; Use of external inspectors
1900-30	The Age of Efficiency and Testing Systematic and empirical studies of program effectiveness, Use of standardized tests; Emphasis on local needs and questions	Establishment of centers specializing in school evaluation; Standardized tests; School survey movement Objective tests
1930-45	The Tylerian Age Goal-driven; Local needs and questions	The Eight-Year Study
1946-57	The Age of Innocence Expansion in standardized testing: Tylerian approach dominant	Key Theorists: Tyler, Gulliksen; Lindquist and Bloom

1958-72	The Age of Expansion Development of evaluation models and theories; Professional specialization of evaluation; Program evaluation	Key theorists: Caro; Campbell; Cronbach; Provus; Sandres; Stake; Stanley; Stufflebeam; Tyler; Weiss; Worthen
1973-present	The Age of Professionalization Development of professional journals; Pluralistic approaches/methods	Key theorist: Guba and Lincoln; Stufflebeam; Weiss, Stake; Worthen; Sanders; Rossi and Freeman

Source: Madaus, Scriven and Stufflebeam (1983), Stufflebeam, D. L., & Coryn, C. L. (2014)

4.1.2 Classification of Evaluation Approaches and Models

In general, the term theory has been used as somewhat synonymous with approaches or models in evaluation studies (Carden & Alkin, 2012). In theories, there are descriptive or empirical theory that would describe what would necessarily happen given particular sets of actions. And there are prescriptive theories. Theories of evaluation present more characteristics of the prescriptive theories because evaluators simply do not know the results of the particular evaluation actions taking place within various contexts. Evaluation is often considered as a professional practice rather than an academic discipline.

Nevertheless, there are studies which did the theoretical survey in evaluation. Following table shows the recent scholars' work of evaluation theory from Scriven to Stufflebeam. They tried to classify approaches and models of evaluation based on their perspectives. Because approaches and models were not developed based on certain theoretical framework, method, range, and purpose of each approach were different. Therefore, those classification made by following theorists: Scriven, Worthen & Sanders, Bohla, Shadis, Cook & Leviton, Alkin and Stufflebeam categorized approaches not based on theories or framework ideas but

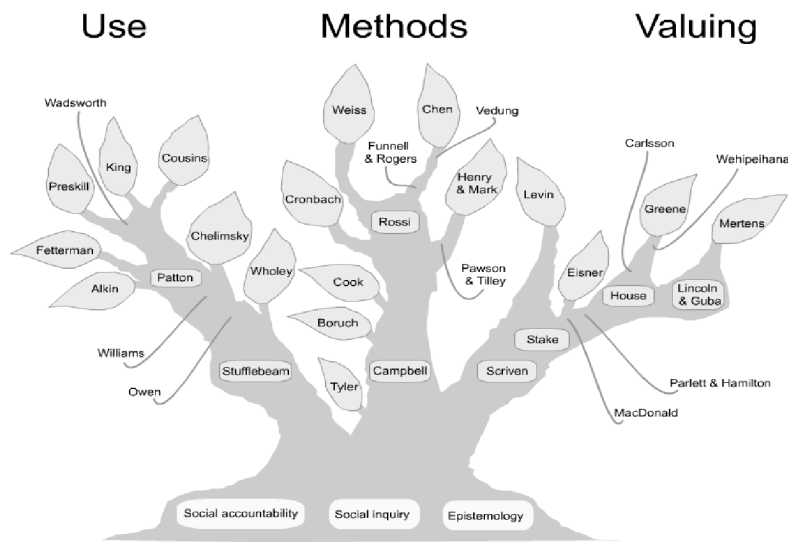
on primarily methods and use of evaluation. As an example, most recent work of Alkin’s theory tree and Stufflebeam’s five classifications will be described in detail.

Theorists	Classification
Scriven, 1976	Summative/formative evaluation
Worthen & Sanders, 1987	Six approaches
Bhola, 1990	Internal/external
Shadish, Cook and Leviton, 1991	Three stages and five theoretical dimensions
Scriven, 1993	Six views
Alkin, 2012	Evaluation theory tree
Stufflebeam, 2014	Five classifications of twenty-six approaches

(1) Alkin’s Evaluation Theory Tree

Alkin (2004) claims that “While theory is conventionally used in evaluation literature, in some ways, it would be more appropriate to use the term approaches or models”, he tried to classify theories and theorists in the field of development under three branches: use, methods and valuing. In the following chapter, the work of Alkin, Christie, and Vo (2012) will be introduced. It is one of the most recognized map of evaluation theory in the field but not necessarily the most comprehensive.

There is an attempt to classify those prescriptive theories in evaluation by Alkin and Christie in the Evaluation Roots (2002, 2012). They argue that all prescriptive theories must consider: (a) issues related to the methodology being used, (b) the manner in which the data are to be judged or valued, and (c) the user focus of the evaluation effort. Each theorist has been presented on the tree on branches of methods, judgment/valuing, and use.



Source: Alkin (2004)

In the central branch of the Evaluation Theory Tree, the Methods branch, evaluation is primarily guided by research methodology. D. T. Campbell (1957) and D. Campbell, Stanley, J. C. (1966), which defines the conditions for appropriate experimental and quasi-experimental designs, are the founding theorists who paved the road to further studies. Shadish et al. (1991) mentioned that the theorists on this branch are typically concerned with knowledge construction in the most rigorous manner possible.

To the right of the Method branch is the Valuing branch, which was initially inspired by the work of Michael Scriven. He claimed that evaluation is not evaluation without valuing (1967). Robert Stake is also an influential theorist to split the branch into two: objectivist and subjectivist. In the objectivist's view, it is the role of the evaluator to do that valuing. On the contrary, in the subjectivist's view, dynamic process and a truth is always relative to some particular frame of reference.

The third branch of the tree is Use, which originally focused on an orientation toward evaluation and decision making (Carden & Alkin, 2012). Theorists on this branch focused specifically on those empowered to use the information. The authors of the tree proclaim that three branches are not meant to be viewed as independent from one another.

(2) Stufflebeam’s Five Classification of Twenty-six Approaches

In 2014, Stufflebeam and Coryn classified program evaluation approaches into five categories. The first category includes approaches that promote invalid or incomplete findings, and the other four include approaches that agree, more or less, with the definition (questions and/or methods-oriented, improvement/accountability, social agenda/advocacy, and eclectic).(Stufflebeam & Coryn, 2014)

Classification	Approaches and Models
Pseudoevaluations	Public Relations-Inspired Studies Politically Controlled Studies Pandering Evaluations Evaluation by Pretext Empowerment Under the Guise of Evaluation
Questions- and Methods- Oriented Evaluation Approaches	Objectives-Based Studies Accountability, Payment-by-Results Studies Success Case Method Objective Testing Programs Outcome Evaluation as Value-Added Assessment Performance Testing Experimental Studies Management Information Systems Cost-Benefit Analysis Clarification Hearing Case Study Evaluations Criticism and Connoisseurship Program Theory-Based Evaluation Mixed-Methods Studies
Improvement- and	Decision and Accountability oriented studies

Accountability-Oriented Evaluation Approaches	Consumer-Oriented Studies Accreditation and Certification
Social Agenda and Advocacy Approaches	Responsive Evaluation or Client-Centered Studies Constructivist Evaluation Deliberative Democratic Evaluation
Eclectic Evaluation Approaches	Utilization-focused Evaluation

4.1.3 Development Evaluation

In previous section, literatures of program evaluation approaches and models are reviewed. Assuming that the background ideas and practices of development evaluation are closely linked with the features of development cooperation at each period, it is meaningful to review the history of international development cooperation along with the evaluation theories which reflected on. The following table is modified based on the work of Hjertholm and White (2000). As authors mentioned, a comprehensive review of global aid evaluation seems missing. I also found its synthesis is not compatible and sometimes contradictory; however, it provides overall picture of available literature and trends in the field of foreign aid.

To give a brief review of the history of aid evaluation, it seems natural to provide a history of foreign aid because the theories and practices are closely linked and reflected at each period. In the mid-1940s, the underlying theory of foreign aid had been anchored by economics. Economists provided the main theories and models for explaining how aid works (Clements, 1996). In this period, the main focus of aid was reconstruction of infrastructures, such as transportation and electricity facilities in Europe. In 1950s, during the cold war, as the United States and the Soviet Union competed on providing aid for development countries. The main focus was community development and the typical type of aid as project-

based aid is formulated and has been used until now. In 1960s, bilateral aid agencies are established including France (1961), Germany (1961), Japan (1961), Sweden (1961), Belgium (1962), Norway (1962), Netherlands (1964), Canada (1968) and others. The U.S Economic Cooperation Agency, originally established in 1951 for the Marshall Plan, was renamed as USAID in 1961. In this period, economic cost-benefit and cost-efficiency analyses gained much methodological attention (McKean, 1996; Dorfman, 1963; and Mishan, 1971). In 1970s, the main sector of aid shifted to social sectors such as health and education, reflecting an ideological change in aid community, from economic development to poverty reduction. During this period, the experimental and rigorous quasi-experimental designs were advocated and some of those methodologies were adopted in large-scale aid evaluation in Colombia (1971-75) and Nicaragua (1974-78). In 1980s, the market-based reform has been adopted through the famous structural adjustment program initiated by the World Bank. It stimulated many NGOs to be established and the movement of participatory and empowerment evaluation approaches raised (Chamber 1988). In 1986, Cassen's remarkable study *Does Aid Work?* was published and it raised issues of aid effectiveness as the first trial by reviewing major aid projects. One of the most influential accomplishment of donor agencies in 1980s is the Project Cycle Management (PCM) system. It is firstly developed by the EC (now, the EU) and it spread quickly to other donor agencies. In 1990s, the famous DAC five evaluation criteria has been approved. The PCM system adopted the DAC evaluation criteria and became widely used as a management tool as well as evaluation approach among major donors. Including USAID and other aid agencies started to focus on evaluating performance

measurement and the outcome monitoring. Its strategy and approach became generally known as Results Based Management (RBM). The Asian Economic Crisis in 1997 affected the approach of aid shifted from independent projects to sector-wide programs.

	Dominant or rising institutions	Donor ideology	Donor focus	Types of aid	Evaluation in int'l development cooperation	Key theorists/ evaluation work
1940s	Marshall Plan and U.S system (including WB)	Planning	Reconstruction	Marshall Plan (largely program aid)		
1950s	United States, with Soviet Union gaining importance from 1956	Anticommunist, but with role for the state (e.g., comprehensive planning & national five-year plans)	Reconstruction continued. Also focused on community development movement	- Food aid - Projects aid		Lewis, Hayes
1960s	Establishment of bilateral programs	As for the 1950s, with support for state in productive sectors	Productive sectors (e.g., support to the green revolution) and infrastructure	- Bilaterals gave TA & budget support - Multilaterals supported projects	The boom in establishing bilateral aid agencies - Sophistication of economic cost-benefit analysis for project preparation	Campbell & Stanley, Hirschman
1970s	Expansion of multilaterals, (especially WB, IMF, and Arab-funded agencies)	Continued support for state activities in productive activities and meeting basic needs	Productive sector continued. Increasing focus on poverty, taken as agriculture and basic needs (social sectors)	- Fall in food aid - Start of import support	The boom in establishment of evaluation units - focus on the process of aid - Large-scale experiments - Logical framework developed by USAID	Gittinger, Tendler, Squire & Van der Tak, Rossi, Freeman & Wright
1980s	Rise of NGOs from mid-1980s	Market-based adjustment	Macroeconomic reform	- Financial program aid	Focus on aid effectiveness by various approaches	Scriven, Casley & Lury,

		(rolling back the state)		- Debt relief	- Rapid, low-cost methods - Participatory approach - Training & empowerment	Cassen, OECD, Riddell, Chamber et al., Cassen
1990s	Eastern Europe & FSU become recipients; emergence of corresponding institutions	Market-based adjustment continued	Environment and gender (but passed quickly)	Financial program aid & debt relief continue	Diffusion of DAC's evaluation criteria vs. adoption of "managing results" approach - Sector level, country level, and thematic evaluation	OECD-DAC, Osborne & Gaebler, World Bank
2000s	Aid coordination forum at local level as well as headquarters level	Move back to role of the state. (balance the market & the state)	Poverty and then governance	Move toward sector support	New agenda: sector program evaluation, MDGs New agenda: Evidence-based evaluation	World Bank, Wlfenson, Hatry

Source: Author modified the work of Hjertholm & White(2000) and Sasak

4.2 Results-Based Approach in Development Evaluation

The results-based approach has been widely adopted in development field since 1990s with great promotion by World Bank and USAID. For last 30 years of dominant use of results-based approach as a management as well as evaluation framework, it provided efficient tool so-called logical framework or results-framework for development providers to monitor implementing process of projects and programs(OECD, 2019). In addition, this approach certainly promoted understanding of what outcomes and results are expected among stakeholders by providing logically valid matrix of project design.

In this research results-based approach is not reviewed as one of program evaluation models or theories but as a conceptual framework for development evaluation as compared to Amartya Sen's Capability Approach. Results-based approach is often explained as a counterpart of implementation-based approach in development evaluation. When the approach was introduced in 1990s, 'activities' had been the main focus of management and evaluation in development practices. As the question of aid effectiveness has been raised, development provider started to look for the results of their programs and changes that occurred in recipient countries.

After a long debate between aid pessimists like Easterly (2006) and Moyo (2009) and aid optimists like Sachs (2014) and Sen (2006), there is now broad consensus that aid does contribute to growth in general, though quite modestly, and that it can continue to do so despite its many flaws (Picciotto, 2018). In both

partner and donor countries, there is increasing pressure to scientifically prove the impact of development interventions in order to maintain aid and respond to accountability concerns. There are two generally agreed purposes of evaluation in development cooperation: accountability and learning. As Kogen (2018) argues that ‘despite decades of practice, we in the West do not fully understand how to “do” international development’; that the accountability and learning terms ‘have been used too close to interchangeably by the world’s major donors’ and that ‘while accountability is relatively well-established as a concept, learning, as a cross-cutting conceptual goal of evaluation’, should be rediscovered.

Yes, accountability is a common requirement for all democratic governments and evaluation helps to demonstrate that official promises are kept through fair, impartial, and transparent processes. In addition, evaluation is needed to draw lessons-learned to improve the next phase of a program or provide information and data to related development programs.

The results-based approach is originally used in program management sector. And it is now widely used both in management and evaluation in development practices. Since the launching of the initiatives of MDGs, government of developing countries were pushed to adopt public management systems that shows results. The effectiveness of aid has been questioned and officially discussed through several occasions including the famous Paris Declaration on Aid Effectiveness and the role of evaluation to prove the effectiveness is reaffirmed.

The results-based evaluation takes dominant position in development evaluation (OECD, 2019). Every OECD DAC donor country and UN agency are adopting and promoting this approach, with the lead of World Bank. Setting

international development agenda, such as MDGs and SDGs, is also making a favorable environment for this approach to be used. The results-based approach is not an evaluation theory nor a method. It is more like a finger pointer that draws out attention towards what we want from evaluation: proving whether it works or not. In evaluation, there are growing number of approaches and models. To make less confusion, the term ‘approach’ used in results-based approach contains rather broader and overarching meaning compare to what it meant in the work of (Stufflebeam & Coryn, 2014). The term approach found in evaluation literatures are used in various levels of notions and sometimes its scope and use seem somewhat ambiguous.

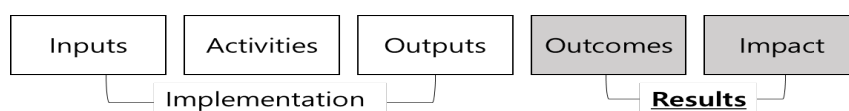
4.2.1 What is results-based evaluation?

1) Concept and Definition (Terminology)

Results-based approach in development evaluation is normally understood as the assessment of outcomes and impacts, or more generally of results of a project or program (Morra-Imas et al., 2009). This evaluation approach was introduced and conducted in other evaluation sectors much earlier than that of development field. For example, it is widely used in education sector by comparing to the process-based evaluation. When results-based evaluation is issued in education sector, the result generally represents the change of students’ behavior or learning outcome. It is often called objective-based evaluation in education sector and it will be explained in more details in following section.

In development evaluation, the results-based approach is notably promoted

by the World Bank and USAID since 1990s. It is often compared with the evaluation approach that they called ‘implementation-based evaluation’ which focuses on the assessment of inputs, activities, and outputs (Kusek & Rist, 2004). Therefore, the term ‘results’ in development evaluation particularly indicates ‘outcome’ and ‘impact’ of a project or program which stated in its theory of change or logic model (figure 1).



<Figure 1> Results Chain

Because the results-based evaluation is often explained through the theory of change, it is important to understand the main components of a theory of change. According to Kusek and Rist (2004), theory of change is a representation of how an intervention is expected to lead to desired results. Theory of change models typically have five main components: inputs, activities, output, outcome, and impacts.

Component	Description
Inputs	Resources that go into a project, program or policy (funding, staffing, equipment, curriculum materials, and so forth)
Activities	What we do. Activities can be stated with a verb ('market', 'provide', 'facilitate', 'deliver')
Outputs	What we produce. Outputs are tangible products or services produced as a result of the activities. They are usually expressed as nouns. They typically do not have modifiers. They are tangible and can be counted.
Outcomes	Why we do it. Outcomes are the behavioral changes that result from the project outputs (quit smoking, boiling water, using bed nets). Outcomes can be increased, decreased, enhanced, improved, or maintained.
Impacts	Long-term changes that result from an accumulation of outcomes. Can be similar to strategic objectives.

Source: Kusek and Rist (2004)

According to the World Bank handbook, results-based evaluation is designed to address the 'so what' questions such as so what about the fact that outputs have been generated? So what that activities have taken place? So what that the outputs from these activities have been counted? Etc. To answer those questions, the main evaluation questions of results-based evaluation are based on these three questions: What are the goals of the organization? Are they being achieved? How can achievement be proven?

2) Results-based evaluation in development

The term results-based management was first used in the 1990s and its concept is originated from other policy areas since the 1960s (OECD, 2019). In development cooperation, RBM has been highlighted and broadly launched in

many aid agencies in the 1990s and its importance has been once again confirmed by the Paris Declaration on Aid Effectiveness in 2005. As the aid effectiveness is questioned, multiple stakeholders in development cooperation had been under pressure to present ‘the results’ that they achieved throughout projects and programs.

There are challenges and unintended consequences of RBM and many ‘alternative’ initiatives and approaches have been presented; however, RBM has been implemented in development co-operation over the last 20 years with tremendous support and promotion of major development agencies notably by OECD and World Bank.

The 2002 OECD DAC Glossary of key terms in evaluation and results-based management provided somewhat narrow definition of RBM as “a management strategy focusing on performance and achievement of outputs, outcomes and impact” (OECD, 2002). And its definition has been enlarged when defining Managing for Development Results (MfDR) as “global development assistance can be made more effective by enhancing country ownership, aligning assistance with country priorities, harmonizing development agencies’ policies and procedures, and focusing more consistently on the achievement of development outcomes” (OECD, 2006).

World Bank (1993) has been the leading organization to stress the importance of results-based monitoring and evaluation (M&E) system. According to World Bank, there are 10-steps guideline for building a results-based M&E system. And following actions are the essential steps: formulate outcomes and goals; select outcome indicators to monitor; gather baseline information on the current

condition; set specific targets to reach and dates for reaching them; regularly collect data to assess whether the targets are being met; analyze and report the results (Kusek & Rist, 2004).

Based on the agreement on what results-based evaluation should contain, the emphasis and focus of what to, why and how to evaluate program seem explicit. It stresses the importance of setting outcomes and goals beforehand because the most important reason for the results-based evaluation is to determine whether intended targets are achieved or not. To have more reliable evaluation results, numerous evaluation models and research methods are adopted including experimental evaluation. Regardless of what kinds of methods and models are used, the fundamental paradigm of results-based evaluation is to find out whether the stated goals are achieved in a systematic and logical way of proving.

3) Objectives-based Evaluation and Results-based evaluation

According to Tyler (2000), objectives-based study is first developed for evaluation activities and published in 1934 under the title *Constructing Achievement Tests* (Tyler, 1934). The background of the study is to improve the instruction of undergraduates in the Ohio State University. There was a great concern over the fact that a large percentage of the freshman students fail or drop out. And the university's Bureau of Educational Research believed that teaching and learning in the university could make improvement with aid of relevant research, particularly with the use of tests and measurement.

Tyler started his research in the biology courses. One of the main problem he

found was that the instructors were using tests that demanded only that students recall specific information. This kind of tests gave the students the wrong notion that they focus on memorizing specific information to get good test results and not care about demonstrating the behaviors that the instructors expect students to obtain; ways of thinking, feeling, or acting. To overcome such limitation of achievement tests, Tyler worked with biology instructors and the first step was to identify the educational objectives of the courses. Instructors were requested to work out definitions of the objectives, expressing them in terms of *behavior* and *content*. Then he published the article “A generalized Technique for Constructing Achievement Tests” that contains what he learned throughout the work in the Ohio State University.

The article involves seven steps to form an educational achievement tests and it became the first work of what is now called ‘objective-based evaluation’. Ralph Tyler is generally acknowledged to be the pioneer of the objectives-based type of study (Stufflebeam & Coryn, 2014). Later Percy Bridgman, and Edward Thorndike are also credited and Bloom, Englehart, Furst, Hill, and Krathwohl (1956); Hammond (1972); Metfessel and Michael (1967); Popham (1969); Provus (1971); and Steinmetz (1983) have developed variations of Tyler’s model. Tyler originally developed the objectives-based approach for evaluating educational programs. However, its influence over many other areas of study is recognized and it is particularly seen in many government funded programs to determine the extent to which each program achieved its objectives.

Goal-achievement (objective-based) model considers that merit of the program is to be equated with success in achieving a stated goal (Stufflebeam et

al., 2000). In this kind of evaluation, it is crucial to identify objectives and goals to be achieved and measured at early stage of both program design as well as evaluation design. Because achieving the stated goals is considered as the success of the program, evaluation questions, indicators, methods are also influenced by the expected results.

What results-based evaluation means in development evaluation; therefore, shares lots in common with the objective-based evaluation in general evaluation as well as education field of studies. As definition and emphasis of results-based evaluation is mentioned earlier, both results-based evaluation and the objective-based evaluation focus on achieving the stated results in other word, objectives.

4.2.2 Logical Framework, Logic Models and Results-based management

Logical framework (also called logframe) approach has been widely used as a classic tool of aid management (Gasper, 2000). Some donor agencies, including KOICA, require that (ex post) evaluations must use logframe. Development programs and projects have their unique target, scope, scale, and content; however, they have all had one solitary linking factor that they are designed, managed and evaluated by use of the logical framework approach (Bell, 2000). This approach is originating at USAID in early 1970s and then spreading rapidly to UNDP.

Until the 1990s, the 'project model' was dominant type of development assistance. The importance of managing those projects were focused on project staff producing 'deliverables'(Conlin & Stirrat, 2008). Logical framework was a tool to contain how these deliverables were to be delivered throughout the project

with the presumed links between the inputs, outputs and outcomes as well as assumptions underlying these links. Purpose of evaluation at that time was to find out whether or not these deliverables are delivered.

This project-based approach of development was associated with the context of donor-centered tendency of development work at that time. Donors provided funds and donor agencies owned the projects rather than the partner countries (Conlin & Stirrat, 2008). However, since the 1990s there has been a major shift in the landscape of the development assistance. The first was the increasing demand for 'management for results.' The MDGs and SDGs required governments to provide aggregated data indicating whether or not the targets were met. The second factor was the increasing variety of types of development assistance from project to program to budget support. The third factor is that the players are changing. Governments and IOs have been replaced by trade and private investment. Finally, there are cross-sectoral issues to consider. Climate change and security have broadened the scope and become global issues.

Logframe is one of tools to describe a project or, intervention in development practices. As the following table shows, the four by four matrix contains: 1) a hierarchy of levels of objectives for a project/intervention; row in the matrix correspond to different levels of objectives, which are described in general terms in the first column. 2) Indicators of the fulfilment of objectives, and typically also targets and sources of information, for each of the objective levels. Usually the matrix has four columns, with the second and third columns on the measurement and data issues. 3) Sets of assumptions, concerning conditions required for the desired project story to happen, and notably about factors external to the project

(Gasper, 2000). Assumptions are filled in the final column. This version of logframe is firstly introduced by USAID in the early 1970s and was long dominant (Coleman, 1987).

Hierarchy of Objectives	Performance Indicators	Data Sources	Assumptions & Risks
Impact/Goal (Longer-term project impact)	(Measurable indicators for Goal)	(Data sources for verifying status of Goal-level indicators)	(Assumptions/risks between Goal and Super-Goal)
Outcome/Purpose (Near-term project impact. The essential motivation for undertaking the project)	(Measurable indicators for End-of-project Impact)	(Data sources for verifying status of Purpose-level indicators)	(Assumptions/risks between Purpose and Goal)
Outputs (The deliverables of the project)	(Measurable indicators for Outputs)	(Data sources for verifying status of Output-level indicators)	(Assumptions/risks between Outputs and Purpose)
Activities (Smaller work packages needed to accomplish each Output)	(Budget Summary)	(Data sources for verifying status of budget and activities)	(Assumptions/risks between Activities and Outputs)

Source: 'A late 1990s version of the logical framework' (Social Impact, 1997)

As Gasper (1997) stresses “logical frameworks – also known as logframe, project frameworks, or project matrices – have thus become a foremost example of the rise in public and development work of the type of modern managerialism which emphasizes statement of hierarchically ordered and, as far as possible, quantified objectives”, it has become widely used and influential tool among aid agencies and organizations around the world. In many cases its use has become obligatory by aid agencies including Sida, the World Bank, JICA, KOICA and numerous NGOs.

Because the results-based approach gives great emphasis on the outcomes and

impact compared to outputs and activities, it is convenient to design and explain the logic of programs with logical framework because they share the same notion of inputs, activities, outputs, outcomes, and impact. There are neither explicit rules or guidelines to use logical framework for results-based evaluation, nor the theoretical background; however, it is likely that the logical framework provided foundational ideas and perspectives that underlies the core concept of results-based approach.

The logical framework provides a convenient overview of programs and their objectives. According to Gasper (2000), a clear hierarchy of components of logframe converges on a single goal and a set of measurable and time-bound performance indicators are presented. It is effective and convenient to manage projects and programs as well as monitor and evaluate its results systematically with set of indicators. Sometimes it is also used as a communication tool between donor and partner countries and various stakeholders to collect extensive opinions and respective interests. For this reason, using logical framework on development projects and programs had been successful to meet the needs of efficient management as well as accountability-oriented evaluation.

However, the world of development assistance has changed over time. According to Conlin & Stirrat (2008), OECD (2016) and OECD (2019), the modalities of development assistance moved from project-based to many such as sector wide approaches and budgetary support. A range of development actors are also varied to private investors and corporates. Furthermore, in relation to adopt new global agenda of 2030, more development goals, targets and indicators are covered.

Those major changes and shifts in development industry call for more context sensitive and alternative development evaluation including scope, methods and theories. According to Conlin & Stirrat, (2008), “various forms of evaluation based on models derived from project logical frameworks are no longer appropriate to complex types and approaches of development assistance in the 21st century, nor to deal with the increasingly wide array of stakeholders involved in.”

To meet the needs of changing development environment Amartya Sen’s capability approach as an evaluation framework will be introduced and discussed in Chapter III. Before moving to explore what is the meaning and value of the capability approach and how to apply in development evaluation, I would like to discuss about three main challenges of RBM in following section.

4.3 Challenges of Results-based approach

First, when compared to what is actually being discussed and promoted among development agencies, RBM appears to be somewhat specious in terms of its utilization. According to OECD (2019), though most of development cooperation providers adopt to implement RBM in their organization, only the handful of agencies including UN, GEF and SDC have formally defined a purpose for RBM. CGIAR (2017) evaluation notes that the fact that the organizations have not specified why they are doing RBM has reduced the learning potential, created confusion about what RBM was meant to do within the organization and undermined the motivation and the merits that RBM supposed to bring.

OECD (2017) reviewed seven case studies of RBM by development

cooperation providers and it found that those organizations are at different stages in implementing RBM. It is surprising that RBM has been implemented in most development agencies over the past 15-20 years; however, there are not many agencies have a clear perception of RBM and a well-defined plan under implication for system-wide operation. To conclude, a number of providers do not communicate and understand clearly about the purpose of the RBM system and how it contributes to their development results even though its definition has been developed and shared explicitly at the OECD MfDR meeting as well as in the Paris Declaration.

Why is RBM system not systematically used in development providers? There are many challenges and difficulties including a weak results culture and lack of guidance on RBM; structural and system issues; measurement and method issues (OECD, 2019). OECD (2017) showed that many development agencies tend to adopt a ‘dual track’ system which serves domestic accountability and communication from management of projects and programs. Because each development providers has their respective context and culture of responding to their domestic political and public requests, RBM requirements that is made at international level could become an extra work with the lack of a results culture within the organization as well as of appropriate incentives. For example, the Norad evaluation notes that a results and learning culture is not yet in place; staff do not systematically seek out and learn from results data and evaluations. The CGIR (2017) evaluation also notes that the lack of a shared conceptual understanding of RBM reduces its learning potential. Some evaluations such as the UN evaluation argue that the lack of a results culture is due to staff attitudes

lacking a results-oriented ‘mind-set’. The Finland evaluation link the lack of a results cultures to management challenges which includes a lack of guidance or understanding within the organization on why RBM is important and how to implement it in practice. The Norad evaluation states that “commitments to being ‘results oriented’ and ensuring, ‘funds deliver results’ are consistently found in government documents, but there is no detail on what this should look like in practice” (Norad, 2018:7).

Besides the lack of a results culture and clear guidance on utilizing RBM in each development organization, there are more problems remained. OECD (2019) found that many development providers are facing inconsistencies and disconnection between policies and budget. While the goals and objectives are often specified at a policy level, these are not operationalized in the funding allocations. This makes it very challenging for the agencies to practice RBM at a strategic level. The United States Department of State evaluation, for example, noted that responsibilities and requests for managing results have increased; however, corresponding policies, guidance, incentives, and staff resources have not expanded (US, 2015).

There are also measurement and method issues with collecting reliable results and data for RBM. From the donor perspective, having inadequate and inconsistency data tracking tools and measurement formats at different period of projects and programs make it difficult and, in most cases, not possible to measure the long-term results and changes. What makes it worse in development evaluation is that the partner countries’ data systems are generally very weak; the absence of impact indicators; poor credibility and inadequate quality of data (WB, 2017; GEF,

2017; Finland, 2015; AfDB, 2018).

Second, the development industry has changed rapidly to more complex and fluid forms over the past 20 years and the usefulness of RBM becomes more questionable. Until the 1990s, the ‘project model’ was the major form of development assistance which tended to focus on whether or not the project ‘deliverables’ had been delivered. These deliverables were set out in a ‘logical framework’ which contains links between the inputs, activities, outputs, and outcomes. Traditionally, those project models of development assistance laid stress upon achieving outputs.

However, since the Millennium Summit of 2000, the ‘Millennium Development Goals’ (MDGs) has provided an overarching results framework within which all development projects should fit. MDGs became a set of development targets with specific indicators to all development providers. What is important about these goals is not just the emphasis on objectives rather than outputs, but that they are associated with a movement away from project-oriented development interventions to a much wider view of what development assistance should do and how it should be delivered (Conlin & Stirrat, 2008).

Moving away from stand-alone development projects, the modalities of development assistance has been varied. The sector wide approaches (SWAPs) and program-based approaches has been promoted among major aid agencies and beyond the sector support, there has been a move to the field of ‘general budget support’ which donors no longer take the leading position of development activities, but the partner countries have responsibilities and roles to accomplish agreed results.

Furthermore, the increasing recognition of the significance of trade, investment and other economic activities in development assistance has changed the landscape of development stakeholders. According to OECD (2017), ODA has shrunk in comparison to private financial flows and to meet Agenda 2030 more than 50 per cent of the financing could be mobilized through private sector. The development assistance now is only one of many ways in the development process and that trade and private investment are in quantifiable terms much more important in relationship between developed and developing countries. However, the challenges of data management system of partner countries still remained and many projects involving private sector partners also have weak monitoring systems or not to share primary objectives of development interventions.

In similar fashion, the scope of development objectives has widened to global issues such as climate change, gender, security, migration, inequality and sustainable development (OECD, 2005; OECD, 2019). After the MDGs, the Sustainable Development Goals are set with much wider sectors engaged with increased number of targets and a set of indicators to be measured.

Overall, the trends, emphasis and context of development industry has been more complex in terms of aid modalities, sectors, and stakeholders than it was 15 or even 10 years ago. When the RBM is promoted in 1990s, dominant form of development assistance was stand-alone projects or programs, but those shifts in the development industry made the RBM less useful for measuring and managing the development results (Conlin & Stirrat, 2008). One of major challenge that evaluation is facing is the issue of attribution. Increasingly development providers are asked to evaluate how the wider modalities, sectors and involvement of

stakeholders generate particular impacts or results on the development process. As far as the evaluation is concerned, reaching firm conclusion to attribution of development results to inputs and the chain of causation becomes more and more difficult if not impossible.

To meet these requests, many evaluation designs, methods, models and approaches have been developed and applied. In the 2000s, many widely-used impact evaluation designs, including randomized control trials (RCTs) and quasi-experimental designs were encouraged in development evaluation. Organizations such as the MIT Poverty Action laboratory and the working group of the Centre for Global Development see RCTs as the gold standard for impact evaluation and the standard which all should aspire to (Conlin & Stirrat, 2008). However, those impact evaluation designs fail to capture important unintended consequences of development (Bamberger et al., 2016). This form of evaluation is relatively rare in the development world (Rugh, 2006) because a matter of expense and other problems that they are only appropriate for highly selected situations with no practical or ethical problems in randomization and exclusion of a control group from development intervention. Such techniques work well only in particular context for instance, to assess effectiveness of drugs on health which has relatively clear and obvious distinction between variables. However, as discussed earlier, in the world of today, the range and factors that influence development trajectories are too complex to be calculated with in terms of simple cause and effect models. These experimental evaluation designs and the methods perhaps are encouraged and welcomed in condition where stand-alone projects or programs still dominate the landscape of the development, for example, the United States (Conlin & Stirrat,

2008). To conclude, in the complex world of program-based approaches, budgetary support, harmonization with partner countries, multi-sectoral approaches, there is less room for the effective use of impact evaluation or even the logical framework, but it is required for more flexible and less rigorous methods.

Third, over last 20 years of implementation, RBM has shown limitations and challenges in terms of methods as well as purpose of evaluation. Logical framework and results chain, the most utilized management and communication tools for development process since 1990s, are neither developed based on RBM nor share theoretical foundation with. Nevertheless, they have been one of the classic tools and still very much promoted by many development providers. For example, in every single project document and evaluation report produced by KOICA contain PDM (logframe) obligatorily regardless of modalities and types of aid: project, program, budgetary support, bi-lateral, multi-lateral, cross-sectoral and so on.

One of the biggest limitation that results-based approach in development evaluation could be its narrow focus on assessing stated objectives what Scriven (1991) said ‘desired intended effects’. Based on what Richards (1985) offered a classification of types of effects:

	Good (Desired)	Bad (Undesired)
Expected	Objectives achieved	Bad objectives achieved
Unexpected	Unexpected benefits	Unexpected harm

The doctrine of preordinate objectives invites us to focus on the upper left-

hand corner of the diagram [good expected effects], and even there it tempts us to oversimplify, because it calls for a focus on stated objectives, whereas real objectives are likely to be unstated. It is irresponsible because it judges the worth of a thing on the basis of an arbitrary subset of its effects. Robert Stake has affixed the label 'responsive evaluation' to studies that attribute no special importance to preordinate objectives, but instead estimate the value of the benefits a program has actually produced (Richards, 1985, p.32).

As Richards stress, division of expected effects (results) into stated and unstated implicitly and sometimes explicitly influence over evaluation process. Because the doctrine of results-based evaluation is to assess whether or not the stated objectives are achieved, the approach is likely to neglect unexpected and undesired results and thus inclines too much towards present data collection in evaluation (Gasper, 2000).

Thus, results-based approach may denounce the achievement of unintended or unexpected results, as well as the achievement of unstated objectives. And neglecting unforeseen routes and unintended results makes development providers to have narrow perspective on what really matters in partner countries. Some (Bamberger, 2000; Conlin & Stirrat, 2008); (Bamberger et al., 2016); GASPER, 1997) argue that when the results-based approach was adopted in development, most of important players and agencies were filled with people whom come from a quantitative, and often economic, tradition. As management-wise matter of efficiency as well as satisfying accountability issue in donor side, setting intended

results at the stage of program design and monitoring the progress towards the target seem effective and reasonable. Nevertheless, its provider-centeredness, what Conlin & Stirrat (2008) stated ‘supply-driven’, seem no longer suitable for current diversified situation of development industry.

Meanwhile, change of the subject of evaluation, so-called evaluand, in development over time has made it harder to answer problems in attribution. In 1990s, when the RBM was rapidly spread to agencies, stand-alone project dominated the development modality. Process of stand-lone project has relatively simple logic and program theory compared to others such as program-based sector wide approaches or general budgetary support. The movement away from projects as the dominant modality has led to a great deal of interest in how to evaluate development results to cope with complexities of SWAPs and budgetary support programs (Conlin & Stirrat, 2008).

CHAPTER V. Capability-based Evaluation Framework for Quality Education

So far, we have discussed what capability approach means from an educational point of view and why it is appropriate to be used in evaluating the quality of basic education. We have also looked at how the quality of education is evaluated and what are the important factors considered in the field of education. Capability approach has been used in a wide variety of academic fields, including education, development evaluation and studies of development cooperation. In terms of evaluation, capability approach contributed to expanding the standards for wellbeing and poverty in a more multidimensional sense from income-centered. This still has a great influence on many economists who measure the development, welfare, and happiness of the country. Before the use of HDI in 1990, it was only natural to evaluate the development of the country based on GDP. However, the definition of the development of a new approach of expanding substantial freedom based on Sen's capability suggested a groundbreaking evaluation paradigm in the field of development cooperation.

This chapter presents a new evaluation framework for the quality of basic education based on the core value of Capability approach as well as essential factors for the quality education that have been discussed for centuries in the field of education. Capability approach has been applied several times to the evaluation of stand alone development projects across various sectors, including education projects. However, as reviewed in the previous chapter, no research has been

conducted that suggests an evaluation framework at the field level beyond the evaluation at the unit project level.

Therefore, it is the first attempt to present an evaluation framework that combines a capability approach to education with a sector-wide approach together with its benefits, limits, and methodological concerns. In this regard, it is anticipated that follow-up research would improve the sophistication and applicability of the framework.

The framework suggested is for multidimensional measurement. The reason for this is that the link between the factors that affect the educational quality that is being evaluated is multidimensional. The variables that make up the quality of education are quite various, and the level is also multidimensional, as I learnt from the literatures on quality evaluation of education in the previous chapter. For example, Save the Children (2017) describes five foundations that support the wellbeing and learning and outlines sixteen components across those five foundations. And the foundations are underpinned by policies and systems that schools and communities greatly depended on.

The capability approach proposes a normative framework from an evaluation standpoint, whereas the quality of education framework provides field exclusivity and characteristics. The capability approach's normative proposition is that social arrangements should be evaluated largely on the degree of freedom people have to promote or fulfill functionings they value. Simply said, progress, development, and poverty reduction take place when individuals have more freedom. This is comparable to the ultimate goal of education, which is to realize holistic humankind.

Before presenting the framework, it is necessary to address one critical point: the capabilities approach's openness. This is because presenting the framework itself goes against some of the principles of openness. Sen (1992, 1993) argues that the Capability Approach is deliberately incomplete since it does not identify a list of valuable capabilities or functionings. Furthermore, he does not provide practitioners or researchers with clear practical guidelines on how to measure or identify capabilities (Comin 2001).

This incompleteness has been criticized by many researchers, including Sugden (1993), as a key weakness in having practical and operational importance of the capability approach. Sen responds to these critiques by claiming that 'an agreement on the usability of the Capability Approach - an agreement on the nature of the "space" of value-objects - need not presuppose an agreement on how the valuational exercise may be completed' (1993: 48). Sen (2005) also refuses to accept a predetermined set of capabilities, stating:

The problem is not with listing important capabilities, but with insisting on one pre determined canonical list of capabilities, chosen by theorists without any general social discussion or public reasoning. To have such a fixed list, emanating entirely from pure theory, is to deny the possibility of fruitful public participation on what should be included and why. (2005:158)

Nonetheless, the distinctive position of basic education in development studies provides an important rationale for proposing an evaluation framework based on the capabilities approach. Obviously, in many cases, those who benefit

from development are barred from participating in activities that are necessary for living a life that people value. In the case of basic education, this exclusion situation worsens to special conditions for young children. Although the basic education target is not solely determined by age, and the definition and scope differ from country to country, the target can be viewed as an age group corresponding to elementary and secondary education in general.

When the target of basic education is primarily focusing on children, it becomes difficult to associate with one's 'ability to choose' which underpins the whole process of the capability approach (Crocker, 2007) because at a younger age, the empirical basis for deciding what is valuable for oneself is more likely to be determined by the surrounding environment and structures. Ibrahim (2006), therefore, stresses the importance of the concept of collective agency which focuses on structural mechanisms and knowledge. Eyben (2004) identifies this kind of power on agency's ability to choose into five perceptions: power to, power over, power with, power as knowledge and power structure.

Furthermore, power influences people's ability to recognize what they value. Sen (1999) refers to this as 'adaptive preferences.' Such a concept is linked to Lukes' (1974) concept of 'false consciousness,' which prevents people from knowing their true interests. People need access to knowledge to develop a critical consciousness and overcome processes of dominance, according to the Freirean pedagogical tradition.

External factors that impede children's ability to choose for themselves can vary when incorporated into the field of basic education. For example, if the minimum conditions for lowering children's well-being are not met, such as

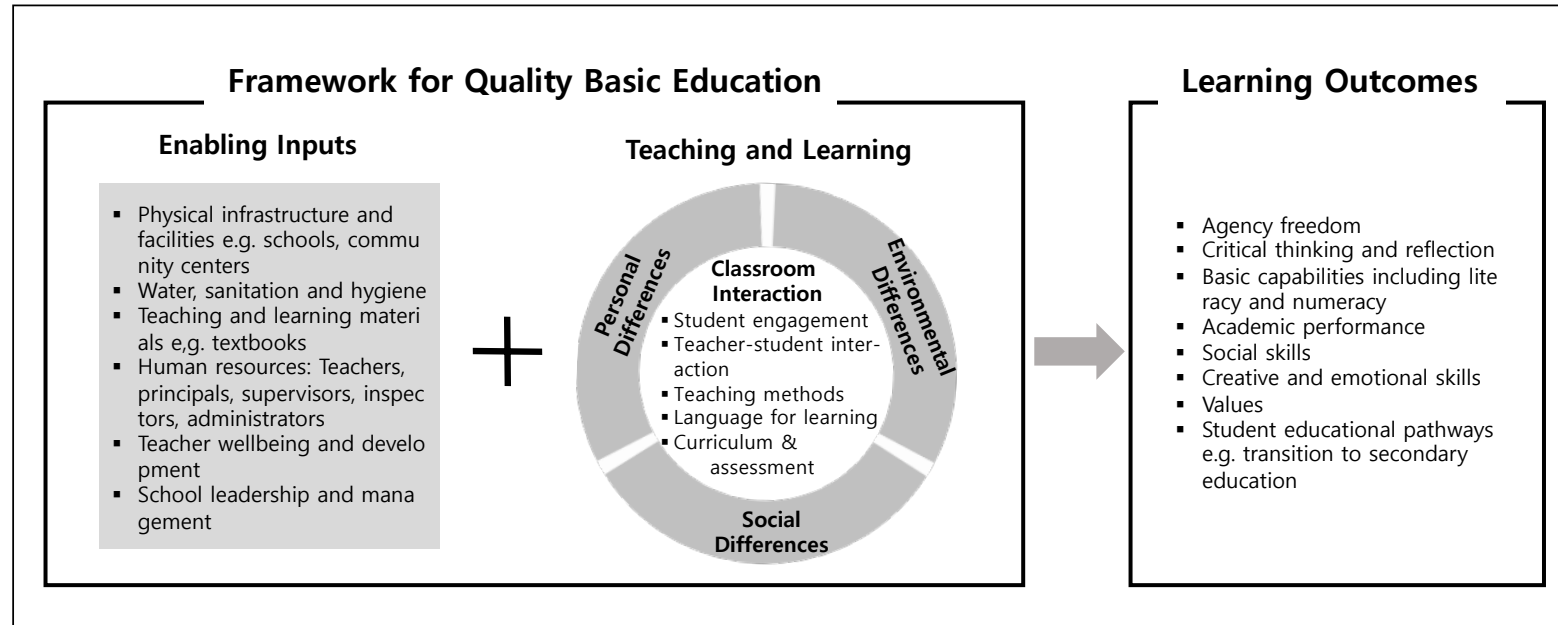
excessive housework, violence, extreme poverty, and unstable nutritional status in the home, it limits their ability to choose a life they value.

The capability approach stresses the importance of individual different ability to convert their resources and commodities into achieved functionings. These commodities could be tangible such as schools and learning materials or intangible such as curriculum and quality of teaching. The conversion of commodities to achieved functionings is affected by a series of factors which vary from person to person and context to context. Sen distinguishes the conversion factors into three different levels: personal, social and environmental and they are included in the following framework at the part of the process of teaching and learning.

5.1 Evaluation Framework for Basic Education Quality

Suggested evaluation framework as following consists of two parts. First part identifies what makes quality basic education, how it is assessed and what kinds of different conversion factors affect quality of basic education. Second part is the achievable results i.e., learning outcomes when quality education is provided.

Evaluation Framework for Basic Education Quality



Source: author

5.1.1 Enabling Inputs

In order to have quality basic education, some basic conditions are required to build an enabling environment for learning. According to OECD CRS, (this much) of development funds and resources are poured into build safe and accessible physical infrastructure such as schools with toilets, water taps and sanitation facilities (yearly or periodical stats provision). This type of education development approach had been accelerated during MDGs period and it is still one of the dominant ways of allocating resources in basic education sector.

As we moved on to the SDGs era, main agenda for education development also shifted from accessibility to quality issue. Influenced by this, the types and sub-sectors of aid has also been diversified. Compared to the 2000s, the average of last five years of aid allocation on teacher development has been (this much) increased (OECD CRS). Training teachers is a crucial and key contribution to improving the quality of education especially in situations where the quality gap among teachers is large, such as in many developing countries.

Furthermore, while interest in the field of development cooperation has recently increased, the emphasis on strengthening school leadership and management capacity is regarded as vital when discussing elements to improve educational quality in the field of education. Therefore, all those ‘enabling’ inputs including provision of safe and accessible learning spaces; water, sanitation and hygiene; adequate teaching and learning materials; quality teachers and school leadership and management become essential as well as fundamental factors that improve quality of education.

However, as mentioned several times in previous chapters, these inputs do not

immediately lead to improvement in learning outcomes. They must, of course, be essential prerequisites for improving learning outcomes. There would be no opposition to it at all. Nonetheless, despite extensive educational aid in development history, no discernible effects have been obtained in terms of learning outcomes.

5.1.2 Learning Outcomes

What are you referring to when you say, ‘learning outcomes?’ The definition of this has not been agreed and will not ever be. There is a significant gap in understanding between learning outcomes in education and learning outcomes in development cooperation. SDG 4 aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. Among 11 SDG 4 global indicators, the first and the most important indicator is about proportion of children achieving literacy and numeracy. In the field of education, it is commonly referred to as academic performance or examination performance. As such, learning outcome has often been evaluated as a result of a standardized as well as measurable form of test.

In other 10 global indicators, there are bodily and psychosocial well-being; accessibility to formal and non-formal education including ICT and global citizenship; adequately equipped schools; scholarships; teacher development and so on. According to what suggested framework indicates, most of those indicators are the matter of enabling inputs for quality education. Excluding these indicators, if only the indicators of learning outcomes are examined, it will be results of how

many children and adults are achieving a certain level of proficiency in literacy and numeracy skills.

Of course, standardized test scores must be a significant outcome indicator of quality of education. Basic education tests such as PISA and TIMSS have been used for a long time to identify gaps in education levels, not only in developing countries but also in many developed countries, and each government has developed its own education development strategy based on decades of analysis.

Learning outcomes could vary widely depending on the objective of learning. And the objectives are set differently depending on the level of education, context, and learners. Nonetheless, when it comes to categorizing learning outcomes, there is a common ground such as intellectual skills, cognitive skills, social skills, emotional skills, attitudes, values, social benefits, and such. Then, which of these do literacy and numeracy belong to? In general, they fall into the category of intellectual skills. In the field of education, Bloom's Taxonomy is frequently used to categorize intellectual skills. Benjamin Bloom, an educational psychologist, developed a model of intellectual skills in the 1950s that defined abilities like application, analysis, and synthesis as building on fundamental knowledge. Since then, a variety of models and approaches have been used by psychologists to better understand intellectual skills. However, some skills, such as problem solving, are widely used across a wide range of intellectual disciplines. Although literacy and numeracy are minor components of learning outcomes, being able to read and write and solve basic mathematical problems serve as foundational skills for further education as well as basic capabilities, they are one of the most significant achievements in the field of basic education.

Sen's capability approach centers on the concept of freedom, which he divides into two types: well-being freedom and agency freedom (Sen 1985). Brighouse and Unterhalter (2010) defined those freedoms as what an individual can achieve through education. These freedoms are related to the social circumstances that enable education to ensure instrumental, intrinsic, and positional values. In education, the field of well-being freedom is concerned with issues such as freedom from harassment in the classroom, freedom to focus in the classroom (not too sleepy, hungry, or nervous), freedom to access a lesson through suitable pedagogies, and excellent management quality (Unterhalter 2005). These freedoms are available to both adults and children. On the other hand, the ability to acquire information about education, engage in conversation, and make up one's choice regarding access to education for an adult without fear of violence or humiliation is referred to as agency freedom. In the case of children, the freedoms involve freedom from interference with their welfare rights as well as preservation of their ability to develop agency independence through school attendance (Brighouse 2002; Saito 2003). Therefore, if we consider the relationship between these two freedoms in education, well-being freedom can be seen as a freedom from restrictions for individuals to engage in the process of education, and agency freedom can be classified as achievements and outcomes that can be obtained as a result of education.

Until now, the question of enabling inputs for improving education quality, as well as the indicators and diverse categories in learning outcomes, has been discussed. In general, it can be said that in development cooperation, an implicit agreement has been reached on the importance and value of enabling inputs, also

known as 'first order' educational requirements (Riddell & Niño-Zarazúa, 2016). However, a variety of critical and less easily measurable factors, far beyond the 'first order' educational requirements, have a significant impact on educational outcomes such as the adequacy of the curriculum, effectiveness of teaching methods, the appropriateness of learning materials, physical closeness to schools, school leadership and management capacities, support and respect from parents and the community.

5.1.3 Teaching and Learning

For decades, astronomical amount of resources have been supported in the form of various enabling inputs for educational development cooperation. Nevertheless, the Preliminary Report of the Mid-Term Evaluation of the Education for All Fast Track Initiative (Cambridge Education et al., 2009) conclude that there is 'no robust evidence that FTI-endorsed countries, which engaged in funded educational programs, have systematically outperformed un-endorsed ones'. There are two important reasons why the enabling inputs appear to have little impact on achieving learning outcomes. First, it is necessary to understand the process of learning from an educational point of view, and the pluralism that an individual has become important variables influencing learning outcomes. Second, from an evaluation point of view, it is important to recognize that the attribution of those variables is typically multifaceted. A detailed discussion of each will be made in the next section, and in this section, we will look at the dynamic of the learning process and discuss the various conversion factors (Sen 1993) that

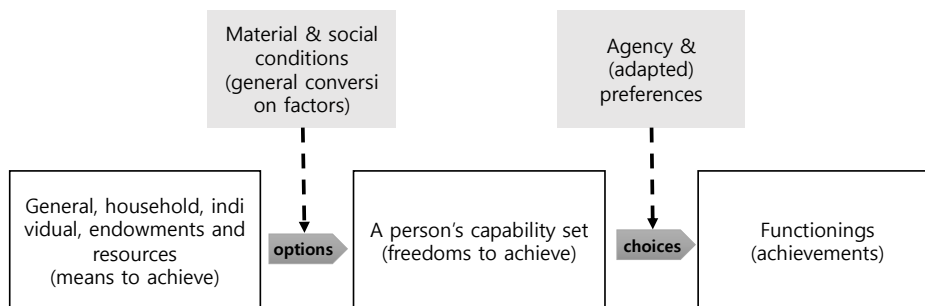
interact with the process.

Although basic education in this study encompasses various types of education in addition to schooling. However, in order to define the learning process that will be discussed in this section, I'd like to discuss the factors that are required for learning in a school setting. Many studies that describe how a learner embraces knowledge through learning activities and interactions with teachers, peers, and the system are worth mentioning. And UNESCO's (2005) framework for quality education, which is one of the world's leading organizations in the field of education, outlines four key elements that must be included in the teaching and learning process: learning time, teaching methods, assessment, feedback, incentives, and class size. Other prominent work (Rachel, 2017) proposes a framework for quality learning based on evidence and Save the Children's 100-year experience that includes important indicators for the teaching and learning process, such as teacher wellbeing and development, teaching and learning materials, learning language, pedagogical practices, planning, assessment, and reporting.

Some foreign aid to education aims to improve the factors mentioned above that are directly related to the learning process. Due to insufficient remuneration of teachers, the attendance rate of teachers in developing countries is only half that of developed countries, and in some cases, the language used by children at home and the language they learn and take exams in at school are different. Furthermore, some countries have been using the curriculum, which is a relic of the colonial era, for decades without revision. And no matter how enabling learning environment is in place, we will never achieve the desired learning outcomes if any of the

aforementioned issues arise during the learning process.

According to the capability approach, those factors directly influence over the learning process are constantly communicating with different conversion factors that learners have in individual, social and environmental levels. As Dejaeghere and Baxter (2017) describes in their work, various conversion factors are affecting a learner's capability set which means freedoms to achieve functionings. The extent to which resources and conditions have been or will be transformed into achieved functionings is determined by the availability of resources and the conversion factors of individuals.



Source: Dejaeghere and Baxter (2017, p.70)

Assuming that learning outcomes are achieved functions in the context of basic education, the personal, social, and environmental conversion factors mentioned by Sen (1992) affect the formation of the individual capability set that determines the choice for such learning outcomes. Although the sources of these factors may differ, all conversion factors influence how a child or learner can be or is free to convert the characteristics of resources into a functioning, being able to read or complete primary education. Internal factors such as sex, metabolism,

physical condition, reading skills, and intelligence are examples of personal conversion factors. If a person is disabled, in poor physical condition, or has never learned to ride a bike, the bike will be of limited assistance in enabling mobility. To put it in a basic education context, if a child is blind, the textbook will not be an effective tool for learning certain concepts.

Public policies, values and norms, practices that unfairly discriminate, societal hierarchical structures, or power structures related to, for example, class, gender, race, or caste are all examples of social conversion factors. Because it is often veiled, the social conversion factor among the three should be considered more carefully in the context of basic education. When a girl is raised by parents who believe that education is harmful to women and thus do not send their daughter to school, providing quality inputs to create an enabling learning environment such as building girls toilets in school has very little impact on getting the girl educated.

The physical environment in which a person lives produces environmental conversion factors. Climate, pollution, earthquake proneness, and the presence or absence of seas and oceans are all factors that influence one's geographic location. The stability of buildings, roads, and bridges, as well as modes of transportation and communication, are all aspects of the built environment. Take the bicycle as an example. The extent a bicycle contributes to a person's mobility is determined by their physical condition (a personal conversion factor), social norms (a social conversion factor), and the availability of decent roads or bike paths (an environmental conversion factor).

The three types of conversion factors all emphasize that knowing what

resources a person owns or can use is insufficient to assess the level of well-being he or she has achieved or could achieve; instead, we need much more information about the person and the contexts in which he or she lives. Sen uses the term "capability" to describe an opportunity made possible and constrained by both internal (personal) and external (social and environmental) conversion factors.

Based on literatures related to the application of the capability approach on education, following is suggested lists of each conversion factors to be considered as evaluative space in the context of basic education.

Personal differences	Social differences	Environmental differences
<ul style="list-style-type: none"> • physical health • emotional and mental health • cognitive skills • Social skills 	<ul style="list-style-type: none"> • power relations related to gender • power relations related to race • power relations related to class • parents support • community support • legal support 	<ul style="list-style-type: none"> • physical distance to learning places (e.g. school) • learning places at home • availability of decent roads, paths, transportation system to learning places (e.g. school)

Source: Author

The substantial freedom to be educated is thus depend on whether these requirements are met, and conversion factors, in this sense, also shape individual needs in question: in a society in which the capability of being educated is important (e.g., in order to vote), it becomes crucial to provide an adequate infrastructure and related policies (Crocker and Robeyns 2009, 68). Learners' personal, social, and environmental differences, as shown in the framework, influence the learning process, which leads to learning outcomes. Each learner has a unique set of conversion factors, some of which are metabolism-related, others

of which are shared with everyone in her community, and yet others of which are shared with people of the same gender, race, or social class.

Like the list of capabilities, conversion factors are not fixed. Above table thus provides the basic conversion factors that need to be evaluated and more factors can be added to them. The absence of a detailed discussion of each conversion factor is not due to a lack of importance in how the list is organized, but rather because the framework should be emphasized that the conversion factors should be included as 'evaluend' when evaluating the quality of basic education.

5.2 Discussions

According to OECD CRS data, the amount of aid to education sector is steadily increasing at a rapid pace. Between 2002 and 2019, total ODA to education increased by 295 percent in real terms, from US\$4.8 billion in 2002 to US\$ 14.3 billion in 2019. Total aid to basic education increased by 239 percent, secondary education by 502 percent, and post-secondary education by 293 percent during the same period. Although the amount of education ODA is steadily increasing, as discussed in Chapter 2, its impact is either negligible or not captured by any indicators other than the enrollment rate. In 2015, at the end of the MDG, organizations such as UNDP, UNESCO and EFA reports proudly highlighted the remarkable achievements in terms of school enrollment rate in primary education level as well as reduced gender gaps. What about completion of basic education, transition to upper level, teacher development, and learning outcomes?

No one can provide a clear answer to whether aid to education was not

effective enough to improve the quality of education or whether the results were not properly evaluated. I cannot provide a clear answer to that question either, but I can say that the evaluation approach that has been practiced in the field of education development cooperation is by no means an appropriate approach to evaluate the quality of education.

In Chapter 2, the issue was thoroughly discussed from both an education and an evaluation perspective. From an education standpoint, the existing evaluation approach had a very limited understanding of the process of improving educational quality and did not take into account education's unique comprehensive characteristics. Furthermore, the Results-based approach, which is currently the most widely used, is an evaluation approach that focuses on measuring achievement of target performance, and when learning outcomes are limited to specific test scores such as literacy and numeracy, it only provides partial information on the quality of education.

In previous section, a new evaluation framework for basic education quality is presented and explained its components. In this section, I will discuss what each of the proposed new evaluation frameworks means from the standpoints of education and evaluation, as well as how to overcome the limitations that have been identified thus far.

5.2.1 From Education Standpoint

There are three important reasons why the new framework is valuable in terms of education. First regardless of measurability, there certainly is the process of

learning and it is the crucial which determines the quality of education. Second, more attention is asked to evaluate various conversion factors which affects the effectiveness of the process of learning. Third, indicators of learning outcomes should be more diverse than literacy and numeracy because their coverage is not only partial but also sometimes irrelevant.

Measuring the unmeasurable: the process of education

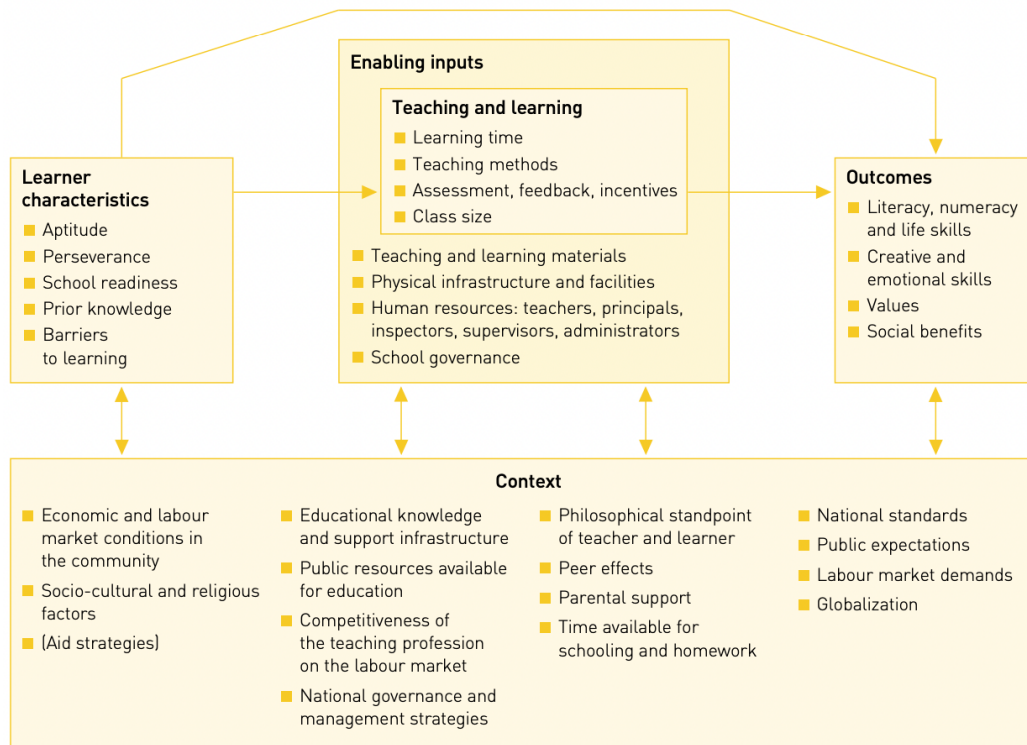
Quality has been declared to be ‘at the heart of education’ since 2000 Dakar framework and it still is in SDG era. It is considered as fundamental determinant of school enrolment as well as learning outcomes. While a decade ago, learning was mentioned in three of the six Jomtien goals, along with quality. Nonetheless, despite these early emphases, quality in the Global Monitoring Reports, and particularly quality in teaching and learning, has remained ambiguous. In chapter 2, we looked closely at the current SDG4 goal setting and indicator establishment process, and know that it is not a sudden result, but an extension of the EFA, which has been discussed in the field of education since 2000.

Following are the SDG monitoring framework global indicators. Some important studies criticized that first, those SDG indicators are formed to measure either inputs or outcomes, but these are not connected, and no indicators of processes are proposed (Unterhalter, 2019). Second, these indicators are proxies to highlight areas for change, but do not cover the structural or human processes needed to support quality education for more equitable manner (Alexander, 2015; K. King, 2017; Unterhalter, 2017).

	Indicators
4.1.1	Proportion of children and young people (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex
4.2.1	Proportion of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being, by sex
4.2.2	Participation rate in organized learning (one year before the official primary entry age), by sex
4.3.1	Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex
4.4.1	Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill
4.5.1	Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated
4.6.1	Percentage of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex
4.7.1	Extent to which (i) global citizenship education and (ii) education for sustainable development, including gender equality and human rights, are mainstreamed at all levels in: (a) national education policies, (b) curricula, (c) teacher education and (d) student assessment
4.a.1	Proportion of schools with access to: (a) electricity; (b) Internet for pedagogical purposes; (c) computers for pedagogical purposes; (d) adapted infrastructure and materials for students with disabilities; (e) basic drinking water; (f) single-sex basic sanitation facilities; and (g) basic handwashing facilities (as per the WASH indicator definitions)
4.b.1	Volume of official development assistance flows for scholarships by sector and type of study
4.c.1	Proportion of teachers in: (a) pre-primary education; (b) primary education; (c) lower secondary education; and (d) upper secondary education who have received at least the minimum organized teacher training (e.g., pedagogical training) pre-service or in-service required for teaching at the relevant level in a given country, by sex

Why do we need to capture the process of learning? And what does that mean? There have been two GMRs which explicitly highlights ‘quality’ issue including 2005 and 2014. In 2002 GMR, a table titled ‘an input-process-outcome framework for assessing education quality’ is proposed. The 2005 GMR also proposes a ‘framework for understanding education quality’. Those two quality frameworks are not very different from each other and both of frameworks emphasizes the

‘process’ of teaching and learning. The following framework is the one from 2005.



Source: UNESCO (2005, p.36)

As Alexander (2015) argues, too often the process of teaching learning has been remained secretly locked in a black box or timorously tiptoeing around the pool of pedagogy. After twenty years of wandering since Jomtien and Dakar, it's time to take the plunge and go one step further into the realm of reality, the 'process' of how we learn something and make changes.

The 2005 GMR explains how the teaching and learning process is closely interrelated with the enabling inputs and other contextual factors. And the most importantly, it puts the teaching and learning at the center of the quality framework to explain that it is the impact of curricula, effectiveness of teaching methods and

the motivations and participation of learners that finally make changes. In the report, it must be highly appreciated that the teaching and learning process is the most direct and core factor in improving the quality of education. Unfortunately, the elements presented by the framework have not been theoretically presented or explained to reflect the teaching and learning process. Proposed four elements of teaching and learning are learning time; teaching methods; assessment, feedback, incentives; and class size. As Alexander (2015) argues, those elements could be just random indicators that is deemed measurable.

Then, what should be assessed and measured and what approach should we take? Fundamentally, assessment for learning is the result of which effective teaching is made: the day-to-day, minute-by-minute observations and interactions through which good teachers constantly monitor children's learning and progress, providing feedback that builds on their understandings and probes and corrects their misunderstandings (Alexander, 2015). In other words, classroom interaction is the most significant and crucial aspect of teaching and learning process, so-called pedagogy, and therefore not looking at interactions and dynamics which occurs in the classroom is grave and unnecessary attempt to measure quality of education.

Measuring only what is measurable is in some way very irresponsible as well as dangerous. If we only measure the inputs and outcomes without the process, it may be the same as assuming a patient with tumor is treated well and cured after ticking the checkbox that medical equipment and doctor is deployed where the patient is hospitalized without asking following questions that monitor the process of surgery and treatment afterwards. That is maybe the key elements to improve

the quality of medical service. It is what we are observing with the list of SDG 4 indicators. We are not monitoring the essential part of the quality of education but rather we choose what is measurable but barely relevant.

In light of this, we need further discourses to set targets and indicators for both teaching and learning. Learning requires both a process and an outcome indicator and based on what we know about the critical conditions for learning, we might try ‘student engagement’. Similarly, if teaching must be reduced to just one indicator, we might try reciprocity in ‘teacher-student interaction’ based on what we know about the characteristics of effective teaching from both Hattie’s (2009) meta-analysis of studies in high-income countries and the 2013 DfID literature review. We should leave what is unmeasurable and develop them in their own terms as qualitative devices for making qualitative judgement and look for appropriate methodology to use them (Alexander, 2015).

Respecting pluralism

The framework does not only put the process of learning at the center of quality of education. It also provides necessary conversion factors that critically affects the learning process as well as the outcomes. As discussed in earlier chapters, pluralism and respecting learner diversity are significant characteristics of the capability approach. Sen (1992) argues that equalizing the ownership of resources and the enabling inputs “need not equalize the substantive freedoms enjoyed by different persons, since there can be significant variations in the conversion of resources and primary goods into freedoms” (p.33). Learners go

who through the process of learning differ in intersecting dimensions including personal, environmental, and social conversion factors.

Specific lists of those three dimensions of conversion factors are not fixed. It is the area of where evaluators and stakeholders might arrange through process of participation and dialogue which Sen consistently emphasizes (1992, 1999, 2004, 2002). The concept of those three different conversion factors are already discussed in previous section. Instead of explaining the concepts and a possible list of them in this section, I'd like to explain why they're important to assess in terms of basic education quality.

Individuals and learners are different in terms of personal differences as well as situated in different environmental and social settings. There is nothing wrong about being different. The problem is that difference or the intersection of differences is not inherently unequal; however, they can become inequalities (Terzi 2005). For example, a learner may value the capability to speak out and express freely but is silenced in a classroom due to specific social arrangements of power and privilege. To convert her capability into a functioning she requires social arrangements that are sensitive to her ways of expressing herself and provide her with opportunities to do so. This requires particular methods of teaching pedagogy and management and the resources for this that are including not only fixed assets such as staff, but also training, cultures of concern with learners' difference, and the capacity to put this care into practice (Walker & Unterhalter, 2007). These differences are amplified dramatically when a learner has a physical disability. Terzi (2005) explained that being a visually impaired learner is a disadvantage when specific learning materials such as Braille texts are

not provided. For learners who are not able to walk freely without help of special devices, appropriately designed physical environments are necessary to stay in school. Thus, it is important to have evaluative framework of enabling inputs as well as different dimensions of conversion factors that determines and influence the learning process to those who have disabilities and/or in basic education level who are not yet ready to make decision on what kind of life they should value and choose. By doing so, we could first get children out of basic conditional deprivation for a quality education and secondly, make a better evaluation.

In the capability approach, education is expected to be empowering and transformative. Sen's approach does not allow for the possibility that education in schools, colleges, and universities does not always function as the unqualified good that he considers it to be (Unterhalter 2003). However, poor-quality education, becomes a disadvantage that can last a lifetime. Our positive and negative school experiences will influence the decisions we make and how we navigate our futures. Curriculum, pedagogy, and assessment, as well as the school's culture, including whether or not all students are equally valued and respected, are examples of such experiences, as proposed by the framework.

As Sen leaves his framework deliberately vague and open, the framework that I proposed in this study does not provide complete set of capabilities, indicators, or variables. Agreeing with Sen's idea that to specify a single list of capabilities is to change the capability approach into the capability theory, a list of capabilities in education or any other area cannot simply be prespecified without public consultation, the main idea of the framework is to propose new *approach* of evaluating quality of basic education.

The specificity of education and learning, in relation to the problem of lists, raises two specific issues for the approach: first, the question of children, and second, the question of what counts as education and learning. How do we evaluate children's capabilities? Should children be allowed to make their own choices about whether or not to accept or reject education or certain elements of their schooling? Is a theoretical understanding of the "education good" required? How do we determine who in education lacks essential capabilities for school learning as well as post graduate opportunities and choices? While one might argue that all valuable capabilities matter at some abstract philosophical or theoretical level, as Robeyns (2003) points out, "this is no option for second-best theorizing or for applications." We must address aspects of the indexing problem and what capabilities matter in order to apply it in education.

Sen emphasizes the importance of schooling to nurture future capabilities when it comes to education and children (Saito 2003). Nussbaum agrees that children should be required to continue in compulsory education (schooling) until they have developed the capabilities necessary to allow them to make genuine and valued choices, such as leaving a traditional religious community. This highlights how capabilities of children cannot be evaluated without an understanding of how they relate to functionings. We might need to promote a relevant capability "by requiring the functioning that nourishes it," as Nussbaum points out (2000, p.91).

It is therefore reasonable to consider functionings rather than just their capabilities in the education of children and young people. Thus, we must understand whether and how capability is being developed, by whom and under what conditions, as well as how this relates to capabilities. And this is why the

proposed framework includes capability spaces to be evaluated in relation to the process of teaching and learning.

Education is not a neutral activity. As Brighouse and Swift (2003, p.367) point out, it always embodies a viewpoint on what is good in human life, or it might "seem vapid, even pointless.". However, is there any education capabilities to claim that it is objectively good for individual education development? We may not want to describe as education the process of generating inequalities or even tolerating, encouraging discrimination, prejudice, exclusion, alienation, or harassment of any student on the basis of difference. The capability approach would be incompatible with education that contributes to *unfreedoms*. As argues that "we need to be clear that respecting a plurality of conceptions of the good life is not the same as endorsing all versions of the good life, and this has clear educational implications".

Therefore, to count as education, the process and outcomes must enhance freedom, agency, and well-being by "enriching one's life with the opportunity of reflective choice" for a life of "genuine choices with serious options" (Sen 1992, p.41 And, in recognition of human and social diversity, in which different forms and roles of education may exist, the process of identifying education capabilities requires to have some form of participatory and inclusive dialogue, however conceptualized.

Nonetheless, we must determine whether we are focusing on developing capability only or also functioning, such as reading, writing, and critically evaluating information. Sen promotes the concept of capability, which allows us to set our own major life goals, and we should not prescribe for adults how they

should live. Capabilities are conceptual and hard to measure. For that reason, Robeyns (2005) and Walker & Unterhalter (2007) argue that we should probably evaluate functioning as a proxy for capability; however, we should do so without attempting to control students what they should do with their own lives and respecting a diversity of conceptions of the good life.

Re-scoping the learning outcomes

It is self-evident and empirically demonstrable that there is a close link between teacher quality and learning outcomes. But what and how should teachers teach? And, more importantly, on what aspects of their teaching should they focus and why? And how can we answer these questions if the nature of education has been misunderstood? Some of these questions are addressed in GMR 2014. It emphasizes diversity in the classroom, gender parity, and children with learning disabilities. It argues for compensating for teachers' lack of subject knowledge and the importance of classroom diagnosis and assessment tools, particularly for at-risk children. It finally enters the classroom in its seventh and final chapter, 'Curriculum and assessment strategies that improve learning.' It is, of course, a significant achievement that we open a new chapter to talk about what is going on inside of the classroom and begun to pay attention to the teaching and learning process, as we should have done since the 2000 Dakar framework for action. Both the curriculum and assessment discussions are useful within the parameters they set for themselves.

GMR 2014 goes beyond the traditional focus on literacy and numeracy to

argue for a broader curriculum and transferable skills. It does, however, adhere to the widely held belief, which dates back to the nineteenth century, that literacy and numeracy are and should always be the sole "basics" of education, regardless of time, place, culture, or national circumstance.

The case for literacy as a tool for individual empowerment and a lever for social and economic progress remains compelling, and successive GMRs have convincingly documented its impact in these terms. But, as heretical as it may seem to some, the case for continuing to give numeracy parity with literacy is neither proven nor entertained; the habit of history, it appears, is sufficient justification, and numeracy gets a free ride because 'literacy-and-numeracy' has effectively become a single curriculum component.

As discussed earlier, there are only few indicators among SDG 4 indicators which is relevant to what we consider the learning outcomes. Quality learning is undoubtedly one of the goals of those who framed Education 2030, particularly the SDG 4 on education and its targets, as we have seen. Quality is an aspiration not only for the overarching goal and, by extension, all of its component parts, but also for three of the key targets, which cover all levels of education with the exception of adult literacy and numeracy. The global indicators are not in the same boat. They do not reinforce the goal's and targets' focus on the quality dimension, but they frequently take a much narrower view of the target than its framers could have ever intended. Indeed, they 'minimalize' the goals of the various stakeholders who came up with this set of ten goals (K. King, 2017).

Minimum proficiency in reading and/or mathematics is considered as the most representative learning outcome global indicator. However, as King (2017)

argues, “the quantitative global indicator tail was wagging the rather quality-oriented SDG 4 dog”. For those countries, many of which are OECD members, but also several so-called emerging economies, where universal primary and secondary education has been in place for many years and where PISA competition is fierce, the process of reporting on global indicators of minimum math and reading proficiency may be almost meaningless. The sheer scale of SDG4 education targets is far beyond what is financially or politically possible for many developing countries, who are far from universal secondary education and whose formal TVET systems are minimal. The language of minimum proficiency may appear more relevant for their primary education systems, but in reality, pupils, parents, and teachers compete in their national selective testing systems for access to secondary school; thus, the narrative about minimum proficiency does not engage with the daily processes of education and the critical role of testing and examinations for them, either.

My argument is based on the assumption that education is more than just one thing, such as a learning outcome linked to test performance or the number of students enrolled in a particular school phase. Many aspects of education are difficult to quantify. Social, emotional, epistemological, normative, political, cultural, and economic relationships in education are difficult to quantify. As a result, many educational writers are critical of the political ways in which measurement has been used as a method of comparative analysis (Cowen 2014; Gorur 2014; Morris 2015; Auld and Morris 2016). Some methods, such as randomized control tests, reduce complex data to apparently simple causal relationships. As a number of historical studies show (Lawn 2013; Meyer and

Benavot 2013; Goldstein and Moss 2014), numerical data on education has come to play a key role in describing education systems and prescribing reform inspired by new public management. However, neither the historical work, which raises questions about data interpretation, nor the critical policy literature's questions have dispelled a desire to define education change as a science based on a set of facts. In education, unmeasurable processes are frequently addressed through the use of measurement or indicators.

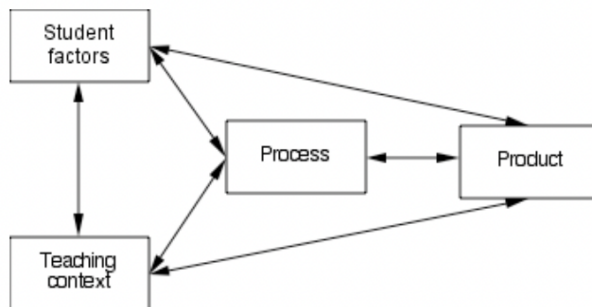
I'm not implying that assessing literacy and numeracy is unimportant. They are unquestionably important and serve as the foundation for a learner's capability to expand freedom. However, they are representing too little of what we can call learning outcome. For example, the SDG target 4.1 is about ensuring “that all girls and boys complete free, equitable, and quality primary and secondary education leading to relevant and effective learning outcomes” (United Nations 2015). In this target, there are important descriptions including free, equitable, quality, relevant and effective. However, the corresponding indicator for this target talks about the ‘proportion of children and young people’ at primary and secondary level “achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex” (United Nations 2015). Where did ‘free’, ‘equitable’, ‘relevant’ and ‘effective’ go? And the ‘quality’ is in lost translation into ‘minimum proficiency’ in just two subjects (K. King, 2017).

This happens not only in target 4.1 but all other SDG 4 targets translated into global indicators. As far as translating ‘quality’ into global indicator in SDG 4 is concerned, the term ‘quality’ is not mentioned in text of literacy and numeracy target that it only mentions “achieving at least a fixed level of proficiency in (a)

functional literacy and (b) numeracy skills” (United Nations, Economic and Social Council 2016). What are we missing out while translating the ambitions and aspirations of the Education SDG 4 goal and its 10 targets into indicators? King (2017) argues that there is an understandable tendency to define indicators in terms of readily available quantifiable data. This tendency is not only found in target 4.1 but also in narrowing down skills development into just ICT skills in target 4.4.

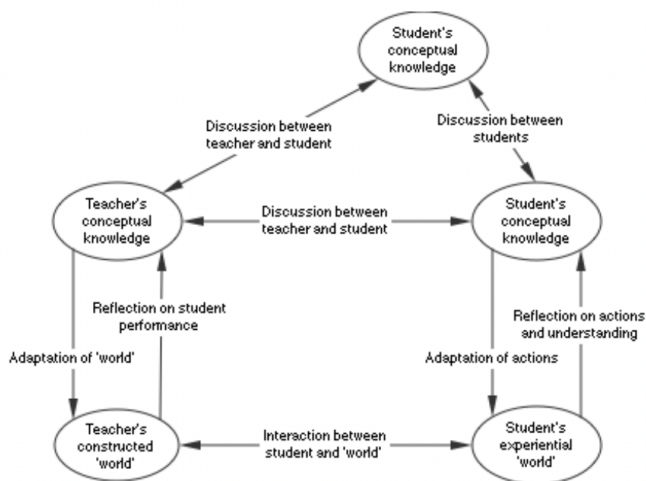
In education, the term ‘learning outcome’ is interchangeably used with ‘learning objectives.’ As it is closely linked to what is taught, learning objectives are the intended goal that learner can do after the process of learning, whereas the learning outcomes mean what learner can actually do. In general, learning outcomes refer to both discipline-specific knowledge and generic skills. The definition contains different dimensions of ‘knowing’ such as cognitive and conceptual understanding (Bloom, 1956); professional skills; learning skills; and societal beliefs and values (Krathwohl et al., 1973). According to Allan (1996), learning outcomes are distinguished into three types: subject-based outcomes; personal transferable outcomes; and generic academic outcomes. In most cases where the evaluation takes place in education setting, the subject-based outcomes represent the learning outcome.

Biggs (1989) proposed three components of learning with 3Ps: Presage, Process and Product. In this case the product represents the learning outcome. As following figure illustrates, 3Ps model shows two aspects of ‘student factors’ and ‘teaching context’ as the presage component.



Source: Biggs (1989)

Focusing on the learning outcome, Laurillard (2002) proposes conversational framework approaches which pragmatically distinguishes the four main aspects of learning and teaching: discussion, interaction, adaptation, and reflection. When Laurillard only counts discussion between the teacher and learner, Phillips & Luca (2000) extended the model to include discussions between students.



Source: Phillips & Luca (2000)

Those above-mentioned models and framework for learning outcomes are applied to all level of education from primary to tertiary and it gives much emphasis on subject-based cognitive knowledge achievement, in other words, the

academic performance. However, when talking of basic education quality, the meaning of learning outcome become broader as well as its scope. This is where I found reason why the capability approach has much to contribute to build what learning outcome means in basic education.

In capability approach, learning is assumed to be a fundamental means that an individual can improve her life (A. K. Sen, 1999). In both Sen's and Nussbaum's works, education is in itself a basic capability and at the same time it can be a means to develop and expand other capabilities (Walker & Unterhalter, 2007). What could be the possible learning outcomes of basic education? As discussed earlier, in terms of education point of view, that the learning outcomes are closely related to the objectives of teaching and learning. And the purpose of basic education can vary from society to society. In line with this, literacy and numeracy are foundational to the learning needs of all children regardless of a diverse range of context-dependent skills necessary for functional communication. In some cases, however, learning outcomes are more specific to the local context, and their application to individual circumstances is pragmatic (Young, 2009).

A study (Young, 2009) provides us a good framework of categorizing basic learning outcomes with basis on the capability approach. Based on data that she gathered, four generic basic learning outcomes are categorized: functional life skills learning, cognitive life skills learning, interpersonal life skills learning and personal life skills learning as agency freedom. She used qualitative methods to interview with 14-year-old children and their parents in Bhutan and Sri Lanka. The respondents provide their own individual and subjective interpretation of 'improvement' in the context of their lives and how learning is valued by the

individual to the extent that it leads to improvement in their lives.

Nonetheless, the study's findings are assisting us in gaining a better understanding of what can be included and excluded from the learning outcome of basic education when there is a deprivation of learning opportunity generically.

Dimensions	Examples
Functional life skills learning	Able to provide shelter, to contribute to the basic and essential daily routine of the household and to grow food for survival, as a baseline level of subsistence
Cognitive life skills learning	Able to achieve at least a minimum level of competency in literacy and numeracy knowledge and skills. Able to access information to learn about new technologies for improved living
Interpersonal life skills learning	Able to live together in harmony and able to participate in community development initiatives
Personal life skills learning as agency freedom	Able to make informed choices and to access information on rights to be able to challenge exploitative situations
Cross-cutting basic capabilities	The practical application of skills as positive learning outcomes Able to achieve financial security, able to remain healthy and able to put into practice basic parenting skills

Source: Young (2009, p.264)

The premise of the capability approach to identifying capabilities greatly value the process of democratic way of public discussion and dialogue and deliberately deny the exclusive list of capabilities. Valued learning has different characteristics in different populations and over time. Depending on their socioeconomic and cultural environments, different populations, even those living close together, may have very different interpretations of what is valued learning. Because of political and economic development, a population's perception of valued learning may change over time.

5.2.2 From Evaluation Standpoint

There are three major reasons why the new framework is valuable from an evaluation standpoint. First it opens room for various stakeholders to participate in evaluation process. Second, in line with the first reason, the framework calls for context reflected index for quality education. Finally, it calls for more discussion on the time frame issue, which is critical in capturing learning outcomes in education and the data collection period in evaluation.

Participatory Evaluation

This is time to be more flexible about evaluation approaches and methods in development evaluation. If the goal of evaluation is 'actually' accountability and learning (OECD), rather than comparing cost-benefit between projects, programs, and policies, it does not appear to be a problem to be more flexible and reflective in designing evaluation that is better suited to what is evaluated. Result-based evaluation and/or results-based management (RBM) is a dominant approach of use in development evaluation since 2005 (OECD, 2019). It is not distinctively acknowledged as an independent model or method in program evaluation; however, its use and the focus is very much similar to the goal-based evaluation; logic model based evaluation; and theory-based evaluation. Fitzpatrick (2011) categorized those three evaluation approaches under the results-based along with goal-free evaluation.

There does not exist one definition of RBM and its purpose; however, a recent

OECD study states that “the ultimate purpose of development co-operation is achieving development results (outcomes and tangible change). Development co-operation contributes to development results and results-based management supports this effort” (OECD/DAC 2017, p.8). As discussed in detail in Chapter 2, many organizations are not providing or defining a purpose for RBM in their evaluation, and it affects the implementation of their RBM reforms as well as reduces the learning potential of it (CGIAR 2017). To conclude, a number of donor agencies and organizations do not clearly communicate the RBM's purpose and how it contributes to development outcomes.

If the use of RBM is neither theoretically nor practically authorized in development evaluation, there is no reason not to use various evaluation approaches and models that have already been proven theoretically or methodologically through numerous studies. Therefore, herewith the new proposed evaluation framework suggests embracing a broader range of evaluation approaches, particularly participatory evaluation in education programs.

Participatory evaluation is one of many program evaluation approaches. It is one of twenty-three evaluation approaches and models in the work of Stufflebeam & Coryn (2014). J. Bradley Cousins is one of the most cited and influential theorists and his work has contributed greatly on contemporary participatory movement in evaluation. Numerous terms have been interchangeably used with participatory forms of evaluation, including participatory rural appraisal, participatory action research, community-based participatory research, and asset-based community development; collaborative evaluation and inclusive evaluation; empowerment evaluation; evaluation capacity building; and practical

participatory evaluation and transformative participatory evaluation (Stufflebeam & Coryn, 2014).

According to Patten (1997), the core principles of participatory evaluation are the following:

- Evaluation process involves participants' skills in goal setting, establishing priorities, selecting questions, analyzing data, and making decisions on the data.
- Participants own (commit to) the evaluation, as they make decisions and draw their own conclusions.
- Participants ensure that the evaluation focuses on methods and results they consider important.
- People work together, facilitating and promoting group unity.
- All aspects of the evaluation are understandable and meaningful to participants.
- Self-accountability is highly valued.
- Facilitators act as resources for learning; participants act as decision makers and evaluators.

As the term implies, participatory evaluation does not follow strict guidelines controlled by experts to data collectors. Instead, data collection guidelines are developed and refined through consensus, reflection, dialogue, and experience (Narayan & Mundial, 1996). It's gaining popularity in the development context, and it's frequently used in development projects, particularly community-based initiatives. People, agencies, and organizations with a stake in an issue, such as children, women, and men in communities, especially those from marginalized groups, are identified and then involved (Morra-Imas et al., 2009).

As the new evaluation framework for basic education quality requires identifying the values that communities pursue in educating children based on

public discussion and dialogue, as Sen emphasizes in the capability approach, involving key stakeholders such as children, teachers, parents, and community members in the evaluation process makes the framework's implementation easier and more sensible. In this regard, considering participatory evaluation approach as an alternative to the RBM seem much more reasonable.

There are both proponents and critics raising critical questions and concerns about participatory evaluation. Brisolara (1998) noted that:

One of the most frequent and serious charges leveled against participatory evaluation is that it violates a long-held evaluation principle (or tradition) by forsaking an objective-as-possible stance for what some see as an inevitable slide into the stance of relativism. (p.34)

Another common criticism of the approach is that it dilutes the technical quality of evaluation by putting many decisions about evaluation methods and procedures in the hands of "nonexperts," reducing external credibility significantly. It also entails risks and challenges that are not found in more traditional approaches, and its viability in practice has been brought into question (J. A. King, 1998; J. A. King et al., 2007).

Responding to those critics, Brisolara (1998) argued that "quality is maintained by evaluators through means similar to those adopted by their non-participatory colleagues – namely, evaluators remain responsible for ensuring the quality of methods and evaluation activities, and their roles as evaluation expert is central to their function" (p.36).

In participatory evaluation, evaluation planning decisions including identifying the questions, measures, and data collection strategies, are made together with participants that it is joint process rather than a traditional top-down process (Morra-Imas et al., 2009). It usually boosts the evaluation results' credibility among program staff, as well as the likelihood of them being used. Participatory evaluation promoters see it as a way to empower participants and build local capacity to participate in the development process.

Features	Participatory Evaluation	Traditional Evaluation
Focus and ownership	Participant	Donor
Purpose	Learning	Accountability and judgment
Design	Flexible	Predetermined
Methods	More informal	Formal
Role of evaluator	Outsiders as facilitators	Outsiders as evaluators

Source: Morra-Imas et al., (2009) p.194

As mentioned in the previous section, the teaching and learning process in the classroom is a critical factor that influences learning outcomes and educational quality. Teachers, learners, and parents are the primary actors in the teaching and learning process, but the group can also include community members, school leadership and/or management personnel, and policymakers as they play important role that forms different conversion factors. The main participants and beneficiary groups of each education program may vary depending on the expected outcomes; however, when the outcomes of the program are related to learning outcomes such as academic performance, it is clear that teachers and learners should be the main participants as well as beneficiaries, and they are invited to the evaluation process to provide key information on the program.

Including teachers, learners, parents, school leadership, officials and policy makers as participants as well as decision in evaluation makes great difference. As Gariba (1998) argues, the participatory evaluation play as a learning tool because it creates an opportunity for stakeholders to learn from their roles in the development intervention. It can also become a part of the development process as the evaluation activity is not separable from the development process itself. Additionally, participatory evaluation could build partnership and sharing responsibility among actors thus the stance of evaluators is transformed from an investigator to a promoter and participants.

As a result, when related stakeholders are deeply involved in the evaluation and have a voice in terms of setting targets, indicators, and data collection methods, the process can help to improve education quality because they have a better understanding of their roles in order to improve the results of the development intervention, including learning outcomes.

Second, the participatory evaluation approach appears to be more appropriate for evaluating basic education quality than the results-based approach because it naturally brings important values that are embedded in societies, culture, and context through the participation of key interest groups. As Sen argues that all the members of any collective or society “should be able to be active in the decisions regarding what to preserve and what to let go,” and emphasizes the importance of the process of public discussion by saying that “people need to be able to take part in these social decisions if they so choose” (1999, p.242).

As previously stated, the proposed framework does not include a comprehensive list of indicators that can be used to assess the quality of basic

education. Instead, it is attempting to secure some space for key stakeholders to participate in the development of the evaluation design, the selection of indicators, and the collection of relevant data throughout the evaluation. Thus, a list of educational capabilities, such as learning outcomes or the teaching and learning process, cannot simply be prespecified or defined without public consultation, because what capabilities and learning outcomes are valuable should be decided by those involved in the development and education process.

What are then the role of evaluators in this context? Because the direct beneficiary as well as key participants of basic education are mostly children and young people, it therefore makes sense to consider functionings (what we manage to achieve) and not just capabilities as indicators and one of important role of evaluators should be finding out the functionings that children need to achieve such as literacy and numeracy. It should, however, be done without prescribing to learners the choices they make about their own lives and while respecting a diversity of perspectives on what constitutes a good life.

Basic Education Quality Index

Another contribution of the proposed framework in terms of evaluation is that it sparks discussion about the need for a basic education quality index. Depends on the scope, scale, level and context of evaluands, its index should vary. It takes time and resources to develop a complex index; however, it is important to distinguish between technical hardship and less meaningful. Since 2000, many indexes have been developed which influenced by the capability approach to

measure things that were previously considered unmeasurable, such as wellbeing and equality. The Human Development Index (HDI), the Human Poverty Index (HPD), UNESCO's indicators of culture, DAC indicators for measuring poverty, the Gender-related Development Index (GDI), the Gender Empowerment Measure (GEM), the Inequality-adjusted Human Development Index (IHDI), the Swedish approach to welfare (Erikson, 1993), and the Dutch index of living conditions (Boelhouwer, 2002) are the examples.

It has been recognized that obtaining universal primary education on its own may not be sufficient until quality education is provided (UNESCO, 2005). While there is widespread agreement on the importance of providing a quality education to all children (UNESCO & UNICEF, 2012) there is limited agreement on what constitutes a quality education. According to Williams (2001), education quality is better understood in terms of output. While student achievement on tests and standardized examinations may be interpreted as a measure of educational quality by some, a more holistic view of education quality should consider the inputs, processes, outputs, and outcomes of education, which include student knowledge (academic and cultural heritage), social preparation (societal trends and needs), and personal development (personal and educational needs and interests) (Thijs & Van Den Akker, 2009).

Numerous researches on the quality of education are found, with substantial evidence pointing to low educational quality in the majority of educational systems (Benavot, 2011; Garira et al., 2019; Spaul, 2015). Nonetheless, there is a shortage of literature on how to improve the quality of education. However, the existing literature and research indicate that continuous monitoring and evaluation

of educational quality by schools through School Self-Evaluation (SSE), a process through which school communities learn about their conditions, processes, and outputs, is effective for school improvement (Carlson, 2009).

Studies including Scheerens's (2000) "Integrated model of school effectiveness," Howie & Plomp's (2001) "Factors related to Mathematics achievement" model, UNESCO's (2005) "Framework for understanding education quality", Griffith's (2008) "Proposed model for assessing quality of education,", Luong & Nieke's (2014) "Conceptualizing quality education from the paradigm of recognition" framework, Rachel's (2017) "Quality learning framework", and Garira's (2020) "Unified conceptual framework for quality of education in schools" provided various version of frameworks and models.

Those frameworks and models provide a wealth of information about how we have and should define education quality. However, there are insufficient evaluation studies and research that develop and use a quality education index for monitoring and evaluating various levels and forms of education in development context. Most of them focus on inputs, processes, and outputs of education without specifying educational levels and the context in which these levels operate.

The proposed framework for basic education quality in this study could be viewed as another framework that explains the constituents of quality education in some ways; however, it differs from others in that it focuses on its application for evaluation rather than conceptualization, narrow down the scope to basic education level and incorporates the capability approach as a theoretical foundation.

It has been more than two decades since the development of both the

capability approach and the EFA quality education goals. Numerous well-known studies are carried out in both areas, contributing theoretical validity and experimental success. Based on those literatures, this study suggested an evaluation framework for basic education quality in development context based on the capability perspective. Technical issues related to evaluation methods and data analysis will be discussed in the following section, with a focus on applying the capability approach to education quality evaluation.

Following table proposes an index for quality basic education based on the list of basic capabilities. The list of basic capabilities is based on reviewed literatures (Scheerens, 2000; Howie & Plomp, 2001; UNESCO, 2005; Griffith, 2008; Rachel, 2017; and Garira, 2020) and nine basic capabilities are categorized by three levels: individual, local, and structural. The list contains key components of both enabling inputs and the teaching and learning process of the framework. Selecting valid indicators as well as methods to collect and analyze should be flexible to evaluators and participants.

Basic capabilities and Index for Quality Basic Education

Informational Space	Basic Capabilities	Specification	Suggestions for potential composition of indicators
A. Individual level	1. Bodily health	Good house, adequate food, personal hygiene, access to healthcare, physically strong and plump	Child development index
	2. Emotional and Psychosocial integrity	Positive and respectful interactions	Survey results from parents, teachers and peer
	3. Parents and community	Learning at home and community; Child, parent and community participation	Attendance, completion rate, transition rate
B. Local level	4. Basic learning environment	Safe and accessible learning spaces	Water, sanitation, and shelter
	5. Teacher capability to provide adequate pedagogical practices	Teaching and learning materials; teachers being able to access in-service training; teachers wellbeing and development	Pupil/teacher ratios, number of classes per day, number of training hours per year, Teacher salary relative to cost of living
	7. School leadership and management	School management, school leadership	Resources spending per pupil
C. Structural level	8. Relevant curriculum	Curriculum relevance to national examination; reflection of local language, culture, and social values	Language of instruction, language of examination, Survey and interview results from teachers
	9. Capability of securing basic education for every child	Free and compulsory basic education law, policy, and financial support; inclusive and protective policies	Resources spending per pupil/GDP, budget allocation to education, free/compulsory basic education law, year of schooling

Source: Author

Time-frame issue

Most development programs in context of ODA have their own timeframe. Based on the timeframe, achievable outputs and outcomes of program is designed at earlier stage of a program. In reality, it is common to design projects that adhere to the project's timeframe and budget constraints. Most development cooperation projects have shorter durations than national development plans, and input resources are limited. Regardless of the shorter timeframe and limited resources, there is a requirement to demonstrate the effectiveness of the intervention through evaluation in development context. The establishment of development goals in the international community, such as MDG and SDG, adds pressure to meet targets on time, forcing development activities to focus on specific indicators and quantitative figures.

Some fields like education requires relatively longer timeframe to achieve certain outcome and if the outcome has to be measurable, not observable, it may take more time. Education indicators that are currently used in development cooperation are the same indicators which are collected and analysed at the national level in most countries. The enrollment rate, graduation rate, and academic achievement indicators are among them. They are close to process indicators for tracking and monitoring, not for analyzing the correlation or causality to specific interventions or inputs. Of course, in countries where the primary school enrollment rate is still low, schools are built, budgets are drawn up, and necessary policies and laws are enacted to increase the enrollment rate. In addition, when the enrollment rate reaches a certain level, other interventions are

planned to improve other educational indicators, such as graduation rates and transition rate to the next level of education.

However, in a development context, results must be measured and achieved within a specific period of time, regardless of the fact that each indicator has a different time requirement. In results-based management and evaluation, it's more important for inputs, activities, outputs, and outcomes to be logically related, which can sometimes be proven through experimental design. Many education indicators that are commonly used in development fields are not easily achievable within few years of projects or programs. Moreover, even though some indicators are achieved during certain timeframe, it is not possible to find what intervention attributed the change.

There is much more to consider about the time factor when the academic achievement indicator is used as an indicator of the project's outcome level. This aspect of time is critical when it comes to educational evaluation planning. First, the time it takes to achieve varies depending on the scope and level of academic achievement. For example, the amount of time it takes elementary school students to grasp the concept of addition versus college students to grasp the concept of calculus should be set differently.

Second, conversion factors have an impact on learning timeframes. Even though they are taught by the same teacher and use the same learning materials, children in the same classroom learn at different rates, have different levels of understanding, and have different test scores. Individual, social, and environmental differences are important factors to consider when determining an appropriate learning timeframe, according to the proposed framework.

Third, learning is a complex process. In most cases, it couldn't be as simple as a matter of knowing certain knowledge or not. Complex and diverse knowledge is frequently required in a comprehensive manner, regardless of how simple the specific concept of teaching is. Assume there is a health-care project that vaccinates people and monitors the production of antibodies as a result. In this case, the results can be measured in a dichotomy based on whether or not antibodies are produced. If, on the other hand, an education project's goal is to improve mathematical performance, key concepts should be chosen by teachers or related experts based on the curriculum taught by teachers for a year or semester in the third grade in order to create a test to measure the indicator.

Academic achievement indicators vary greatly in difficulty within the same field of education. When teaching welding technology for a week in a technical school, for example, the tools, materials, processes, and techniques required for welding can be segmented to assess knowledge and attitude based on the curriculum. However, compared to testing welding techniques, the information needed to measure critical thinking, social skills, and communication skills, is very broad and methodically challenging.

As a result, because the content and purpose of basic education is relatively normative the evaluation approach and method of the field should be different. This does not mean that academic achievement is not important in basic education. Literacy and numeracy are very important basic capabilities that these indicators need to be measured and accumulated steadily. However, there are more learning outcomes depending on curriculum and each of them has their own timeframe for learners to acquire. Additionally, the time it takes for a child to learn to read and

write varies depending on a variety of factors such as teaching methods, teaching and learning materials, learning time, instruction language, and so on.

5.3 Application and Methodology

5.3.1 Famous Case Studies of S. Alkire

‘Social impact assessment’ and participatory social assessments are two widely used methodologies that aim to identify human impacts and use this information to shape public and development activities to beneficiaries' values and institutions. As I've stated several times throughout the study, my argument is not about 'replacing' or overlooking impact assessment's contribution; rather, it's about identifying what information is missing and how it can be supplemented through participatory assessment, such as the capability approach.

In this section, the work of Sabina Alkire (2002), a well-known capability scholar and economist, is examined, with a particular focus on research methodologies. By contrasting economic cost-benefit analysis of three projects in Pakistan with capability analysis, it suggests a tool for improving the evaluation of participatory projects. Her research shows how different results can be derived from different evaluation approaches. She evaluated three OXFAM projects conducted in Pakistan including loans for goat, literacy community development and rose cultivation. Her research team first performed a cost-benefit analysis (CBA) by measuring and analyzing all tangible costs and benefits, then comparing the results to an alternative method of evaluation that included the key intangible impacts described by beneficiaries.

According to Alkire, the literacy project is a prime example that, if evaluated only by traditional criteria, would no longer be funded. An analysis based on criteria such as the total number of graduates, the unit cost per graduate, the income generated by the income generation activities, the projected future socio-economic benefits would conclude that this project is too expensive, too risky, and too weak institutionally to merit funding. She states that:

The literacy project is a prime example of the value of assessing 'intangible' human impacts of activities. For in light of these, it become entirely clear that the project has had a fundamental and transformative impacts on the women students. Had an assessment of the project only reported on economic, or even 'literacy' impacts, it would have omitted the most significant impact. (p.256)

To conduct CBA, her team calculated the costs of the project including wages, utilities, stationery, training, rent, travel, and miscellaneous expenses related to the administration of the literacy classes. The costs are converted to constant prices using the consumer price index. To calculate the economic benefits that literacy graduates gain through income generation activities, the study adopted previous research by Burki and Ubaidullah (1996) that the estimated returns per year of schooling to be 5 per cent. Then she estimate the mean annual income of a woman graduate.

$$\frac{\text{Mean annual income of literate woman} - \text{Mean annual income before literacy class}}{15 \text{ months} \times \text{income loss/month} + \text{cost of study}} \Rightarrow 5\%$$

Earning function of illiterate women: (w)
 Potential earning function of literate women: x(w)
 Social constraints on female employment: S
 Earning function of literate women in presence of s (w)
 Earning function of literate women without x x(w)

She estimated three ranges of likely future wages for graduates based on the survey and presented each of the internal rates of return (IRR). As the result, the likeliest best-case situation (Colum A) among seven estimations shows the negative internal rate of return i.e., it fails to prove its effectiveness based on the social impact assessment which is the most widely used.

Assumptions	A*			B		C	
Additional income and time	5 yrs	10 yrs	20 yrs	3 yrs	5 yrs	1 yr	2 yrs
	Rs 500	Rs 500	Rs 500	Rs 1,000	Rs 1,000	Rs 2,400	Rs 2,400
IRRs							
High shadow wages	-10%	8%	14%	5%	23%	-4%	71%
Med. Shadow wages	-6%	12%	17%	14%	36%	16%	93%
Low shadow wages	-5%	12%	17%	16%	37%	20%	95%

*Colum A is most likely.
 Source: Alkire (2002) p.263

For capability analysis, she first made a list of significant additional benefits and costs of the literacy classes. It included values for its social returns, not only the economic returns, such as the effect on child's health and attitudes towards

children's education, or on contraception prevalence. For identifying and ranking the impacts, the participatory impact assessment exercises were conducted twice: first with the graduates and second with currently enrolled students.

The research provides four and half pages of interview descriptions based on the impact categories as other empirical studies provide numbers with statistical formula. The interview description plays an important role because it contains diverse thoughts, values, conversion factors of each participant and social contexts. Finally, the evaluation is concluded with the rankings of impacts by different groups including coordinator, graduates, current students, and evaluators. The ranking is graded by using a scale of 0 to 5 that 5 means intense impact; 0 means no impact.

According to the ranking results, the fact that one or the other was consistently cited as the literacy project's most powerful impact can be interpreted as an overall sign of 'empowerment.' Empowerment was expressed by participants' own language that: "women are equal to men; women do not need to suffer abuse; women can decide what is good and bad; women can solve own problems; hope to develop own judgement; learn to interact with strangers".

Alkire stresses that ranking results reflects the participants' perceived impacts. However, it does not give a good idea of how deep in fact these impacts were compared to the impacts of similar projects in other areas. Therefore, in order to evaluate the comparable depth of very different impacts, qualitative ranking across projects along the 'dimensions' of impacts to which participants responds is required. The study provides the qualitative ranking of impacts of three cases and the starred numbers indicate the three most significant impacts as identified by the

participants and the evaluators.

Category	Goats Project	Literacy Project	Roses Project
Life/health/security	4*	3-4	1
Knowledge	3-4*	5*	5*
Work/play	3-4	1-5	0-5
Relationships	4-5	4-5	5*
Beauty/environment	2	0-1	5*
Self-integration/inner peace	4	4*	5*
Religion	3-4	2-3	3
Empowerment	3-5*	5*	3

Source: Alkire (2002) p.282

Comparing the ranking results to the cost-benefit analysis of three projects provide us an important insight. If Oxfam had done a careful cost-benefit analysis of each project, the information available for comparison would have been as shown in the table below.

	Goats Project	Literacy Project	Roses Project
Total Oxfam grant	63,400	506,329	16,764
Number of direct beneficiaries	140	66	10
Number of years in operation	5	3	2
Plausible IRR without distributional weights	20%	-6%	-52%
Total estimated income per beneficiary from project	6,058	190 (?)	3,630
Annual estimated income per beneficiary	1,102	190 (?)	2,286
Socio-economic status of women	Poor	Poor	Poor
Market viability	yes	no	Possibly

Source: Alkire (2002) p.280

The goat-rearing project is likely to be funded if Oxfam has extra funds to support and is only looking for cost-effective projects. What if Oxfam's budget is cut and the organization decides to withdraw resources? The literacy project is

likely to be considered because it offers little in terms of income and has a negative IRR. If Oxfam were only interested in making social investments that addressed obvious market failures, it might shift funds away from goat rearing and rose cultivation and toward female literacy. The rationale for doing so could be that the literacy project's 'output' was a 'good' in and of itself, and that it could also be a way to reduce the social constraints that prevented women from participating in the labor market, potentially leading to private economic returns in the long run.

How does the qualitative ranking of impacts information complement the cost-benefit analysis? It draws attention to aspects of that project which might be improved. For example, the literacy project had the greatest impact on empowerment and knowledge of the three projects, but failed to generate significant income. This highlights the importance of including empowerment impacts in its evaluation, as the cost-benefit analysis excludes these benefits entirely. Furthermore, the information assistance comparisons are more subtle and significant. More information will not result in a unique optimum; the decision cannot be made on technical grounds, but rather on moral grounds. Nonetheless, the identification and comparative valuation of impacts by multiple agents has clarified the implications of various courses of action. Lastly she argues that while the choice remains underdetermined, the identification of valuable capabilities did bring to light more complete information about each of the three projects than alternative evaluation methodologies might have.

5.3.2 Application methods

Just because it is not measurable does not mean it is not important. Difficulty in measuring is a problem that must be solved methodologically and should not be dismissed. Reviewing literatures that used the capability approach, the selected functionings and capabilities vary not only by individual preferences but also by being influenced by the country's and region's environment, institutions, and systems, and their weight may also differ. As a result, in order to assess educational development cooperation interventions, it is necessary to first develop or choose measurable indicators and analysis methods by defining the learning outcome to be achieved through the project/program. If this is not possible due to technical or practical constraints, an alternative method for identifying the problems that have the greatest impact on this learning outcome and establishing them as project/outcome program's problems can be considered.

In this section, we look at how the capability approach has been used in the field of evaluation over the last 20 years, as well as the research methods used to identify and assess functionings and capabilities. In addition, we will discuss the unresolved issues and limitations. To use the capability framework, data collection for the selected capabilities must be completed, which can take a significant amount of time and resources. However, given that impact evaluation using RCT was once carried out and actively encouraged, I do not believe it is an insurmountable challenge.

Clark (2017) stresses by saying that “It seems to be important to note that empirical research techniques are not always required for all applications of the capability approach. Some applications rely on analytical reasoning or critical analysis, such as when using the capability approach to critique an existing social

practice or when drawing on empirical findings already available.” However, many applications of the capability approach rely on new empirical analysis, necessitating the application of empirical research methods (Brandolini & D’Alessio, 1998; Kuklys, 2005; Qizilbash & Clark, 2005; Ramos & Silber, 2005).

In terms of methodology, provide categorization of the evaluation and analysis methods proposed within the capability literature. There are fifteen methods divided into four general methodological categories: ad hoc, foundational, procedural, and mixed methods.

General methodology	Description	Specific methods	Selected authors
Ad hoc methods	Selection based on <i>ad hoc</i> considerations Characteristics: Provisional; contextual; pragmatic or intuitive justification	Purpose of study Researcher’s own values/normative views Pragmatic concerns and practical issues	Human Development Index (UNDP, 2015), Sen (1985a, b), Slotje (1991), Klasen (2000)
Foundational methods	Normatively relevant capabilities and functionings are derived from normative values and principles Characteristics: Absolute; universal; theoretical justification	Deductive reasoning Overlapping consensus Technical knowledge	Internalist essentialism Dialectically necessary judgments Human rights theory Others
			Early Nussbaum (1990; 1992), Qizilbash (1998) Claassen and Düwell (2013) Later Nussbaum (2011b) Vizard (2006, 2007, 2010), Fukuda-Parr (2011), Osmani (2005), Sengupta (2002)

Procedural methods	Normatively relevant capabilities and functionings are contingent on the social, cultural, and political context Characteristics: Democratic; open-ended; contextual; empirical or discursive justification	Deliberative democracy		Anderson (1999, 2003), Crocker (2005, 2008), Drydyk (2005), Sen (1999a, b, 2001, 2002), Sen and Scanlon (2004), Drèze and Sen (2002)
		Qualitative analysis	Narrative approaches	Hodgett and Deneulin (2009), Biggeri et al. (2006), Phelps (2006), Clark (2002, 2005)
			Participatory approaches	Biggeri et al. (2006), Clark (2009), Frediani (2007), Doyal and Gough (1991)
		Quantitative analysis	Data-driven selection	Schokkaert and Ootgem (1990), Lelli (2001)
Mixed (or multi-stage) methods	Integrates two or more technocratic and procedural methods Characteristics: Holistic; dialectical; conflicting; both theoretical and empirical/discursive justification	Comparative methods		Brock (2009), Burchardt and Vizard (2011), Vizard (2010), Holland (2014a,b)
		Synthesizing methods		Alkire (2002), Claassen (2011), Khader (2011)

Source: Comim et al. (2018) p.202

Ad hoc methods are characterized by a lack of methodological justification. These methods are not based on any normative theories, but rather on ad hoc considerations such as the purpose of the study, the researcher's own values, or practical and pragmatic concerns.

According to foundational methods, a list of capabilities and functions is derived from some ultimate normative values or principles that grounds the selection, such as human agency, human dignity, or human rights. As a result, foundational lists frequently claim to be universally applicable, even though the lists may be further adjusted to suit local context.

Though procedural methods, a list of capabilities is derived from an open-ended empirical or deliberative discourse. As such, these methods rely on the data of the selection on people's reasoned subjective preferences. Thus, a list of

capabilities is contingent on the social, cultural, and political context and is justified by empirical findings and/or political and deliberative procedure.

According to Comim et al. (2018) the mixed or multi-stage methods combine one or more foundational methods with one or more procedural methods. As a result, mixed methods seek to integrate various sources of knowledge (e.g., normative theory, human rights theory, qualitative and quantitative analysis, democratic deliberations) in order to provide a dialectical and holistic approach.

CHAPTER VI. CONCLUSION

6.1 Evaluation of Education Quality and Capability

Certainly, the global education agenda such as EFA, MDGs and SDGs along with enormous amount of foreign aid to education made more children to have opportunity for education. However, according to UNESCO Institute for Statistics (UIS), the official source of cross-nationally comparable data on education, currently 617 million children still cannot read and do basic math, less than 40% of girls in sub-Saharan Africa complete lower secondary school and 262 million or 18% of all children and youth are out of school.

From MDG 2 to SDG 4, one of the major shift of global education agenda is its focus on the quality. Quality has been always at the center of discussion in education development. Apart from school entry and completion, quality agenda seeks to ensure that all children achieving minimum proficiency level of literacy and numeracy. Progress has been made since 2015. There are many assessment tools, and they make negative prospect to reach the target by 2030.

The results of the Latin American Laboratory for Assessment of the Quality of Education (LLECE) suggest that nearly 60% of grade 3 reach minimum proficiency on average and this proportion will stay the same. The Progress in International Reading Literacy Study (PIRLS) assesses the reading skills of grade 4 students in mostly high-income countries and 81% of them achieve at least minimum proficiency level. The results of the Programme d'Analyse des

Systemes Educatifs de la CONFEMEN (Analysis Programme of CONFEMEN Education Systems, PASEC) show that 42% of grade 6 students in assessed Francophone African countries achieve the minimum proficiency, but the trend is negative, and if continues, the proportion could drop by nearly 30%. The Programme for International Student Assessment (PISA) indicates that two-thirds of 15-year-old students in middle- and high-income countries perform at the minimum level. Different types of assessments are showing slightly different results in different regions; however, all of them prospects that at the current rate of progress, the result will be more or less the same or even worsen in 2030.

There can be many ways to explain why the quality indicator lags behind the target. Depends on context and perspectives, responses and solution to this problem will be different. How we interpret the phenomenon can affect future decisions, and a scholar's responsibility in a society is to inform the world of the perspective what she believes is correct. In that regard, I tried to present a slightly different viewpoint than what was previously available. In the context of development cooperation, the method and approach for evaluating the quality of education should be deviated from donor-centered evaluations that overlook the process aspect of education. This is because such an approach not only fails to accurately assess the actual outcomes of education, but also fails to provide useful policy implications or lessons.

In order to propose a new evaluation framework for assessing quality with specific emphasis on basic education, first I examined the dominantly used evaluation approach in development evaluation and show critical limitations that needs to be supplemented in order to capture qualitative aspects of education

process and learning outcomes. Amartya Sen's Capability Approach has been applied various ways to measure wellbeing and poverty in development context. It defines development as substantial freedom and considers education as core capability to be achieved as an end and as an instrument to expand other capabilities as a mean. Its core concepts such as functionings, capabilities, conversion factors and agency are adopted in this study to explain the essence of education and to determine what to evaluate.

What is the better way to evaluate the quality of education? Can evaluation be more educational? I suggest an answer with the capability approach and propose an evaluation framework which embeds the core value of the capability approach to measure quality of basic education. Though the target of international education agenda has changed from quantity issue of accessibility to quality matter, mostly the learning outcome, the way of evaluation has not been changed in terms of target and indicator setting, evaluation approach and method to collect data.

From the perspective of evaluation, the capability approach opens up space for discussion on who decides the purpose and objective of development intervention and what method to take to collect reliable data to measure the results. Some argues that development evaluation practice has been donor-centered that the objective of intervention, target, and indicator are decided by the donor countries. It caused alienation of local demands and context affects the ownership, effectiveness, and sustainability of development intervention.

In program evaluation, there are many approaches and models (Stufflebeam & Coryn, 2014). What is called development evaluation has its root in program evaluation and the most encouraged and widely promoted approach is called the

Results-based evaluation or results-based management approach. It shares many things in common with the theory-based evaluation, logical approach, goal-based evaluation (Fitzpatrick, 2011). According to OECD (2017), the twelve DAC members have all made efforts to improve results-based management and progressively move to measure outcomes instead of outputs. The evaluation of results-based approach normally provides a flow chart or matrix that shows relationship between inputs, activities, outputs, outcomes, and impact through theory of change model or logical framework. However, according to Vähämäki (2017), a common finding in RBM research is that different objectives can conflict with one another. If organizations do not define what type of results information they want to collect, for example, they may end up collecting a lot of data in vain and overburdening partner organizations with requests. Meanwhile, decision-makers may receive information they cannot use due to a lack of understanding of why, when, and who requires results information for decision-making. As a result, it's critical to understand what types of information are appropriate for various purposes and audiences at various levels. According to the OECD, results-based management approaches benefit from a defined objective and commitment that is aligned with the agency profile (size, modes of operation, and so on) and strategy (OECD, 2017).

One of the problems with the results-based framework is that as it requires results information that is easy to aggregate, organizations naturally give more attention to the outputs or short-term outcome data rather than changes or outcomes which take more time to observe and collect (OECD, 2019). OECD called this consequence as 'measure fixation' which means the tendency to focus

on what can be measured easily when designing evaluation. In practice, RBM tends to prioritize what can be easily measured, preferring quantitative data over qualitative evidence (even though the latter can better reflect the reality of program implementation) and short-term goals over long-term outcomes.

In contrast to what RBM focuses, quantitative data and short-term outcomes, what we expect from education is not easily measurable because the process has to be proven by qualitative evidence and it sometimes takes longer time than we expected at first place. Therefore, when education program or projects are evaluated based on RBM, what is easily measurable in shorter timeframe is very likely to be assessed. In education we express as ‘black box’ to indicate the difficulty of capturing the process of teaching and learning in classroom. The correlation between number of schools and enrollment rate is probably stronger than the correlation between school enrollment and literacy rate. There is a certain gap which is mostly invisible and unmeasurable between resources and the learning outcome. And the black box is normally indicating the ‘teaching and learning process’ in education.

Whatever learning outcome is defined, it is not easily achieved in short-term framework. In development context, finding out objectively calculatable impact and correlation between intervention and results is one of important mission of the evaluation. This tendency invited experimental research such as RCT into development evaluation to assess social changes based on monetary values and statistics. The limitations of RCT have been already discussed in detail in this study. In summary, given the increased time and resource requirements, the results of RCT as one of the approaches of impact are not as useful as they should be in

measuring a variety of social changes and non-monetary values. And for several reasons, including 'general equilibrium effects,' results from experimental research are difficult, if not impossible, to replicate successfully across an entire system (UNESCO-UIS, 2019). Additionally, its strict control of experimental design leads the exclusion of critical conversion factors that may affect the actual outcomes.

Therefore, if an evaluation concludes that an intervention is effective at improving learning outcomes without examining the teaching and learning process critically, it is very likely to be classified as one of 'pseudo' evaluations (Stufflebeam et al., 2000). As King (2017) notes, the 'quantitative indicator tail wagging the rather quality-oriented intervention dog', the contents and approaches of development projects or programs can be purposefully designed to be 'easy to measure' regardless of their quality dimension.

6.2 Limitation of the Study

Before concluding present study, I would like to review challenges and unsolved problems of the proposed evaluation framework which incorporates the capability approach. In a way, this study may cause some discontent to some readers. First, some might expect profound philosophical analysis on the capability approach. This study focuses on how the capability approach can be applied to develop an alternative framework of evaluation in education development cooperation. Therefore, core theoretical arguments made by Sen and other prominent scholars are included in the literature review part; however, the

majority of the literatures examined in this study are empirical research rather than philosophical or theoretical.

Second, further evaluation research which applied the proposed evaluation framework are requested to examine its applicability in development evaluation practice. This study is not going further to present how the framework is applied in real world evaluation. Conducting evaluations requires not only a reasonable amount of time and resources, but also special permissions to access aggregate data created by both donor and partner agencies, as well as a diverse group of participants. Recognizing this limitation and attempting to address it, this study included a chapter that examined the work of Alkire (2005). By presenting the evaluation results of three development projects, she demonstrates how capability-based evaluation can replace traditional social cost-benefit analysis.

Furthermore, it could be argued that the proposed evaluation approach overlooks practical issues too much in favor of focusing on capability. It is a valid criticism; however, there are enough evaluation practices, methods and their scholarly articles that are heavily affected by the results-based approach, so in this study I would like to stress capability elements while acknowledging the benefits and contribution of the results-based in development evaluation.

I look forward that this study will further develop through opinion and reviews from experts who has more knowledge in the fields of evaluation, methodologies, indicators and even education.

6.3 Final Remarks

The EFA has been in place for more than two decades. Many children have received a school education as a result of the international community's efforts, which is a true accomplishment. However, now that we've reached the point where we're talking about quality, we need to start having discussions about how to 'evaluate' quality. The danger of evaluating quality through the lens of the existing quantitative approach is that it is very partial, biased, and can even lead to incorrect results has been discussed throughout the study. In development cooperation, approaches and methods of evaluation has been evolving. It has been discovered that experimental research cannot provide sufficient answers to how to improve the quality of education because the once widely used impact evaluation has serious limitations in its application in specific fields, including education.

Two types of studies would further strengthen the knowledge base which can be useful for evaluating quality of education. One is more frequent and rigorous both empirical and qualitative data analysis based on functionings and capabilities. So far only few studies which applied the capability approach was able to collect and analyze data that are initially designed for the purpose of the capability evaluation. Rest are using data that is available. The more studies we have that compare capabilities in various contexts, the more implications and methodologically sound approaches will be drawn. For example, each level of education may have their unique list of basic capabilities, as this study suggested for the basic education, and it will differ from communities, regions, and countries.

Secondly, there are many studies out there to prove which country and specific

intervention made great change and how fast they achieved (UNESCO-UIS, 2019) in numeric terms. However, not many have asked how the process design and implementation made difference in learning outcomes in development context. This study does not argue that qualitative evaluation should replace the current dominant trend of quantitative measurement; rather, it raises concerns what important factors have been missing in current evaluation framework when evaluating education quality and what are the alternative ways to include quality-oriented interventions in education, such as teaching and learning, can be assessed. To give an answer to the question, this study proposed an evaluation framework for basic education quality. The framework distinguished between enabling inputs and the process of learning which it always has been the most decisive yet has been easily neglected in development evaluation. And the conversion factors introduced by Sen which affect the process of learning are also included as one of important evaluation subjects.

Lastly, debates on whether to take education as end of development or mean to expand other capabilities will continue. In many articles, Sen focuses more on instrumental value of education how can girl's education be valuable to health, or good health to productivity, political participation, and social practices and so on. And some education scholars, on the other hand, argue great detail on the intrinsic value of education how education be the end of development. In this study, this argument was not considered as central issue because developing evaluation framework is more like a work of building a structure of house but not of deciding what to put on the top of it. Also, in a way that this study already limited its scope to basic education, whatever the learning outcome is will take the end of

educational development practice.

BIBLIOGRAPHY

Alexander, R. J. (2015). Teaching and learning for all? The quality imperative revisited. *International Journal of Educational Development*, 40(1), 250–258. <https://doi.org/10.1016/j.ijedudev.2014.11.012>

Alkire, S. (2005). *Valuing freedoms: Sen's capability approach and poverty reduction*. Oxford University Press on Demand.

Allan, J. (1996). Learning outcomes in higher education. *Studies in Higher Education*, 21(1), 93–108.

Allen, R. L. (2021). Critical race hermeneutics: A theoretical method for researching the unconscious of white supremacy in education. In *The Handbook of Critical Theoretical Research Methods in Education* (pp. 15–30). Routledge.

Andresen, S., Otto, H.-U., & Ziegler, H. (2010). Bildung as Human Development: An educational view on the Capabilities Approach. In H.-U. Otto & H. Ziegler (Eds.), *Capabilities – Handlungsbefähigung und Verwirklichungschancen in der Erziehungswissenschaft* (pp. 165–197). VS Verlag für Sozialwissenschaften. http://link.springer.com/chapter/10.1007/978-3-531-91909-6_10

Bamberger, M. (2000). The Evaluation of International Development Programs: A View from the Front. *American Journal of Evaluation*, 21(1), 95. <https://doi.org/10.1177/109821400002100108>

Bamberger, M., Tarsilla, M., & Hesse-Biber, S. (2016). Why so many “rigorous” evaluations fail to identify unintended consequences of development programs: How mixed methods can contribute. *Evaluation and Program Planning*, 55, 155–162. <https://doi.org/10.1016/j.evalprogplan.2016.01.001>

Barrett, A. (2009). *The Education Millennium Development Goal Beyond 2015: Prospects for Quality and Learners*.

Bell, S. (2000). Logical frameworks, Aristotle and soft systems: A note on the

origins, values and uses of logical frameworks, in reply to Gasper. *Public Administration and Development*, 20(1), 29–31.

Biggs, J. B. (1989). Approaches to the Enhancement of Tertiary Teaching. *Higher Education Research and Development*, 8(1), 7–25. <https://doi.org/10.1080/0729436890080102>

Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York: McKay*, 20(24), 1.

Brandolini, A., & D'Alessio, G. (1998). Measuring well-being in the functioning space. *General Conference of The International Association for Research in Income and Wealth, Cracow, Poland*.

Brisolara, S. (1998). The history of participatory evaluation and current debates in the field. *New Directions for Evaluation*, 1998(80), 25–41.

Butin, D. W. (2010). *The education dissertation: A guide for practitioner scholars / Dan W. Butin*. Corwin.

Camfield, L., & Tafere, Y. (2011). 'Good for Children'? *Local Understandings versus Universal Prescriptions: Evidence from Three Ethiopian Communities* (pp. 200–221). https://doi.org/10.1057/9780230308374_9

Carlson, B. (2009). School self evaluation and the critical friend perspective. *Educational Research and Reviews*, 4(3), 078–085.

Clark, D. A. (2005). *The Capability Approach: Its Development, Critiques and Recent Advances*. Global Poverty Research Group.

Clark, D. A. (2017). Valuing and revaluing education: What can we learn about measurement from the South African poor? *Comparative Education*, 53(1), 54–80. <https://doi.org/10.1080/03050068.2017.1254954>

Comim, F., Fennell, S., & Anand, P. B. (2018). *New frontiers of the capability*

approach. Cambridge University Press.

Comim, F., Qizilbash, M., & Alkire, S. (2008). *The Capability Approach: Concepts, Measures and Applications*. Cambridge University Press.

Conlin, S., & Stirrat, R. L. (2008). Current Challenges in Development Evaluation. *Evaluation, 14*(2), 193–208. <https://doi.org/10.1177/1356389007087539>

Drèze, J., & Sen, A. (2002). *India: Development and Participation*. Oxford University Press.

Easterly, W. (2006). *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So*. So.

Entwistle, V. A., & Watt, I. S. (2013a). A capabilities approach to person-centered care: Response to open peer commentaries on “Treating patients as persons: a capabilities approach to support delivery of person-centered care.” *The American Journal of Bioethics, 13*(8), W1–W4.

Entwistle, V. A., & Watt, I. S. (2013b). Treating patients as persons: A capabilities approach to support delivery of person-centered care. *The American Journal of Bioethics, 13*(8), 29–39.

Fitzpatrick, J. L. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Pearson Education.

Fukuda-Parr, S. (2003). The human development paradigm: Operationalizing Sen's ideas on capabilities. *Feminist Economics, 9*(2–3), 301–317.

Fukuda-Parr, S. (2011). Theory and Policy in International Development: Human Development and Capability Approach and the Millennium Development Goals. *International Studies Review, 13*(1), 122–132. <https://doi.org/10.1111/j.1468-2486.2010.01003.x>

Fukuda-Parr, S., & Cid-Martinez, I. (2019). Capability approach and human

development. In *The Palgrave Handbook of Development Economics* (pp. 441–468). Springer.

Garira, E. (2020). A Proposed Unified Conceptual Framework for Quality of Education in Schools. *SAGE Open*, 10(1), 2158244019899445. <https://doi.org/10.1177/2158244019899445>

GASPER, D. (1997). “*Logical Frameworks*”: A critical assessment managerial theory, pluralistic practice.

Gaspar, D. (2000). Evaluating the “logical framework approach” towards learning-oriented development evaluation—ProQuest. *Public Administration and Development*, 20(1), 17–28.

Gore, C. (1997). Irreducibly social goods and the informational basis of Amartya Sen’s capability approach. *Journal of International Development*, 9(2), 235–250. [https://doi.org/10.1002/\(SICI\)1099-1328\(199703\)9:2<235::AID-JID436>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1099-1328(199703)9:2<235::AID-JID436>3.0.CO;2-J)

Griffith, S. A. (2008). A proposed model for assessing quality of education. *International Review of Education*, 54(1), 99–112.

Gutwald, R., Leßmann, O., Masson, T., & Rauschmayer, F. (2014). A Capability Approach to Intergenerational Justice? Examining the Potential of Amartya Sen’s Ethics with Regard to Intergenerational Issues. *Journal of Human Development and Capabilities*, 15(4), 355–368.

Hart, C. S. (2012a). *Aspirations, education and social justice: Applying Sen and Bourdieu*. A&C Black.

Hart, C. S. (2012b). The capability approach and education. *Cambridge Journal of Education*, 42(3), 275–282.

Hatakka, M., & Lagsten, J. (2012). The capability approach as a tool for development evaluation – analyzing students’ use of internet resources.

Information Technology for Development, 18(1), 23–41.
<https://doi.org/10.1080/02681102.2011.617722>

Hinchcliffe, G., & Terzi, L. (2009). Introduction to the Special Issue “Capabilities and Education.” *Studies in Philosophy and Education*, 28(5), 387–390.
<http://dx.doi.org/10.1007/s11217-009-9133-7>

Holland, B. (2008a). Ecology and the limits of justice: Establishing capability ceilings in Nussbaum’s capabilities approach. *Journal of Human Development*, 9(3), 401–425.

Holland, B. (2008b). Justice and the environment in Nussbaum’s “Capabilities Approach” why sustainable ecological capacity is a meta-capability. *Political Research Quarterly*, 61(2), 319–332.

Holland, B. (2014). Nussbaum, Rawls, and the Ecological Limits of Justice: Using Capability Ceilings to Resolve Capability Conflicts. *Capabilities, Gender, Equality: Towards Fundamental Entitlements*, 382.

Howie, S. J., & Plomp, T. (2001). *English Language Proficiency and Other Factors Influencing Mathematics Achievement at Junior Secondary Level in South Africa*. <https://eric.ed.gov/?id=ED455250>

Kaufman, A. (2006). Capabilities and Freedom. *Journal of Political Philosophy*, 14(3), 289–300. <https://doi.org/10.1111/j.1467-9760.2006.00239.x>

King, J. A. (1998). Making sense of participatory evaluation practice. *New Directions for Evaluation*, 1998(80), 57–67.

King, J. A., Cousins, J. B., & Whitmore, E. (2007). Making sense of participatory evaluation: Framing participatory evaluation. *New Directions for Evaluation*, 2007(114), 83–105.

King, K. (2017). Lost in translation? The challenge of translating the global education goal and targets into global indicators. *Compare: A Journal of*

Comparative and International Education, 47(6), 801–817.
<https://doi.org/10.1080/03057925.2017.1339263>

Klees, S. J., & Thapliyal, N. (2007). The Right to Education: The Work of Katarina Tomasevski. *Comparative Education Review*, 51(4), 497–510.
<https://doi.org/10.1086/520863>

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1973). Taxonomy of educational objectives, the classification of educational goals. Handbook II: Affective domain. David McKay Co. Inc., New York, 1, 956.

Kronlid, D. (2014). *Climate change adaptation and human capabilities: Justice and ethics in research and policy*. Springer.

Kuhn, T. S. (1970). *The structure of scientific revolutions* ([2d ed., enl]). University of Chicago Press.

Kuklys, W. (2005). *Amartya Sen's capability approach: Theoretical insights and empirical applications*.

Kusek, J. Z., & Rist, R. C. (2004). *Ten Steps to a Results-based Monitoring and Evaluation System: A Handbook for Development Practitioners*. World Bank Publications.

Laurillard, D. (2002). *Rethinking university teaching: A conversational framework for the effective use of learning technologies*. Routledge.

Luong, M. P., & Nieke, W. (2014). Conceptualizing quality education from the paradigm of recognition. *Journal of Education and Practice*, 5(18), 178–191.

Mair, S., Jones, A., Ward, J., Christie, I., Druckman, A., & Lyon, F. (2017). A Critical Review of the Role of Indicators in Implementing the Sustainable Development Goals. In *World Sustainability Series*. https://doi.org/10.1007/978-3-319-63007-6_3

Matias, C. E. (2021). *The handbook of critical theoretical research methods in education / edited by Cheryl E. Matias*. Routledge, Taylor & Francis Group.

Millar, M. R. (2013). A capabilities perspective on healthcare associated infection. *The American Journal of Bioethics*, 13(8), 53–54.

Morra-Imas, L. G., Morra, L. G., & Rist, R. C. (2009). *The Road to Results: Designing and Conducting Effective Development Evaluations*. World Bank Publications.

Moyo, D. (2009). *Dead aid: Why aid is not working and how there is a better way for Africa*. Macmillan.

Narayan, D., & Mundial, B. (1996). *Toward participatory research*. World Bank Washington, DC.

Nielsen, L. (2015). Why health matters to justice: A capability theory perspective. *Ethical Theory and Moral Practice*, 18(2), 403–415.

Nussbaum, M. (2003). CAPABILITIES AS FUNDAMENTAL ENTITLEMENTS: SEN AND SOCIAL JUSTICE. *Feminist Economics*, 9(2–3). <https://www.tandfonline.com/doi/abs/10.1080/1354570022000077926>

OECD. (2002). *Glossary of Key Terms in Evaluation and Results Based Management*. OECD.

OECD. (2006). *Emerging good practice in managing for development results* [Source Book]. OECD.

OECD. (2019). *Learning from Results-Based Management evaluations and reviews* (OECD Development Co-Operation Working Papers No. 53; OECD Development Co-Operation Working Papers, Vol. 53). <https://doi.org/10.1787/3fda0081-en>

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The Measurement of*

Meaning. University of Illinois Press.

Otto, H.-U., & Ziegler, H. (2006). Capabilities and education. *Social Work & Society*, 4(2), 269–287.

Phillips, R., & Luca, J. (2000). Issues involved in developing a project based online unit which enhances teamwork and collaboration. *Australasian Journal of Educational Technology*, 16(2).

Picciotto, R. (2018). Accountability and learning in development evaluation: A commentary on Lauren Kogen's thesis: *Evaluation*. <https://doi.org/10.1177/1356389018781361>

Qizilbash, M., & Clark, D. A. (2005). The capability approach and fuzzy poverty measures: An application to the South African context. *Social Indicators Research*, 74(1), 103–139.

Rachel, M. (2017). *The Quality Learning Framework*. Save the Children. https://resourcecentre.savethechildren.net/node/12460/pdf/00_quality_learning_framework_september_2017.pdf

Riddell, A., & Niño-Zarazúa, M. (2016). The effectiveness of foreign aid to education: What can be learned? *International Journal of Educational Development*, 48, 23–36. <https://doi.org/10.1016/j.ijedudev.2015.11.013>

Robeyns, I. (2005). Selecting Capabilities for Quality of Life Measurement. *Social Indicators Research*, 74(1), 191–215.

Robeyns, I. (2006). The Capability Approach in Practice. *Journal of Political Philosophy*, 14(3), 351–376. <https://doi.org/10.1111/j.1467-9760.2006.00263.x>

Ruger, J. P. (2004). Health and social justice. *The Lancet*, 364(9439), 1075–1080.

Sachs, J. (2014). The case for aid. *Foreign Policy*, 21, 2014.

- Saito, M. (2003). Amartya Sen's Capability Approach to Education: A Critical Exploration. *Journal of Philosophy of Education*, 37(1), 17–33. <https://doi.org/10.1111/1467-9752.3701002>
- Scheerens, J. (2000). *Improving school effectiveness*. Unesco, International Institute for Educational Planning.
- Schlosberg, D. (2012). Climate justice and capabilities: A framework for adaptation policy. *Ethics & International Affairs*, 26(4), 445–461.
- Schlosberg, D., & Carruthers, D. (2010). Indigenous struggles, environmental justice, and community capabilities. *Global Environmental Politics*, 10(4), 12–35.
- Scriven, M., & Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation*, 57–70.
- Sen, A. (1985). Well-Being, Agency and Freedom: The Dewey Lectures 1984. *The Journal of Philosophy*, 82(4), 169–221. <https://doi.org/10.2307/2026184>
- Sen, A. (1992). *Inequality reexamined*. Oxford University Press.
- Sen, A. (1993). Capability and well-being. *The Quality of Life*, 30, 270–293.
- Sen, A. (2003). *Development as Capability Expansion*. New Delhi: Oxford University Press.
- Sen, A. (2006). The Man without a Plan-Can Foreign Aid Work. *Foreign Aff.*, 85, 171.
- Sen, A. (1999). *Development as freedom / Amartya Sen*. Knopf : Anchor Books.
- Smith, N. L. (1993). Improving Evaluation Theory Through The Empirical Study of Evaluation Practice. *Evaluation Practice*, 14(3), 237–242. <https://doi.org/10.1177/109821409301400302>

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation Theory, Models, and Applications*. John Wiley & Sons.

Stufflebeam, D. L., Madaus, G. F., Scriven, M., & Kellaghan, T. (2000). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Springer Science & Business Media.

Terzi, L. (2007). Capability and Educational Equality: The Just Distribution of Resources to Students with Disabilities and Special Educational Needs. *Journal of Philosophy of Education*, 41(4), 757–773. <https://doi.org/10.1111/j.1467-9752.2007.00589.x>

Thijs, A., & Van Den Akker, J. (2009). *Curriculum in development*. Netherlands Institute for Curriculum Development (SLO).

Tikly, L. (2017). The Future of Education for All as a Global Regime of Educational Governance: El futuro de la educación para todos es un régimen global de gobierno educativo. *L'avenir de l'Éducation Pour Tous Comme Régime Mondial de Gouvernance Éducative.*, 61(1), 22–000. <https://doi.org/10.1086/689700>

Tikly, L., & Barrett, A. M. (2011). Social justice, capabilities and the quality of education in low income countries. *International Journal of Educational Development*, 31(1), 3–14. <https://doi.org/10.1016/j.ijedudev.2010.06.001>

Tomaševski, K. (2001). *Free and compulsory education for all children: The gap between promise and performance*. Gothenburg.

Ul Haq, M. (2003). The birth of the human development index. *Readings in Human Development*, 127–137.

- UN. (2015). *The Millennium Development Goals Report 2015*. UN.
- UNESCO. (2005). *UNESCO's Conceptualization of Quality: A Framework for Understanding, Monitoring and Improving Education Quality*.
- UNESCO. (2017). *Unpacking Sustainable Development Goal 4: Education 2030; guide* (p. 32).
- UNESCO & UNICEF. (2012). *End of Decade Note on Education for All: Goal 2*. <https://www.yumpu.com/en/document/view/23145927/end-of-decade-note-on-education-for-all-goal-2-pdf-unicef>
- UNESCO-UIS. (2019). *How fast can levels of proficiency improve?* (No. 62; Information Paper).
- Unterhalter, E. (2003). Education, capabilities and social justice. *Chapter Prepared for UNESCO EFA Monitoring Report, 4*.
- Unterhalter, E. (2009a). Social Justice, Development Theory and the Question of Education. In R. Cowen & A. M. Kazamias (Eds.), *International Handbook of Comparative Education* (pp. 781–800). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6403-6_50
- Unterhalter, E. (2009b). What is equity in education? Reflections from the capability approach. *Studies in Philosophy and Education, 28*(5), 415–424.
- Unterhalter, E. (2014). Measuring Education for the Millennium Development Goals: Reflections on Targets, Indicators, and a Post-2015 Framework. *Journal of Human Development and Capabilities, 15*(2–3), 176–187. <https://doi.org/10.1080/19452829.2014.880673>
- Unterhalter, E. (2017). Negative capability? Measuring the unmeasurable in education. *Comparative Education, 53*(1), 1–16. <https://doi.org/10.1080/03050068.2017.1254945>

Unterhalter, E. (2019). The Many Meanings of Quality Education: Politics of Targets and Indicators in SDG4. *Global Policy*, 10(S1), 39–51. <https://doi.org/10.1111/1758-5899.12591>

Venkatapuram, S. (2013). *Health justice: An argument from the capabilities approach*. John Wiley & Sons.

Walker, M. (2005). Amartya Sen's capability approach and education. *Educational Action Research*, 13(1), 103–110. <https://doi.org/10.1080/09650790500200279>

Walker, M. (2008). A human capabilities framework for evaluating student learning. *Teaching in Higher Education*, 13(4), 477–487. <https://doi.org/10.1080/13562510802169764>

Walker, M., & Unterhalter, E. (2007). *Amartya Sen's Capability Approach and Social Justice in Education*. Springer.

Williams, J. H. (2001). On school quality and attainment. *Learning for a Future: Refugee Education in Developing Countries*, 85–108.

Wolff, J., & De-Shalit, A. (2013). *On Fertile Functionings: A Response to Martha Nussbaum: Journal of Human Development and Capabilities*. 14(1). <https://www.tandfonline.com/doi/abs/10.1080/19452829.2013.762177>

Yoo, S.-S., Mosrur, R., Lee, E., & Andrea, D. T. (2019). Development as education for social justice. *Asia Pacific Education Review*, 20(2), 259–272. <https://doi.org/10.1007/s12564-019-09603-2>

Young, M. (2009). Basic Capabilities, Basic Learning Outcomes and Thresholds of Learning. *Journal of Human Development and Capabilities*. <https://doi.org/10.1080/19452820902941206>

국문초록

교육의 질 평가:

Capability 기반의 교육개발협력 기초 교육의 질

평가프레임워크 개발 연구

이은혜

서울대학교 교육대학원

글로벌교육협력전공

대표적으로 MDG2, EFA, SDG4와 같은 국제사회 교육 개발 협력의 목표는 전세계 수많은 아이들에게 교육의 기회를 제공하는데 기여하였다. 특별히 MDG 두번째 목표인 ‘Achieving universal primary education(보편적 초등 교육의 달성)’을 통해 많은 수의 아이들을 초등학교에 등록하게 하는 획기적인 양적 성과를 이뤘다. 이후 국제 사회는 이러한 성과를 바탕으로 보다 도전적인 목표로서(SDG4) ‘모두에게 통합적이고 평등한 양질의 평생학습 기회’ 확보를 세웠다(UN 2015). 실제 교육 분야로 투입된 원조의 양은 지난 20년간 꾸준히 증가하여 2019년에는 약 159억 달러에 달하였다(OECD CRS 2021).

하지만 국제사회의 꾸준한 정책적 재정적 지원에도 불구하고 많은 통계자료는 MDG이후 국제 사회 교육 지표의 변화가 멈추거나 심지어 일부 지역의 경우 악화되는 모습을 전망하고 있다(UNESCO 2019; UN 2015). 특히 MDG 목표 달성의 주요 척도였던 등록률 지표를 제외하고는, 중도탈락률, 졸업률 등의 지표는 전혀 개선의 여지를 보이지 않고 있다. 지표(Indicator)를 어떠한 상태의 본질이나 과정의 가장 적절한 순간을 포착하는 도구 중 하나라고 가정하였을 때, 등록률 지표는 교육 발전의 긴 과정의 문을 여는 기초 단계라고 할 수 있다. 교육 발전을 논의할 때 가장 보편적이고 기초적인 지표로 활용되는 이유이다. 뿐만 아니라 등록률 지표는 UN이 Tier I 으로 구분한 것과 같이 데이터 수집의 용이성 측면에서도 기초적이며, 지표 측정의 방법론적 측면에서도 기초적이다. 그렇다면 중도탈락률, 졸업률, 혹은 학업성취율과 같은 지표는 어떠한가? 저자가 과도한 일반화의 위험을 감수하고도 상식적 수준에서 말할 수 있는 것은, 해당 지표들은 등록률 지표와는 다른 수준의 복잡성을 내포한다.

본 논문은 등록률 외의 교육 지표의 변화가 멈춘 이유를 단순히 지표의 해석적 다양성 혹은 데이터 수집의 어려움에서 찾고자 하지 않는다. 오히려 우리가 특정 목표의 달성을 점검하기 위해 수행하는 평가에서 지표를 활용하는 것이 얼마나 단편적이며 결과 중심적인 해석을 가져오는지 그 한계를 드러내고자 한다. 그리고 그것을 극복하기 위한 대안으로서 개발평가에서 교육의 질을 평가할 때 지금까지 활발하게 논의되지 않았던 ‘학생과 교수자와의 학습 공간에서의 상호 작용’을 평가의 한 가운데로 놓는 평가틀을

제안한다.

후생 경제 이론으로 노벨 경제학상을 받은 인도의 경제 철학자 아마티아센은 1990년대 Capability Approach를 통해 개발협력의 수많은 분야에 가히 혁명적 영향을 미쳤다. 특히 삶의 질을 평가하는 개념적 틀을 제공하는데 큰 기여를 했는데, 그간 GDP로만 평가되었던 국가의 발전 정도를 교육과 보건의 영역까지 확장 시켰기 때문이다. 본 논문이 제시하는 교육의 질 평가틀을 개념화하는데 있어 Capability Approach가 이론적 근거로서 갖는 핵심적 기여는 개인과 사회의 ‘다양성’을 평가의 매우 중요한 요소로 포함시킨다는 것에 있다. 즉, 개인의 선택의 과정과 그 결과는 개인적, 사회적, 환경적 요인의 다양한 조합에 의한 영향의 결과임을 드러낸다. 이것은 센에 의해 conversion factor라는 용어로 표현되었다.

교육의 질을 평가하기 위한 질적 접근은 상대적으로 많은 노력과 재원이 필요하다. 그럼에도 불구하고 이러한 대안적 평가틀을 제안하는 것은 다음의 이유가 있다. 첫째, 이미 많은 교육학자들의 연구를 통해 교육의 ‘과정’의 질적 개선과 교육 발전과의 상관관계를 보여줬기 때문이다. 여기서 말하는 과정은 학습자와 교수자 간의 상호작용을 이야기하며 공교육의 형태에서는 교실 안 수업 시간의 모습을 예로 들 수 있다. 둘째, 현재의 개발 평가의 접근은 이러한 학습 과정에 대한 평가를 철저하게 배제하고 있기 때문이다. 특히 개발협력의 프로젝트와 프로그램 평가에서 활용되는 지표들은 데이터 수집의 용이성, 측정 가능성 등의 경제적 이유로 가장 핵심적이라고 할 수 있는 학습 과정, 질적 변화에 대한 반영이 너무나 미흡한 실정이다.

본 논문에서 제안하는 기초교육의 질 평가들은 아마티아 쉐의 Capability Approach의 핵심적 가치를 기반하였으며, 기존 교육 분야에서 질 평가를 위해 활용되어온 이론과 접근 방법을 상당 부분 차용하였다. 또한 ‘기초교육의 질 지표(Index)’를 제안함으로써 보다 포괄적인 접근에서의 교육의 질 평가가 이뤄질 수 있도록 하였다. Capability에 기반한 평가는 ‘어떻게’ 평가하느냐에 대한 방법론적 질문 뿐만 아니라 ‘무엇을’ 통해 교육의 질을 평가하느냐에 대한 보다 근본적인 질문에 대한 답을 찾게 한다. 본 연구가 교육개발협력에서의 기초 교육의 질을 정의하고 평가하는데 보다 교육적 관점을 제공하는데 기여할 수 있기를 기대한다.

주제어: 교육개발협력, 개발평가, 교육의 질, capability approach

학번: 2014-31210