

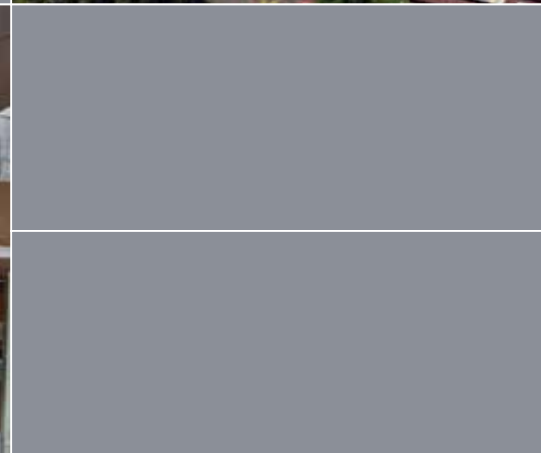


Evaluation Series of NORHED

Higher Education and Research for Development

Theory of Change and Evaluation Methods

Report 4/2014



Norad
Norwegian Agency for
Development Cooperation

Postal address
P.O. Box 8034 Dep. NO-0030 OSLO
Visiting address
Ruseløkkveien 26, Oslo, Norway

Phone: +47 23 98 00 00
Fax: +47 23 98 00 99

Photo Ken Opprann

ISBN: 978-82-7548-795-5

Evaluation Series of NORHED
Higher Education and Research for Development
Theory of Change and Evaluation Methods

October 2014

University of Southern California's
Development Portfolio Management Group

Preface

Norad's Evaluation Department has commissioned an evaluation series of the Norwegian Programme for Capacity Development in Higher Education and Research for Development (NORHED) which will follow the programme from its beginning. The assignments are commissioned by the Evaluation Department of the Norwegian Agency for Development Cooperation (Norad) and are conducted by the University of Southern California's Development Portfolio Management Group.

The overall purpose of the evaluation is to enable learning within the field of higher education and development, so that future investments can become more effective in building capacity as a contribution to development.

In the first assignment, presented in this report, the evaluation team has used existing research to develop a general theory of change for how NORHED as a programme can increase institutional capacity. In essence, this is simply an evidence-based discussion of the programme logic, its underlying assumptions and the context within which the program operates. The general theory of change will be tested in the upcoming evaluations of NORHED.

Through a comparison of the general theory of change developed by the evaluation team and NORHED's implicit theory of change, the evaluation team finds that the programme largely has the potential to achieve its objectives. However, the success of NORHED depends, in part, on system and university level enabling factors outside the programme's control. Due to these uncertainties, the evaluation team recommends that baseline data should be obtained to ensure that the NORHED programme can be evaluated and plausibly attributed to the programme. This is crucial to ensure that future policy decisions regarding the programme are evidence-based.

In the next evaluation assignment, this first study will be followed up by undertaking a baseline study of the current level of institutional capacity in higher education institutions. This will include assessments of the existing monitoring and results framework and the selection mechanism for awarding funds.

We hope that this study will contribute to a debate on how capacity development of higher education institutions can be conducted most effectively with respect to achieving programme objectives, and to a more general discussion on using theories of change to improve development assistance.

The consultants are responsible for the content of the report, including the findings, conclusions and recommendations.

Oslo, October 2014



Director, Evaluation Department

Acknowledgements

This report rests on the contributions of several individuals. Norad's Evaluation Department, NORHED's management and staff, and other Norad staff generously made time available to meet with the team and ensured that the team had access to the documents needed.

The Development Portfolio Management Group managed the work. The DPMG team consisted of Sue Berryman, Jamil Salmi, Lauritz Holm-Nielsen, Gema Zamarro, Juan Saavedra, Margaret McLaughlin, Kara Godwin, Ann Louise Lie, and Linda Morra Imas. Sue Berryman led the team. Jamil Salmi took the primary responsibility for developing a general theory of capacity development in higher education relative to producing more and higher quality graduates and higher quality and more relevant research. In that context he designed and oversaw a special purpose literature review by Kara Godwin on capacity building and organizational change in higher education, with a special focus on those capacities key to producing graduates and research. On behalf of the team, Mr. Salmi also attended the workshop convened by NORHED for the new grantees and their Norwegian university partners in Addis Ababa in mid-March 2014.

Lauritz Holm-Nielsen took the primary responsibility for developing the theory of change evidenced by the design of the NORHED program and the 46 projects that were funded. Ann Louise Lie worked with Mr. Holm-Nielsen to constitute the logical frameworks of each of the 46 projects and to analyze the assumptions explicit and implicit in NORHED's program documents. With Mr. Salmi, Mr. Holm-Nielsen compared the theory of change that emerged for the NORHED program with the general theory of change that reflected empirical studies in the literature.

The methodology team for the study, led by Gema Zamarro, consisted of Gema Zamarro, Juan Saavedra, and Margaret McLaughlin. Gema Zamarro and Juan Saavedra focused on quantitative data collection and analysis; Margaret McLaughlin, on qualitative data collection and analysis. They worked with the theory of change for the NORHED program and the realities of its implementation to identify possible indicators, data sources, and ways to increase the probabilities that effects could be attributed to the program.

Linda Morra Imas provided an independent peer review of the draft document. Her review was complemented by quality assurance reviews by the DPMG Director, Xavier Legrain, and the project's manager, Sue Berryman.

This report is the product of its authors, and responsibility for the accuracy of data included in this report rests with the authors. The findings, interpretations, and conclusions presented in this report do not necessarily reflect the views of Norad's Evaluation Department.

Table of Contents

Preface	ii
Acknowledgements.....	iv
Table of Contents.....	vi
Acronyms and Abbreviations	viii
Executive Summary	x
I. Purpose, Approach and Organization of Report	2
Purpose	2
Approach of the Report.....	3
Organization of the Report	3
II. Preliminary Observations and Assumptions	6
Synergies between Research and Teaching	6
Institutional Complexity and Alignment of Factors	6
The Time Dimension	7
III. A General Theory of Capacity Development in Higher Education	10
System-level Enabling Conditions.....	10
Institutional-level Enabling Conditions	13
Key Dimensions in the Sequence of Causal Links within Universities.....	15
IV. Theory of Change Implied by the Design of the NORHED Program and Projects.....	18
Main Features of NORHED Projects.....	18
NORHED’s Implicit Theory of Capacity Development.....	20
V. Comparison of NORHED’s Implicit Theory of Change with the General Theory of Change	24
VI. Evaluating the Effects of the NORHED Program:	
Causal Pathways, Indicators, Methods, and Data	25
Preliminary Considerations.....	25
Causal Pathways and Indicators to Measure NORHED Effects.....	27
Methods and Challenges for Establishing Attribution of Potential Effects	33
Possible Data Sources.....	36
VII. Conclusions and Recommendations	42
Conclusions.....	42
Recommendations.....	43
References.....	44

Annex 1.....	49
Annex 2.....	56

Figures

Figure 1: General theory of change for producing more/better-qualified graduates and more/higher quality research	11
Figure 2: NORHED’s implicit theory of change for producing more/better-qualified graduates and more/higher quality research.....	21
Figure 3: Decision tree to guide methodological choices for establishing attribution.....	34

Tables

Table 1: Modalities of NORHED Interventions	19
Table 2: Causal pathways and indicators to measure the production of more and better graduates.....	28
Table 3: Causal pathways and indicators to measure the production of more and better research.....	30

Acronyms and Abbreviations

AHELO	Assessment of Higher Education Learning Outcomes
EBSCO	Information/library resources for academia, medicine, law, education, public libraries, corporations, government.
ECDPM	European Centre for Development Policy Management
EconLit	American Economic Association's electronic bibliography
Eldis	Electronic Development and Environment Information System, produced by University of Sussex, England
Embase	Not an acronym. A search engine in medicine.
FRANCIS	Academic bibliographic data base maintained by Institut de l'information scientifique et technique (<u>INIST</u>)
Harzig Publish or Perish	Resources in international & cross-cultural management, research quality and impact, including software for citation analysis
HEI	Higher education institution
IDEAS	Central index of economics and finance research, including working papers, articles and software code
INTRAC	International NGO Training and Research Centre
LMICs	Low- and Middle-Income Countries
LogFrame	Logical framework. The rows of the matrix describe four different types of events that take place as a project is implemented: <i>Activities</i> , <i>Outputs</i> , <i>Purpose</i> and <i>Goal</i> . The four columns provide different types of information about the events in each row: a <i>Narrative</i> description of the event; <i>Objectively Verifiable Indicators: Means of Verification</i> ; and the <i>Assumptions</i> -- external factors that could influence (positively or negatively) events described in the narrative column.
MDGs	Millennium Development Goals
NGO	Non-governmental organization
NOMA	Norad's Program for Master Studies
Norad	Norwegian Agency for Development Cooperation
NORHED	Norwegian Program for Capacity Development in Higher Education and Research for Development
NSSE	National Survey of Student Engagement
NTNU	Norges teknisk-naturvitenskapelige universitet (Norwegian University of Science and Technology)
NUFFIC	Netherlands Universities Foundation for International Cooperation
NUFU	Nasjonalt Utvalg for Utviklingsrelatert Forskning og Utdanning (Norwegian Program for Development, Research and Education)
OECD	Organization for Economic Co-operation and Development
PISA	Program for International Student Assessment
PBL	Problem-based learning
SAT	Scholastic Aptitude Test
SERU survey	Student Experience in the Research University survey
STEM	Science, technology, engineering and mathematics
TIMSS	Third International Mathematics and Science Study
UNICEF	United Nations International Children's Emergency Fund.
WASS-Sense global rankings	Rankings of academic book publishers of Wageningen School of Social Sciences
WHOLIS	Bibliographic database of World Health Organization
WIPO	World Intellectual Property Organization

Executive Summary

Introduction

Norad's program to support higher education in developing countries, the Norwegian Program for Capacity Development in Higher Education and Research for Development (NORHED), has two immediate objectives. These are to produce more and better research relevant to six identified areas/sub-programs¹ and to produce more and better qualified graduates, men and women, in these same areas/sub-programs. By strengthening capacity in higher education institutions in low- and middle-income countries (LMICs), NORHED's longer term objective is to sustainably contribute to a) a more and better qualified workforce, b) increased knowledge, c) evidence-based policy and decision-making and d) enhanced gender equality.

The terms of reference for this first study in connection with evaluating the NORHED program were specified as follows:

This preparatory study is...undertaken to learn about what works and why in NORHED's support to capacity development of higher education institutions in low- and medium income countries. The purpose of this particular study is...to provide a basis for future evaluations of NORHED strategies... through the development of a testable theoretical framework, including a proposal of methods and indicators, with input from, and verified through stakeholder consultation...With respect to time horizon, the theoretical framework shall take into account NORHED's long term perspective, while being explicit on what is testable in the short term, i.e., a three year period.

In response to its terms of reference, the team developed a general theory of change of those factors required to develop the capacities needed to achieve NORHED's objectives. A general theory of capacity development in higher education did not exist, but the materials for constructing such a theory did exist. They reflected a large body of cross-national research findings and experience, including research findings on higher education reform in developing countries. This knowledge could be and was used to ground a robust theory of capacity development and change.

Although NORHED's objectives were clear and even a cursory review of NORHED's documents indicated thinking about causal pathways, the NORHED program lacked an explicit theory of capacity development that could be used to frame the evaluation of the program. However, it was still essential to determine the implicit theory of change that had informed the design of the program and to benchmark this implicit theory against the one developed from cross-national knowledge. Was NORHED's implicit theory of capacity development broadly consistent with the general theory, or did the design of the NORHED program involve implicit assumptions not supported by international research and experience?

Accordingly, the report offers a general theory of change and capacity development in higher education to be tested as part of the planned evaluation of the NORHED program and projects. It explicates NORHED's implicit theory of change, based on reviews of NORHED documents and the construction of the logical frameworks for the NORHED program as a whole and for all 46 projects. It compares the general and implicit theories to determine the consistency between the two. Finally, the report identifies options for attributing results to the NORHED program, appropriate indicators for inputs, outputs, outcomes, and impact, and possible data sources that might be used to measure these indicators.

¹ The six priority programme areas are: Education and Training; Health; Natural resource management, climate change, environment; Democratic and economic governance; Humanities, culture, media, and communication; and Capacity development in South Sudan.

Assumptions common to both theories of change

Both the general and the implicit NORHED models of capacity development for higher education make three assumptions. One is that the joint production of—synergy between—education and research makes the separation of results between research output and graduates somewhat artificial. For these reasons, both theories of change look at the causal chain leading to improved research and graduates in a single framework.

The second assumption, based on recent work on the determinants of excellence in research universities, is the need for full alignment among the main drivers of success to achieve high levels of performance. In complex environments such as the universities supported by the NORHED program, better quality research and teaching outcomes cannot be obtained without the proper combination of all key factors contributing to these outcomes.

The final assumption involves the time needed to achieve results. Improvements in research and teaching do not happen instantly or even quickly. They are long-term processes that require many years of sustained interventions. The time factor has implications not only for the design and duration of donor interventions, but also for the measurement of results and impact.

Findings and conclusions

General theory of higher education capacity development

The general theory of higher education capacity development models the change process at the system level, the university level, and within the university. System-level and university-level factors constitute enabling conditions for changes within the university that lead directly to more and better graduates and research. When the system-level and university-level conditions are met, they support changes within the university and increase the likelihood that a NORHED-funded project will meet its development objectives. When they are not met, these conditions convert into risks to the achievement of the outcomes sought by NORHED.

System-level factors that most systematically influence the situation of individual universities come under governance and funding: i) the degree of institutional autonomy and academic freedom enjoyed by higher education institutions; ii) the accountability instruments (quality assurance and accreditation, assessment of learning outcomes, labor market results, research assessment) linked in one way or the other to performance measures of higher education institutions; iii) student support and financial aid to protect students that belong to disadvantaged groups; and iv) the funding available for research and teaching, especially for the public universities that are heavily dependent on state resources.

University-level factors that most directly affect endeavors within the university are university leadership that ensures that the departments, faculties or schools involved in project implementation are fully supported by the university and aligned with key processes such as internal quality assurance and institutional research, allocation of financial resources, performance rewards systems, and student support mechanisms.

Input factors within the university that are expected to affect the quality and number of graduates and research products are: i) the selection process for and academic qualifications of incoming students; ii) the qualifications, experience and motivation of academic staff that are determined by institutional recruitment policies, faculty development programs and incentive systems; iii) the linkages with industry and the

community² that can increase the relevance of students' learning experiences and the production of university research that can be used to solve priority a country's economic and social problems; and iv) the learning environment and research infrastructure, including well-equipped scientific laboratories, libraries and an appropriate digital infrastructure.

When these input factors (students, academics, linkages with industry, learning and research infrastructure) are present, they are assumed to combine to produce a relevant curriculum that provides a collaborative and proactive learning process and dynamic research activities. These, in turn, ultimately translate into high quality graduates and high quality research outputs.

Implicit NORHED theory of change

NORHED's implicit theory of capacity development was explicated by reviewing NORHED documents, the 46 grantee proposals, and the logframes that were developed for each proposal and for the overall NORHED program. The NORHED program is designed to directly affect factors within the university that will in turn lead to more and better graduates and research. It is not designed to affect the system-level or university-level enabling conditions.

The factors expected to promote the achievement of the desired outcomes are: i) academic recruitment and staff development; ii) enrolment system and procedures; iii) institutional capacity building (via activities that improve physical capacity, research capacity and support, and managerial and leadership capacity); and iv) translational research and outreach (via activities such as research conferences and workshops, research publication and dissemination, knowledge communities, translation of research findings into policy recommendations).

The NORHED implicit theory of change compared to the general theory of change

The theory of change implicit in the NORHED program is broadly consistent with the general theory of change within the university. Most of the main differences are in the details. The one exception is NORHED's limited attention to improving the study environment. Out of an understandable concern for sustainability, the program supports up to 30% of the student intake from marginalized groups. Thus, it implicitly assumes that the system will take care of the financing and social needs and study environment requirements of the rest of the students.

As a collective program, then, NORHED has most of the necessary conditions in place for becoming a successful program if the enabling assumptions are met. However, individual projects are implementing in highly variable local conditions in a large number of lower capacity countries that span three global regions. It is only realistic to expect that at least some, if not most, participating institutions will fail to meet one or more of the enabling conditions key to success, such as a good governance environment, adequate financing, or meritocratic, transparent and fair staff recruitment procedures. Thus, the individual projects are at risk of implementation problems that can affect the likelihood that they can achieve their objectives.

Reform at the higher education level is complex, and NORHED seeks reform in difficult contexts. The inescapable conclusion is that NORHED and the grantees must constantly monitor inputs and outputs and track progress toward the projects' intended objectives. This investment, if acted on, can yield huge benefits. It catches projects that are failing in time to be able to get them back on track.

² "Industry and community" is defined broadly as all key stakeholders that are likely to employ the graduates of higher education institutions and to use their research products. These include firms, policymakers, government agencies, local authorities, and NGOs.

Attribution of results to NORHED indicators and data sources

A thorough analysis of the options for attributing outcomes to the NORHED program is summarized in a decision tree that can be used to guide the choice of approach. The most robust approaches to establishing attribution will probably not be available for the NORHED program.

All options for determining the effects of the NORHED program and being able to attribute these with even modest confidence require baseline data for the outcomes sought. They also require the systematic monitoring of inputs and outputs and the tracking of progress toward achieving the intended outcomes during implementation.

The team concluded that it was impossible to overstate the importance of measuring inputs, outputs, and outcomes before implementation (baseline data) and during implementation. Without baseline data and monitoring and evaluation during implementation, Norad will not know what was achieved, whether any observed achievements can be attributed to the NORHED program, whether a project is failing when there is still time to get it back on track, or why a project might be succeeding or failing.

Norad's recent and thoughtful report (*Can We Demonstrate the Difference that Norwegian Aid Makes?*) found that "none of the reports on grants that were evaluated could reach firm conclusions about the results being achieved. Reports showed well what money was being spent and what direct activities or services were being delivered. But critical questions about whether those services gave rise to real benefits for poor people and other target groups proved elusive." (p.xv)

NORHED and Norad will not be able to learn which aspects of the program worked or did not work and why unless projects have baseline data and systematic monitoring and evaluation of implementation progress. The NORHED program will become just one more Norad program for which firm conclusions about the results being achieved cannot be determined.

The report also identifies the causal pathways (inputs, outputs, outcomes, and impacts) and indicators for measuring the production of more and better graduates and the production of more and better research. It identifies a variety of data sources for measuring these indicators and discusses the advantages and disadvantages of each--for example, the ease and cost of data collection and the likelihood of bias or reliability problems. Although indicators of NORHED's desired outcomes should be common across projects, the 46 specific projects differ in their causal paths to reach the outcomes desired by NORHED. The specifics of the projects being evaluated will determine the final selection of indicators and data sources for inputs, outputs, and interim outcomes.

Recommendations

All recommendations are relevant to Norad's Evaluation Department and NORHED's program management, but recommendations 1 and 5 are especially relevant to the Evaluation Department; recommendations 2, 3, and 4, to NORHED's program management.

1. Use the general theory of change as the theoretical framework for evaluating the NORHED program. Although NORHED's design is implicitly and broadly consistent with the general theory of change, the general theory of change has the advantage of being closely linked to international research findings that expand the options for interpreting data from assessments of NORHED's investments.
2. Ensure that all projects:
 - a. Have adequate baseline data for inputs, outputs, and outcomes; and

- b. Systematically monitor inputs and outputs and track progress toward objectives during implementation.
3. As part of the measurement initiative, a limited set of outcome indicators needs to be selected that can be used across the projects. The indicators selected should be measurable with data that are commonly available across grantees, easy and reasonably inexpensive to collect, and minimally subject to bias and reliability problems.
4. Monitor the status and effects of the enabling conditions on how well grantees can use NORHED-financed inputs to achieve their objectives. As noted, NORHED has most of the necessary conditions in place for becoming a successful program if the enabling conditions are met. However, the NORHED program is operating in environments where these enabling conditions are at risk. If the enabling conditions are often found to be absent and crippling to the best efforts of grantees, it may be advisable to adjust the NORHED program design by collaborating with national and international partners to take concerted action to get the necessary conditions in place. In the meantime, NORHED can focus attention on these conditions through dialogue with national and university-level leaders.
5. Select the approach to be used for assessing the effectiveness of the NORHED program. DPMG recommends that Norad draw interim conclusions about the effectiveness of the NORHED program by evaluating a sample of projects that have been implementing for a sufficient number of months to allow progress and problems to surface. In the spirit of the formative evaluation tradition, DPMG also advises that projects be evaluated a sufficient number of months prior to closing to give grantees and NORHED a chance to get failing projects back on track. DPMG understands that Norad may wish to separate the evaluation of the NORHED program from its operation. However, it has found that the formative model adds more value in terms of salvaged projects and in terms of learning by program staff and grantees.

I. Purpose, Approach and Organization of Report

Purpose

1. The Evaluation Department in Norad specifies its objectives for this report in its Terms of Reference (TOR) that Annex 1 presents in full:

This preparatory study is a first in a line of evaluation-assignments undertaken to learn about what works and why in NORHED's support to capacity development of higher education institutions in low- and medium income countries. The purpose of this particular study is to contribute to a discussion with the NORHED administration and other stakeholders about how the program can make a difference, and to provide a basis for future evaluations of NORHED strategies. This will be done through the development of a testable theoretical framework, including a proposal of methods and indicators, with input from, and verified through stakeholder consultation. It is expected that a discussion of NORHED strategies will both contribute to learning among stakeholders and ensure that future evaluation assignments are relevant. The focus of this call-off will be on the effects of capacity development at the level of institutions in the countries where projects are based. With respect to time horizon, the theoretical framework shall take into account NORHED's long term perspective, while being explicit on what is testable in the short term, i.e., a three year period (pp.1-2).

2. The TOR focused this work on a theory and measures of change in capacities (capacity development) that could allow grantees to improve their graduate and research outputs.
3. Higher education institutions play critical roles in fostering the knowledge, insights, innovative abilities and creative thinking needed for designing and implementing effective poverty alleviation strategies in the context of a green economy that is low carbon, resource efficient and socially inclusive. Efforts to develop the capacity of higher education institutions in low and middle-income countries, as exemplified by Norway's recently launched NORHED program and projects, are of great significance in that respect.
4. Higher education institutions contribute to poverty reduction and sustainable development by: i) training a qualified and adaptable labor force, including high-level scientists, professionals, technicians, teachers in basic and secondary education, medical professionals, and future government and business leaders; ii) generating new knowledge; and iii) providing the capacity to access existing stores of global knowledge and adapt this knowledge to local use. Universities are unique in their ability to integrate and create synergy among these three dimensions (World Bank, 2002).
5. Of particular importance is the role of higher education in furthering the United Nations Millennium Development Goals (MDGs). It is doubtful that any low-income country could make significant progress toward achieving the MDGs for education—universal enrolment in primary education and the elimination of gender disparities in primary and secondary education—without a robust higher education system. Higher education plays a key role in supporting basic and secondary education through the training of teachers and school principals, the contribution of specialists to curriculum design and educational research, and the establishment of admission criteria that influence the content and methods of teaching and learning at the secondary level.

6. What is true for education applies also to the other MDGs. For instance, the contribution of medical education at the tertiary level, especially the training of medical doctors, epidemiologists, public health specialists, midwives, nurses and hospital managers, is essential for meeting the basic health MDGs (World Bank, 2002). Improving the environment and mitigating the devastating effects of climate change is impossible without the participation of researchers, scientists and other relevant professionals.

Approach of the Report

7. Norad has defined the main two results of the NORHED projects as: i) producing more and better research; and ii) preparing more and better-qualified graduates.³ To ensure that the planned project activities achieve these results, it is essential to establish their alignment with the factors most likely to bring about the expected positive transformation of participating higher education institutions. This requires a proper understanding of the causal sequence of how capacity development can help higher education institutions improve their research and graduates output. In other words, the TOR focuses this work on a theory and measures of change in capacities (capacity development) that can allow grantees to improve their graduate and research outputs.

8. In response to its terms of reference, the team developed a general theory of change of those factors required to develop the capacities needed to achieve NORHED's objectives. A general theory of capacity development in higher education did not exist, but the materials for constructing such a theory did exist. They reflected a large body of cross-national research findings and experience, including research findings on higher education reform in developing countries. This knowledge could be and was used to ground a robust theory of capacity development and change.

9. Although NORHED's objectives were clear and even a cursory review of NORHED's documents indicated thinking about causal pathways, the NORHED program lacked an explicit theory of capacity development that could be used to frame the evaluation of the program. It was nonetheless essential to explicate the implicit theory of change that had informed the design of the program and to benchmark it against the general theory of change. Was NORHED's implicit theory of capacity development broadly consistent with the general theory, or did the design of the NORHED program involve implicit assumptions not supported by international research and experience?

10. Finally, the report identifies options for attributing results to the NORHED program, appropriate indicators for inputs, outputs, outcomes, and impact, and possible data sources that might be used to measure these indicators.

Organization of the Report

11. Chapter 2 discusses challenges and assumptions underlying both the general theory of change and NORHED's implicit theory of change. Chapter 3 offers a general theory of change and capacity development in higher education to be tested as part of the planned evaluation of the NORHED program and projects. In formulating this theory, which reflects a set of hypotheses and assumptions about relationships among the key variables influencing the expected outcomes, the authors combined two main sources of information: i) an extensive review of the relevant literature; and ii) lessons of experience derived from analytical and policy work on higher education reform and capacity development in more than 80 developing and transition countries over the past two-and-half decades.

³ By strengthening capacity in higher education institutions in LMICs, NORHED's longer term objective is to sustainably contribute to a) a more and better qualified workforce, b) increased knowledge, c) evidence-based policy and decision-making and d) enhanced gender equality.

- 12.** Chapter 4 explicates the theory of change implicit in the NORHED program. It is based on a review of NORHED program documents, a detailed analysis of all 46 projects selected in the first round of grants, and the reconstruction of the logical frameworks linking inputs to results for the overall program and each of the 46 grants.
- 13.** Chapter 5 compares both theories of change to detect possible gaps or elements of misalignment between NORHED's implicit theory of change and that based on international research findings.
- 14.** Chapter 6 identifies options for attributing results to the NORHED program, appropriate indicators for inputs, outputs, outcomes, and impact, and possible data sources that might be used to measure these indicators. It discusses the advantages and disadvantages of alternative data sources, such as the potential for bias or ease and cost of data collection.
- 15.** Although indicators of NORHED's desired outcomes should be common across projects, the 46 specific projects differ in their causal paths to reach the outcomes desired by NORHED. The specifics of the projects being evaluated will determine the final selection of indicators and data sources for inputs, outputs, and interim outcomes.
- 16.** Chapter 7 summarizes the conclusions of this work and recommendations for the Norad Evaluation Department and NORHED's program management.

II. Preliminary Observations and Assumptions

17. Any theory of capacity development in higher education has to confront at least these important considerations.

Synergies between Research and Teaching

18. Unlike what happens at the lower levels of education, the interaction between research and teaching adds a dimension of complication to the measurement of results in higher education. Besides inculcating appropriate values and attitudes to young people, the main purpose of primary and secondary education is to transmit existing knowledge. By contrast, the core mission of universities is to create new knowledge, in addition to teaching existing knowledge. As a matter of fact, the synergy between the production and dissemination of knowledge is one of the characteristics and strengths of universities (Boyer, 1990; Clark, 1995). Quality teaching at the university level, whether in highly applied or academic fields of study or in the hard sciences or the humanities, requires that teachers be conversant with the current state of knowledge in their subject areas. Currency of knowledge is best secured by integrating the teaching staff with the international community and by exchanging their knowledge and research results freely with their peers.

19. The joint production of education and research makes the separation of results between research output and graduates somehow artificial. The quality of research influences, to a large extent, the quality of teaching and learning. And some studies argue that teaching can also positively influence the direction and quality of research (Becker and Kennedy, 2004). This feature of academic life needs to be taken into consideration when examining the determinants of the quality of graduates. For these reasons, the proposed theories of change—the general one and the one specific to NORHED—look at the causal chain leading to improved research and graduates in a single framework.

20. It is worth noting, however, that the interconnections of teaching and research do not play out in the same way in all countries and in all types of universities. They depend on the structure of the system into which universities operate and on the nature of the prevailing incentives (Halliwell, 2008). Interaction is less strong in dual systems, such as France, Germany and Russia where a large share of research has traditionally been concentrated in separate research institutes. In these cases, research institutes do not always have direct access to the brightest students, while university students have little exposure to the leading researchers who are active in the specialized research academies/institutes. By contrast, the interaction between research and teaching is higher in research-intensive universities, such as in Denmark, Switzerland, the U.K. and the U.S.

21. Furthermore, the range of monetary and prestige incentives that countries and higher education institutions use to reward good research and teaching affect directly the behavior of academics and the likelihood of a strong link between the two activities. For example, under the influence of the global rankings, a growing number of universities are offering special monetary incentives for publications in prestigious international journals, which can result in excessive recognition of research contributions compared to teaching.

Institutional Complexity and Alignment of Factors

22. Recent work on the determinants of excellence in research universities points to the need for full alignment among the main drivers of success to achieve the highest levels of performance (Altbach and Salmi, 2011). Along the same line, the theory of change and capacity development proposed in this document rests on the premise that, in complex environments such as the universities supported by the NORHED program, better quality research and teaching outcomes cannot be obtained without the proper

combination of all key factors contributing to these outcomes. It is not sufficient to focus on one aspect, for example injecting significant amounts of additional financial resources, while neglecting the other drivers of performance.

23. Theories of change in basic education point to a large number of factors affecting educational outcomes, such as the quality of teaching, time spent on tasks, the quality of educational facilities, the curriculum, or language of instruction. This element of complexity is likely to be even greater in the case of higher education institutions, considering their multiple missions of teaching, research and service to the community and the local economy.

24. To give an illustration of critical linkages that are not always systematically considered but would need to be taken into account, a recommendation from the evaluation literature could be that time spent on active and collaborative learning activities—as opposed to traditional lecturing—is an important factor of effective learning. However, if low-income students must work at the same time as they study, it is more difficult for them to participate in active learning. Adequate financial aid is therefore indispensable to allow all students to benefit from new pedagogical approaches.

25. The appreciation of these elements of complexity is consistent with the findings of a recent review of theories of change in development. “In tension with the drive for more assurance of results, there is a growing recognition of the complexities, ambiguities and uncertainties of development work, involving complex political and social change in dynamic country contexts. Theory of change thinking is viewed as one approach to help people deal positively with the challenges of complexity (Vogel, 2012, p. 11).”

The Time Dimension

26. The last assumption that needs to be mentioned concerns the time needed to achieve results when it comes to the transformation of higher education institutions. Improvements in research and teaching do not happen instantly or even quickly.⁴ Institutional change and capacity development are long-term processes that require many years of sustained interventions. For example, to start a new program at the master’s level would require designing an appropriate curriculum and possibly sending a few promising young academics overseas to obtain a Ph.D., a process likely to take four to five years. Then it would take another two years to graduate a first cohort of master students. Similarly, to build up the research capacity of a university department in a partner country, one would need to set up new laboratories and train a minimum core of researchers. Once the new team is in place, it would take several years to see the results of the new research activities.

27. However, donor agencies supporting higher education institutions in developing countries are not always able to factor into the design and organization of their programs the fact that it takes a long time to bring about meaningful change. Their support is often embedded in projects whose duration does not exceed four or five years, reflecting regular budget cycles and common restrictions affecting the length of

⁴ A poem by Nikos Kazantzakis (van Deuren, 2013, p. 23) graphically reminds us that development cannot be rushed. “I remember one morning when I discovered a cocoon in the bark of a tree, just as the butterfly was making a hole in its case and preparing to come out. I waited a while, but it was too long appearing and I was impatient. I bent over it and breathed on it to warm it. I warmed it as quickly as I could and the miracle began to happen before my eyes, faster than life. The case opened, the butterfly started slowly crawling out and I shall never forget my horror when I saw how its wings were folded back and crumpled; the wretched butterfly tried with its whole trembling body to unfold them. Bending over it, I tried to help it with my breath. In vain. It needed to be hatched out patiently and the unfolding of the wings should be a gradual process in the sun. Now it was too late. My breath had forced the butterfly to appear, all crumpled, before its time. It struggled desperately and, a few seconds later, died in the palm of my hand.”

financial commitments to aid programs and projects.

28. The time factor does not have implications only for the design and duration of donor interventions but also for the measurement of results and impact. Donor agencies must integrate this important variable into their evaluation framework, being aware in particular that the effects of their capacity development interventions may not become visible until many years after the end of the program or project.

III. A General Theory of Capacity Development in Higher Education

29. As already noted, a general theory of capacity development in higher education did not exist, but the materials for constructing such a theory did exist. They reflected a large body of cross-national research findings and experience, including research findings on higher education reform in developing countries. This knowledge could be used to solidly ground a robust theory of capacity development and change and to expand the options for interpreting findings for the NORHED program.

30. Figure 1 represents the theory of change for increasing the supply and quality of graduates and improving the production and quality of research by strengthening the capacities that affect these two outcomes. The figure identifies two sets of context factors—system-level and institutional level dimensions—that affect the performance and sustainability of higher education institutions by directly influencing their mode of operation. The figure then models the inputs and intermediary results that, according to the literature and international experience, lead to better graduates and research.

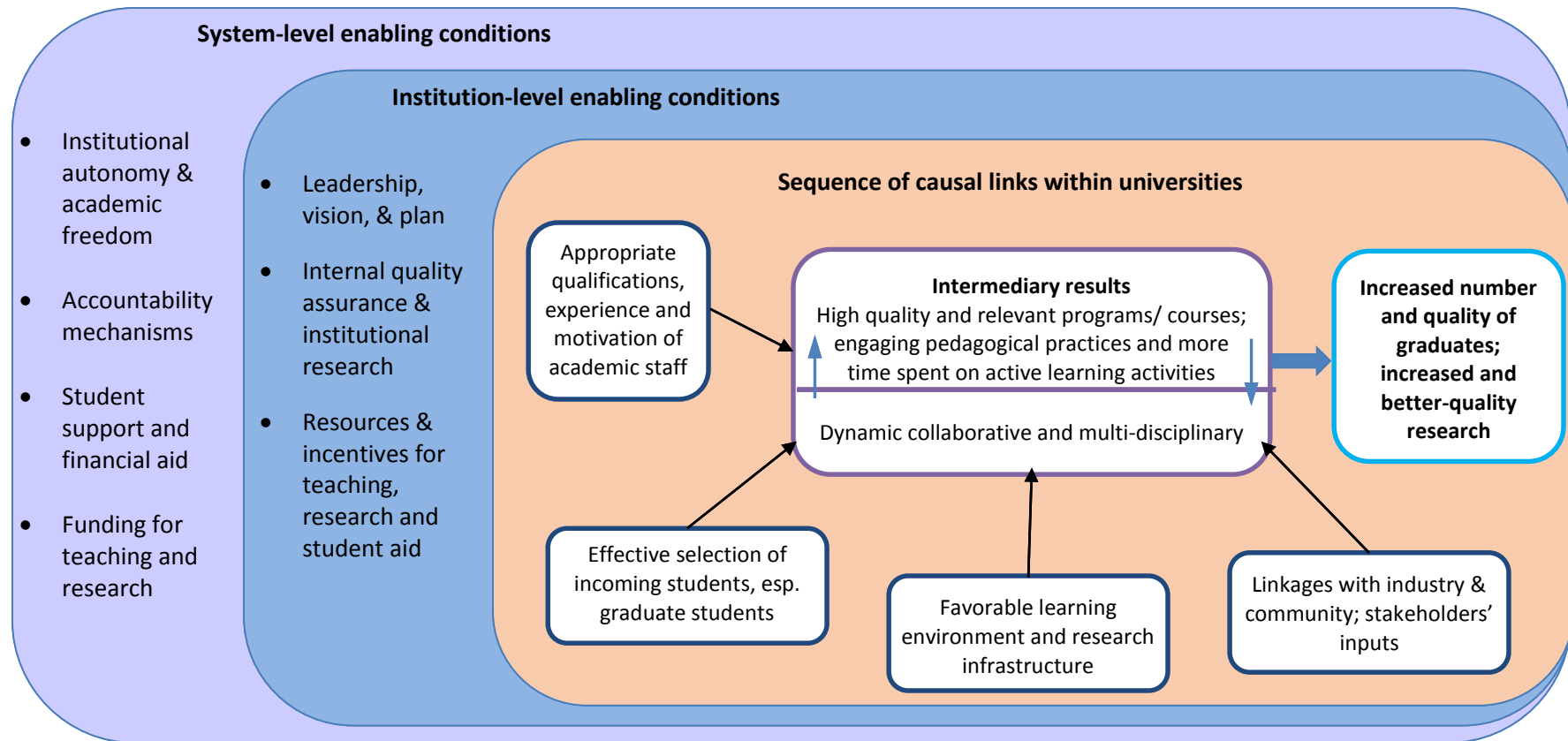
System-level Enabling Conditions

31. Higher education institutions do not operate in a vacuum. To understand the determinants of their performance, it is essential to take into account key forces at play at the level of what could be called the tertiary education ecosystem within which universities and other institutions evolve. These forces can have a facilitating or constraining effect, depending on the circumstances (Salmi, 2011).

32. The tertiary education ecosystem includes the following key elements:

- **Macro environment:** the overall political and economic situation of a country, together with the rule of law and the enforcement of basic freedoms that influence the governance of tertiary education institutions (*e.g.*, appointment of university leaders), their level of funding, their academic freedom, and safety in the physical environment.
- **Leadership at the national level:** the existence of a vision and a strategic plan to shape the future of tertiary education and the capacity to implement reforms.
- **Governance and regulatory framework:** the governance structure and processes at the national and institutional levels that determine the degree of autonomy that higher education institutions enjoy and whether and how they are held accountable. This is especially important for human resource policies and management practices that affect the abilities of higher education institutions to attract and keep qualified academics.
- **Quality assurance framework:** the institutional setup and the instruments for assessing and enhancing the quality of research, teaching, and learning.
- **Financial resources and incentives:** the absolute volume of resources available to finance tertiary education (mobilization of both public and private resources) and the ways in which these resources are allocated to various institutions.
- **Articulation and information mechanisms:** the linkages and bridges between high schools and tertiary education and among the various types of higher education institutions, all of which affect the academic characteristics of incoming students.
- **Location:** the infrastructure and the economic, social, and cultural characteristics of the geographical location of any institution that determine institutions' abilities to attract outstanding scholars and talented students.

Figure 1: General theory of change for producing more/better-qualified graduates and more/higher quality research



- Digital and telecommunications infrastructure: the availability of broadband connectivity and end user devices to enable higher education institutions to deliver educational, research, and administrative services in an efficient, reliable, and affordable way. This is particularly important for the capacity of higher education institutions to implement active and collaborative pedagogical practices, develop distance education programs, and participate in international collaborative research networks.

33. International experience shows that governance and funding are the factors that most systematically influence the situation of individual universities--specifically: i) the degree of institutional autonomy and academic freedom enjoyed by higher education institutions; ii) the accountability instruments (quality assurance and accreditation, assessment of learning outcomes, labor market results, research assessment) linked in one way or the other to performance measures of higher education institutions; iii) student support and financial aid for students from disadvantaged groups; and iv) the funding available for research and teaching, especially for the public universities that are heavily dependent on state resources.

34. In the case of institutional autonomy, Aghion and colleagues (2009) have observed that, besides public funding level and weak development of competitive funding, the stark difference in management autonomy is one of the main factors explaining the lower performance of European research universities in international rankings, compared to their North American counterparts. Recent work on the poor performance of Ibero-American universities has also identified inadequate governance as a key determinant (Bernasconi, 2013; Salmi, 2013).

35. Financial resources affect higher education institutions in two related ways: not only through the volume of resources mobilized and available, but also through the allocation modalities and the extent to which they are performance-linked (Salmi and Hauptmann, 2006). Competitive research grants, for instance, are a powerful mechanism to motivate academics and to stimulate research. Needs-based scholarships and loans can increase opportunities for students from low-income and other minority groups.

36. Three examples from the authors' field experience illustrate how the ecosystem may affect the effectiveness of donor support. First, if the purpose of a project is to assist a partner university in modernizing its curriculum, for instance, implementing the new curriculum could run into difficulties in countries where the Ministry of Education controls the curriculum centrally and sets national exams applying to all public universities. Second, a project to upgrade the quality of a teacher training college may be hampered by the existence of a national university entrance examination with a traditional bias in favor of high prestige careers such as medicine or engineering. Top scorers are apt to select these higher status fields than pedagogical studies. Finally, the success of projects aiming at developing the capacity of academics to use e-learning is conditioned by the availability of broadband, in terms of technical capacity and price, and the possibility of accessing the Internet without government censorship.

37. Furthermore, it is critical to underline the fundamental differences that exist between the environment in which universities function in industrial countries, including the Norwegian universities, and the developing country environment. These differences shape the characteristics and mode of operation of partner universities that participate in the NORHED program. One example is the preparedness of high school graduates. The much lower results of developing countries in international assessments of learning outcomes, such as TIMSS and PISA, imply that the academic level of incoming university students may not be adequate in many cases. The recent publication of the results of the PISA assessment of high school students' problem solving abilities reveals a large gap between high-income economies and developing countries. On a progressive scale of one to six, the proportion of South Korean students reaching level 6 is 7.6%, but in Brazil only 0.1%. At the other extreme, only 6.9% of Korean students are in the lowest level group, compared to 47.3% for Brazil (OECD, 2013). Second, many developing nations, especially those among the group of least developed countries, lack the critical mass of highly qualified academics that

universities in industrial countries enjoy.

38. A second example is that many developing nations, especially those among the group of least developed countries, lack the critical mass of highly qualified academics that universities in industrial countries enjoy. The rapid expansion of higher education in developing countries during the last decade has resulted in a dearth of qualified researchers and university professors (Lindow, 2009; Mouton, 2007; Nickson, 2012; Yizengaw, 2008). Reporting for the *Chronicle of Higher Education* in 2009, Lindow described the gap in needed academics as reaching “crisis proportions.” While this circumstance undoubtedly impacts university teaching functions, its most adverse impact is on knowledge production, research capacity, and the development of new scholars. Across disciplines in African institutions, for example, Yizengaw (2008) notes that only 70% of required faculty positions are filled, and that figures are as low as 30-40% in some departments. This situation is most dire with regard to senior faculty. In many developing countries, the availability of faculty with a doctorate is low compared to high-income nations (Kotecha, 2012; World Bank, 2012; van Deuren, 2013). Even Brazil, a middle-income country, is suffering serious shortages of academics, especially in the STEM disciplines (Knobel, 2014).

39. A third example is the resources gap between high and low-income countries. This gap can be staggering, as illustrated by comparing the budgets of two public universities of similar size, the University of Michigan in the United States and the University of El Salvador in Central America. Both have an enrollment of about 50,000 students. While the first enjoys yearly resources to the tune of US\$1.6 billion, the Central American country’s flagship university operates with only US\$60 million. Even though one dollar will generally buy more in El Salvador than in Michigan, the ratio between the two is still huge: 27 to 1. The availability—or non-availability—of financial resources is likely to make a major difference in terms of teaching and research infrastructure and levels of academic salaries, influencing in turn the quantity and quality of teaching and research, especially when academics must moonlight to make ends meet.

40. Finally, the governance conditions of many developing country higher education systems are often less than favorable. The political appointment of university presidents/rectors, the lack of institutional autonomy, restrictions to academic freedom, and the civil service status of academics do not make for a favorable environment for stimulating high quality teaching and research.

41. These differences are amplified to the extreme in fragile states that have recently emerged from conflict situations or serious natural catastrophes. This is the case, in particular, of post-conflict countries such as Afghanistan, Liberia, Sierra Leone, or South Sudan—one of the six priorities of the NORHED program.

Institutional-level Enabling Conditions

42. Many if not most university partnerships between higher education institutions in the North and in the South start as the result of individual meetings or contacts among academics and students. While this spontaneous and unstructured approach fits well the spirit of academic life, it is not always conducive to proper institutional capacity development. The 2009 evaluation of the NOMA and NUFU programs supported by Norad found, for instance, that the results of the research capacity development projects had been undermined by the absence of a holistic approach.

... both the NOMA and the NUFU program seem somewhat old-fashioned, for instance compared with similar programs in the Netherlands and Sweden. The NOMA and NUFU programs apply quite traditional capacity building measures. In particular they support primarily education of individual researchers/students and only to a limited extent the wider research environment or research management at both the departmental, faculty and central university level (Norad, 2009, p. 67)

43. Similarly, a recent evaluation of the two Finnish higher education programs concluded that the effectiveness of the partnership projects has been directly linked to their degree of alignment with the overall strategic plan of the concerned universities. “The projects have worked best when the planned activities were consistent with on-going institutional efforts to improve the quality of teaching and learning. But in several cases, lack of institutional support weakened the impact of the projects supported by Finnish aid beyond the departments directly involved” (Salmi, Kärkkäinen, Mukherjee and Uuusihakala, forthcoming).

44. This observation is consistent with an important distinction in the recent development literature, between “capacity building” and “capacity development”. While some academics (Ubels, Acquaye-Baddoo, and Fowler, 2010, for example) and donor agencies such as Nuffic and UNICEF (Nuffic, 2012; Sethi, 2012) use both terms interchangeably, this report acknowledges the need to distinguish between the two concepts. In a 2006 publication, the OECD indicated its preference for “capacity development” because the “‘building’ metaphor suggests a process starting with a plain surface and involving the step-by-step erection of a new structure based on a preconceived design. Experience suggests that capacity is not successfully enhanced in this way” (OECD, 2006, p. 12). In the same vein, in their praxis paper for INTRAC, the International NGO Training and Research Centre in the United Kingdom, Simister and Smith (2010) describe capacity development as an endogenous process and capacity building as more exogenous. Similarly, Ubels, Acquaye-Baddoo, and Fowler (2010) refer to capacity as the “ability of a human system to perform, sustain itself, and self-renew” (p. 4).

45. The implication of this distinction is that traditional capacity building—defined as a purposeful external intervention to strengthen capacity over time” (Simister and Smith, 2010, p. 3)—is associated with the risk of having less effective interventions when they are not adequately aligned with the partner university’s own capacity development strategy. The assumption in this report, therefore, is that successful projects must be designed and implemented with a capacity development rather than a capacity building approach in mind, driven by the partner university’s needs and plans instead of the donor’s own strategy and priorities as the primary concern. Concretely, a capacity development approach implies that, as part of project design and preparation, the concerned Norwegian university would work together with the partner institution to obtain full endorsement from the leadership of the partner institution and assess the degree of alignment between the proposed project and that institution’s development plan.

46. Good institutional governance is also important to ensure the success of capacity development activities (Cloete *et al*, 2002). This requires transparent and objective decision-making criteria and processes, for example when it comes to selecting scholarship beneficiaries among partner university academics and graduate students on purely meritocratic considerations.

47. International experience shows that support from the university leadership is indispensable to integrating donor-funded projects, thereby ensuring that the departments, faculties or schools involved in project implementation are fully supported by the university and aligned with key processes such as internal quality assurance and institutional research, allocation of financial resources, performance rewards systems, and student support mechanisms (student services, financial aid, psychological and academic support).

48. In addition, university leaders must be able to manage constructively and effectively the political dimensions involved in any attempt to bring about significant change in a higher education institution. This involves three basic considerations. The first one is to identify who stands to gain and lose from the proposed change. The second and perhaps more crucial step is the consensus-building phase. Translating a transformation program into reality depends to a large extent on the ability of university leaders to build consensus among the diverse constituents of their institution. The third key ingredient for facilitating acceptance of reforms that challenge the status quo is the availability of additional resources that can be channeled towards supporting all the concerned groups (Barth, 2013; Salmi, 2010).

49. The institutional enabling conditions represented in Figure 1 shape the organizational capabilities identified in the literature on capacity development as key determinants of institutional outcomes. These capabilities are the collective abilities of a “sub-system” or organization, in this case the higher education institutions receiving support under the NORHED program (Baser & Morgan, 2008; ECDPM, 2008; Ubels et al., 2010, p. 6). Morgan (2006) explains that capabilities cover a broad range of collective “hard” skills including those that are technical like financial management or policy analysis, and “soft” skills like the ability to “earn legitimacy, to adapt, and to create meaning and identity” (p. 8).

50. With respect to research capacity development, the institutional enabling conditions would include the formulation of a research strategy, the definition of criteria to decide on priority research topics, and agreements on methods to select and fund research proposals.

51. Besides the organizational capabilities mentioned above, the development literature views individual competencies as the other key sub-component of capacity development (ECDPM, 2008; Morgan, 2006; Nuffic, 2013; Ubels et al., 2010; van Deuren, 2013). In the higher education context, competencies are individual “specific abilities” that contribute to the results and outcomes of universities. In addition to abilities, competencies include the skills, energy, motivations, behaviors, and influence of individuals (Morgan, 2006; Nuffic, 2013). These competencies would be embedded mainly in the students and in the academics.

Key Dimensions in the Sequence of Causal Links within Universities

52. Selection of incoming students. The ability of a university to select academically qualified and highly motivated students is a key determinant of the quality of the future graduates. Many researchers have identified students’ prior academic performance to be the best predictor of academic success (McKensie and Schweitzer, 2001; Zeegers, 2004).

53. As mentioned earlier, the level of preparedness of high school graduates may be insufficient in many developing countries, due to the poor quality of secondary education, making it all the more important to have appropriate selection procedures and criteria in place for access to university education. However, this is usually not possible in open admission systems, such as the francophone Sub-Saharan higher education systems, which operate with a “default” selection process by failing a large proportion of their students in the first two years of undergraduate studies.

54. While most university admission systems still follow a traditional assessment approach, testing mainly the language and math abilities of incoming students, innovative higher education institutions have widened the range of abilities that they consider when looking at potential candidates. Building on Howard Gardner’s theory of multiple intelligences—linguistic, logical-mathematical, spatial, bodily-kinesthetic, musical, interpersonal and intrapersonal—Olin College of Engineering, for example, which started operating in 2007 with a design-based multi-disciplinary curriculum bringing together engineering, entrepreneurship and values, selects its incoming students on the basis of a broader set of criteria than SAT scores, to get a sense of its candidates’ artistic, creative and social abilities (Gardner, 1983).

55. Appropriate selection process and criteria to identify promising young scholars are particularly critical at the graduate level, not only for the preparation of competent professionals with appropriate values, but also for the formation of the new generation of researchers. Building a critical mass of scholars cannot happen without increasing and improving the supply of doctoral students in the pipeline future faculty (Kotecha, 2013; Mouton, 2007; Olsson, Meek, & Cooke, 2013).

56. Academic staff. The qualifications, experience and motivation of academic staff are certainly among the most important determinants of the performance of higher education institutions. They are

shaped, in turn, by institutional recruitment policies, faculty development programs and incentives systems.

57. Experience indicates that effective faculty development programs are built on the recognition of teaching as a scholarly activity and the need to respect faculty knowledge. They work better when academics are brought together into communities of practice, when the program considers the development of the entire career of participating academics, and when they emphasize professional growth rather than a remedial philosophy. Furthermore, the literature on effective practices for improving teaching and learning in universities clearly indicates that a few isolated workshops are not the most appropriate way of building faculty capacity. What is needed is to design capacity development activities as a systematic training program at the institutional level (Qualters, 1995 and 2009).

58. Cloete and Bunting (2013) proposed a framework for analyzing the literature on research capacity development strategies, identifying three key functions found in universities seeking to strengthen the capacity of their researchers:

- Development functions, focusing on research capacity practices and policies that promote a research culture, and ensure the quantity and quality of academic human resources.
- Support functions, focusing on faculty and students as individuals and building micro level capacity. This includes specific skill development, legal and financial assistance for faculty developing grant applications and managing research funds, infrastructure improvements, and mentoring and supervision.
- Incentive functions that can be direct or indirect. Direct incentives are "output oriented" where research is concerned (p. 44). For increasing the number and quality of graduates, direct incentives usually refer to scholarships and financial aid. Indirect incentives are more oriented toward career development for faculty, job opportunities for graduates, and high-status awards for both.

59. Linkages with employers and the community. Employers and the community are defined broadly as all key stakeholders likely to employ the graduates of higher education institutions and to use their research products. They include firms, policymakers, government agencies, local authorities, and NGOs. Close linkages with employers and the community are indispensable to improving the relevance of students' learning experiences and for ensuring that university research is developed and used to solve priority economic and social problems in developing countries (World Bank, 2002; Yusuf and Nabeshima, 2007).

60. The ability of higher education institutions to relate to external stakeholders is an important dimension of capacity building, as it can facilitate access to resources and support, help build legitimacy, and create an "operating space" for relevant teaching and research (ECDPM, 2008; Morgan, 2006; Nuffic, 2013). The participation of practitioners from the world of work in curriculum development and teaching, internships and cooperative programs where students alternate between study periods at the university and periods of practical experience in companies, and attachment of academics to firms are among the most effective mechanisms to make the educational experience of students as relevant as possible (Boeren and Holtland, 2005).

61. To enhance the research activities of universities, university-industry linkages can take the form of collaborative research projects with firms, government and NGOs, secondment of industry practitioners to university labs, and incubators. Kotecha (2013) also suggests university-industry linkages as a solution for retaining and encouraging doctoral students to obtain both experience and more relevant research skills. Higher education institutional capacity, according to Cloete and Bunting (2013), is based on several factors including the ability to attract high quality graduate and postgraduate students and to engage in contract work with local industries and government.

62. Learning environment and research infrastructure. Universities in industrial economies enjoy a rich physical environment in support of teaching, learning and research. By contrast, universities in developing countries—and this is likely to be the case for partner universities participating in NORHED projects—often struggle to provide adequate facilities, including well-equipped scientific laboratories, libraries and an appropriate digital infrastructure, due to their lack of financial resources. Yzengaw (2008) observed that “higher education institutions in Africa have seen little or no infrastructure improvements for the last few decades. Learning infrastructure is widely deficient due to insufficient budget and overdependence on public financing. Access to infrastructure such as the Internet, library, textbooks, equipment, laboratories and classroom space is often limited, resulting in deterioration of quality of education and learning. The poor state of facilities also affects the quality of research and its ability to contribute to societal development and progress.”

63. Intermediary results. The intermediary results box in Figure 1 illustrates how, in high-quality universities, the key input factors (students, academics, learning and research infrastructure, linkages with industry) are combined into producing a relevant curriculum, offering a collaborative and proactive learning process, and undertaking dynamic research activities. These, in turn, translate ultimately into high quality graduates and high quality research outputs.

64. Many authors have discussed the importance of building relevant curricula and provided recommendations on how to achieve this result (Olsson and Cooke, 2013; van Deuren, 2013; Wang, 2012). But much less research has been published on the impact of new pedagogical approaches. A recent OECD review of the literature offers useful initial findings on the effect of innovative pedagogical practices, concluding that active learning approaches that offer opportunities for autonomous and group learning, such as problem-based learning (PBL), tend to be more effective in preparing students to work and live in societies driven by innovation than traditional teaching methods.

Compared to more conventional higher education teaching, PBL can be an effective way to develop different discipline-specific and transferable skills for innovation. Research focusing mainly on medical education suggests that students in PBL programs outperform students in more traditional programs in applying their knowledge to unfamiliar real-world situations. PBL appears to be beneficial for developing thinking and creativity skills such as critical thinking and problem-solving. It seems to also benefit the development of different social and behavioral skills such as motivation, interest, self-confidence, self-directed learning and teamwork (Hoidn and Kärkkäinen, 2014, p. 48).

65. One important feature reflected by the two arrows within the box of intermediary results is the close connection between research and teaching. As underlined in a recent World Bank report, “active researchers can provide high-quality teaching,” by sharing their knowledge, their research results, and their research experience in the classroom (World Bank, 2012, p. 81). Further, when research and teaching are linked, there are greater opportunities for “interactions between researchers and students” creating critical opportunities for mentorship and student development at the graduate and postgraduate levels (p. 81).

66. Second, drawing a connection between research and teaching is an opportunity to provide positive incentives to faculty. It illustrates a university’s commitment to the kind of work and the long-term investment that many researchers make to their discipline and their scholarly interests. Doing so serves as an incentive to faculty to stay in developing countries, to feel appreciated and to take their research knowledge into the classroom. Also, given the doctoral training received by most faculty, the culture of the academic community, “appealing to academics’ intellectual integrity through rigorous educational research, offering research-informed educational approaches and strategies that can be effective in accommodating a diverse student body, and providing faculty with a conceptual framework for engaging in educational scholarship themselves,” will increase an institution’s support for its faculty (Scott, 2009, p. 6).

IV. Theory of Change Implied by the Design of the NORHED Program and Projects

67. Although the NORHED program lacked an explicit theory of change, it was essential to determine the implicit assumptions about change that had informed the design of the program and to benchmark this implicit theory against the one developed out of cross-national knowledge. Was NORHED's implicit theory of capacity development broadly consistent with the general theory, or was NORHED making implicit assumptions not supported by international research and experience?

Main Features of NORHED Projects

68. Under the NORHED program, Norad launched a first call for proposals within six priority areas: i) Education and Training; ii) Health; iii) Natural Resource Management, Climate Change and Environment; iv) Democratic and Economic Governance; v) Humanities, Culture, Media and Communication; and vi) Capacity Development in South Sudan. Norad conducted a thorough selection procedure. A total of 173 applications were submitted in this first round, out of which 46 were selected to receive funding. Each of the selected projects has been designed with a specific results framework and adequate policies and administrative procedures in place to ensure: a) diversity and gender balanced recruitment; (b) fair selection of fellowship holders; and c) appropriate knowledge management systems and data collection protocols.

69. The team analyzed NORHED program documents and all of the selected NORHED project proposals and available grant agreements. It met with Norad and NORHED's managers and staff and participated in the stakeholder workshop in Addis Ababa on March 13-14, 2014. Using these data sources, the team developed a logical framework for each of the 46 NORHED projects⁵ and a comprehensive logical framework (logframe) for the overall NORHED program that is displayed in Annex 2

70. The team's understanding is that NORHED conceptualizes capacity development as strengthened capacity for institutions in developing countries to educate more and better qualified graduates, to increase gender equality, and to increase the quantity, quality, and relevance of research conducted by the countries' own researchers. The immediate outcomes would be:

- Producing more and better qualified graduates, men and women, relevant to the identified areas/subprograms
- Producing more and better research relevant to the identified areas/subprograms

71. The analysis of the 46 projects undertaken for this report finds that the modalities of NORHED interventions could be systematized within the following categories:

- NORHED provides funding for the development of selected study programs and improvement of teaching capacity. Each of the 46 projects has its own design and focus, within the framework of the six selected fields and the range of eligible countries and institutions. The projects support master programs, including sandwich-master programs, defined as programs in which courses may take place in more than one of the participating institutions, and joint degrees and exchange programs, tailor- made short courses, and diploma programs. The specific interventions make use of a selected set of educational tools such as needs assessments, teaching program design, curriculum development, a targeted selection of course material, and study program specific workshops for teachers. The projects have careful procedures for enrolment of the first student

⁵ The 46 grantees have been revising their results or logical frameworks, especially since the NORHED workshop in Addis Ababa in mid-March. The team had hoped to obtain these revisions so that it could update the logframes that had been constructed for all projects. However, to our knowledge, these have not yet become available.

cohort and the development of gender-balanced student support. The projects aim at having in place efficient quality assurance systems for teaching capacity, curriculum development, and development of teaching material; they also encourage accreditation of the study programs.

- NORHED supports staff recruitment and staff development by: i) introducing or enhancing human resource practices; ii) mainstreaming gender perspectives; and iii) establishing staff upgrading programs, including in-service teacher training programs, short courses and training in general teaching methodologies, e-teaching and e-courseware development, and formal academic qualification upgrading. The program provides Master’s degree and Ph.D. scholarships and travel grants for academic staff.
- NORHED encourages joint research in the six selected fields by supporting institutional cooperative actions. These activities can be undertaken under Memoranda of Understanding and in formal networks of institutions. The program supports comparative studies, action research, fieldwork, baseline surveys, and establishment of shared databases. In addition to providing funding for research support such as supervision of younger researchers, post-doctoral grants, and travel grants, the program includes outreach in the form of the organization of research conferences and workshops and research publication and dissemination.
- NORHED supports institutional capacity in the participating institutions in partner countries. These institutions are eligible for NORHED funding to strengthen their physical capacities, for example E-learning platforms, video laboratories, library facilities, laboratories and community laboratories, minor construction and equipment for training, research laboratories, and field stations. Further, the program seeks to enhance managerial and leadership capacity, Learning Management Systems, change program support, diversity sensitivity programs, mentor programs, management training and recruitment programs, women’s leadership training, and the establishment of ethics review boards.
- NORHED encourages social outreach and translational research by supporting knowledge communities, capacity building for external stakeholders, and network activities that include business, governments, civil society, communities and academia. The program includes support of knowledge exchange, translation of research findings into policy recommendations, communication and dissemination of results, and student and graduate placement programs.

72. Table 1 summarizes the main modalities of NORHED interventions and the range of activities to be financed under the program.

Table 1: Modalities of NORHED Interventions

Purpose of NORHED Support	Activities Financed
<i>Development of selected study programs and improvement of teaching capacity</i>	
Establishment of new master’s programs	Technical assistance, academic exchanges
Curriculum design & reform	Technical assistance, training workshops, academic exchanges
Introduction of active and interactive teaching and learning methods	Technical assistance, training workshops, academic exchanges
<i>Staff recruitment and staff development</i>	
Enhancement of the capacity of existing academics	Technical assistance, mentoring, training workshops, academic exchanges, scholarships for master’s and PhDs in Norway
<i>Joint research</i>	
Collaborative research projects	Mentorship, planning grants, research grants

Purpose of NORHED Support	Activities Financed
PhD Studies and Post Doc Programs in-country and in Norway	Scholarships / fellowships
Development of research labs	Technical assistance and equipment purchase
Strengthening research leadership competencies	Technical assistance, mentoring and capacity building workshops
<i>Institutional capacity building</i>	
Improving physical facilities	Technical assistance, equipment and minor construction
Strengthening institutional management	Training, including women's leadership training and learning management systems
Establishing ethics review boards	Technical assistance, transparency and diversity sensitivity programs
<i>Social outreach and translational research</i>	
Systems for knowledge management, information and dissemination of results	Technical assistance and IT equipment purchase

Source: Elaborated by the authors

73. All NORHED projects are subject to annual project performance reviews, based on a dual process of self-assessment and external review process by Norad.

NORHED's Implicit Theory of Capacity Development

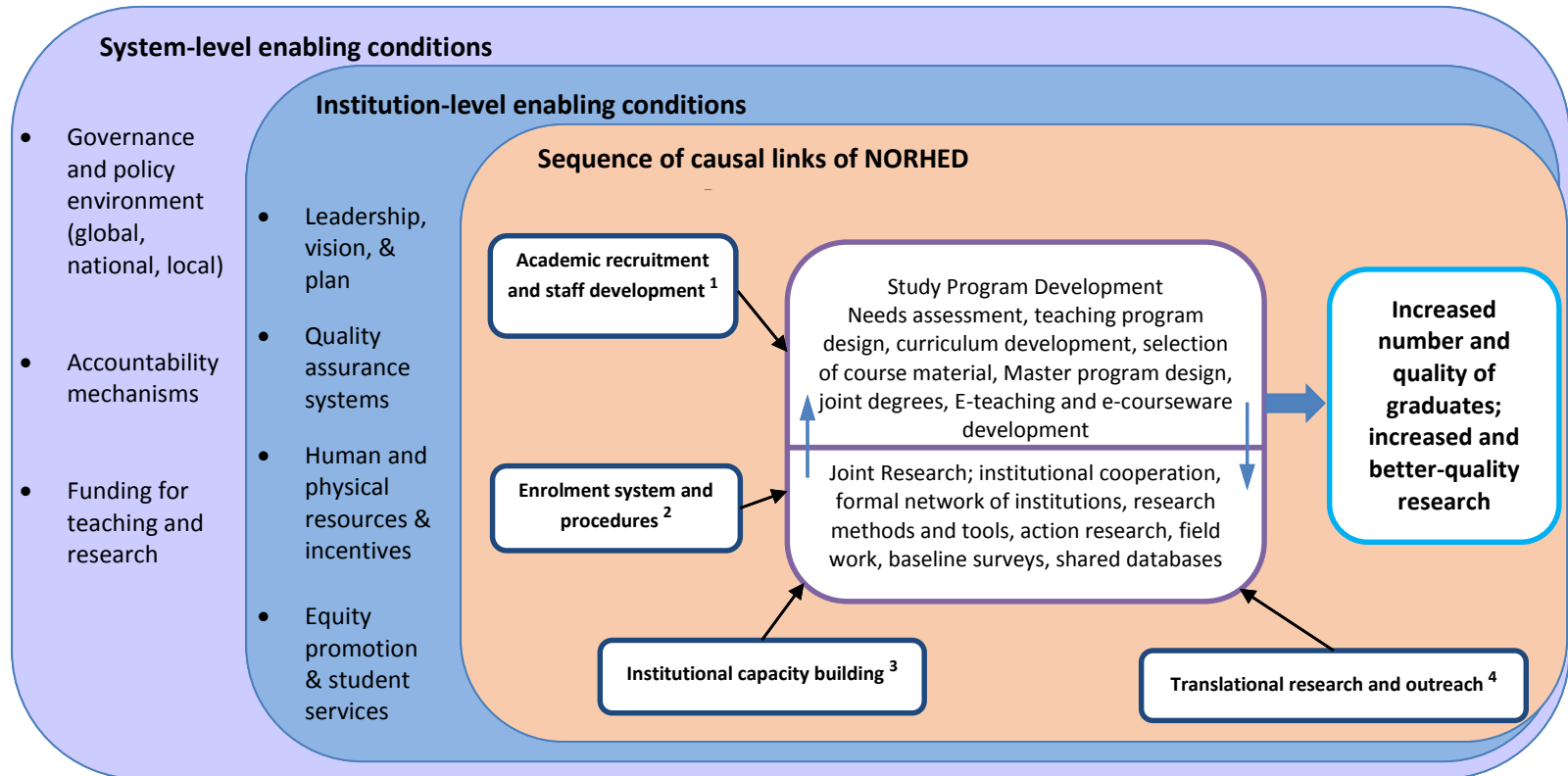
74. Through NORHED, Norad seeks to develop capacity in higher education and research in selected low and middle income countries and in selected economic and social sectors, by producing: i) more, better, and more relevant research; and ii) more and better qualified graduates. Figure 2 attempts to describe the theory of change implicit in the design of the NORHED program and projects, using an approach similar to that in Figure 1, to allow for meaningful comparisons.

75. The main dimensions illustrated by Figure 2 can be described as follows, based on the review of NORHED program and projects documents (see the reconstituted NORHED logical framework in Annex 2):

- Improved study environment, through reduced staff student ratios, gender balance in student body, student safety on campus and in student hostels, availability of health services and food program for all students, and availability of internships and social programs, internationalization by enrolment of more foreign graduate students;
- Improved research environment through empowerment of individual researchers in terms of freedom of choice for research objectives and of publication strategy, balancing the diversity of staff, and by introducing symmetry in international research networks, and improving percentage of co-publication, the status of principal investigators etc.; and
- Staff development and increase in the percentage of graduate degree holders, merit-based career structures, tenure track and full tenure for qualified staff.

76. The NORHED program focuses more on the two latter dimensions than on the first. Out of an understandable concern for sustainability, the program supports up to 30% of the student intake from marginalized groups. Thus, it implicitly assumes that the system will take care of the financing and social

Figure 2: NORHED's implicit theory of change for producing more/better-qualified graduates and more/higher quality research



¹ Good HR practices, mainstreaming of gender perspectives; staff upgrading, in-service teacher training programs, teaching methodologies; formal academic qualification upgrading, Master and Ph.D. scholarships, travel grants

² Gender-balanced recruitment of students; effective selection of incoming graduate students

³*Physical capacity:* E-learning platforms, video labs, library facilities; laboratories; minor construction; equipment for training.

Research capacity and support - Research labs; supervision of young researchers

Managerial and leadership capacity - Learning Management Systems, change program support, Gender sensitivity programs; mentor programs, management training, women's leadership training, ethics review board

⁴Research conferences and workshops; research publication and dissemination; knowledge communities, translation of research findings into policy recommendations

needs and study environment requirements of the rest of the students.

77. Assumptions internal to the NORHED program are that resources are used in accord with the overall purpose and individual NORHED contracts and agreed results frameworks. Specifically, it is assumed that NORHED participants have effective strategies in both Northern and Southern partner institutions, that NORHED complements Norad's research support through Norway's research council and Norad's master program in Petroleum and Energy at NTNU, and that the research base contributes to educational and outreach programs. Finally, it is a prerequisite for the success of NORHED that there is labor market and social demand for graduates in the selected priority areas and that strict anti-corruption rules under Norad's budget are observed.

V. Comparison of NORHED’s Implicit Theory of Change with the General Theory of Change

78. This section assesses the consistency between NORHED’s implicit theory of change and the general theory of change based on international research findings.

79. The theory of change implicit in the NORHED program is broadly consistent with the general theory of change at the level internal to (or within) the university. Most of the main differences are in the details. The main exception is NORHED’s limited attention to improving the study environment, such as scholarships, loans, housing, food programs, campus security, travel grants for undergraduate and graduate students, and fellowships for Ph.D. students. Out of an understandable concern for sustainability, the program supports up to 30% of the student intake from marginalized groups. Thus, it implicitly assumes that the system will take care of the financing and social needs and study environment requirements of the rest of the students. Another difference is that the general theory includes “Linkages with industry”. The NORHED program is more demanding and proactive in that it expects that the program will result in the “translation of research findings into evidence-based policy recommendations”.

80. Through NORHED, Norad can potentially help improve capacities in higher education and research in 16 selected low and middle income countries and in 6 selected economic and societal sectors. By looking at the twin goals of achieving more and better research and more and better graduates in a connected way, NORHED’s strategy builds effectively on the Humboldt tradition of research-based higher education. It seeks to enhance institutional governance (within participating universities), help put meritocratic academic recruitment and annual performance reviews in place, and support increased autonomy and accountability within the selected fields of study and institutions. The assumption is that resources will be used in accord with the overall purpose of the NORHED program, as well as the individual NORHED contracts and agreed results frameworks.

81. As a collective program NORHED has most of the necessary conditions in place for becoming a successful program if the enabling assumptions are met. However, individual projects are implementing in highly variable local conditions in a large number of lower capacity countries that span three global regions. It is only realistic to expect that at least some, if not most, participating institutions will fail to meet one or more of the enabling conditions key to success, such as a good governance environment, adequate financing, or meritocratic, transparent and fair staff recruitment procedures. Thus, the individual projects are at risk of implementation problems that can affect the likelihood that they can achieve their objectives.

82. Reform at the higher education level is complex, and NORHED seeks reform in difficult contexts. The inescapable conclusion is that NORHED and the grantees must constantly monitor inputs and outputs and track progress toward the projects’ intended objectives. This investment, if acted on, can yield huge benefits. It catches projects that are failing in time to be able to get them back on track.

VI. Evaluating the Effects of the NORHED Program: Causal Pathways, Indicators, Methods, and Data

83. In response to the terms of reference, this section of the report focuses on the options for and challenges to measuring the causal paths for the NORHED program and attributing effects to the program. It proposes qualitative and quantitative indicators to measure the main drivers of change (inputs and processes) and the expected results and identifies a range of possible data sources for measuring these indicators.

Preliminary Considerations

84. Rigorous impact evaluation requires a theory-based approach linking inputs to outputs and outcomes. Successful implementation of a theory-based approach to impact evaluation requires: i) mapping out a causal pathway; ii) understanding the context of implementation; iii) defining appropriate counterfactual scenarios and methods to attribute effects, *whenever possible*; iv) undertaking rigorous factual analyses—a particularly important principle when it is not possible to identify viable counterfactual scenarios; and v) collecting and analyzing quantitative and qualitative data to measure relevant changes in inputs, outputs, and processes for each project (White 2011).

85. Determining whether a counterfactual scenario for a NORHED project can be appropriately defined needs to be judged on a case-by-case basis. We anticipate that defining counterfactual scenarios for many of NORHED's grantees is an unrealistic expectation. When defining a counterfactual scenario is not feasible or when only one unit in a given context receives the NORHED support, we propose to emphasize factual analyses.

86. Particularly when n is small, the factual analysis will seek to specify a detailed theory of change and potential alternative hypotheses. In small- n approaches, one may establish attribution when the evidence collected by means of quantitative and qualitative data strongly support the causal links (White and Phillips 2012). In the methodology sub-section that follows later, we provide an overview of methods for small- n evaluation that may be drawn upon to establish attribution of cause and effect.

87. Regardless of whether or not a counterfactual scenario can be defined, finding appropriate and feasible indicators to evaluate NORHED objectives is not straightforward. Tracking changes over time in output and outcome indicators such as research publications and the number of graduates is relatively simple, but measuring effects, such as the quality of research and the relevance of educational programs for local and regional labor markets, is not. In many cases, the best approach will involve combining both quantitative and qualitative data sources and using mixed-method evaluation approaches.

88. For the case of research quality, while indicators exist to measure research effects (e.g. citation counts, H-index or normalized citation impact) there is considerable variation in publication times across disciplines, countries and journal ranks. Publishing in top general interest international journals usually takes much longer than in specialized local ones, even though citations are much higher for articles published in the former than the latter. Similarly, the process of publishing in the social sciences usually takes much longer than in the biomedical sciences. Acknowledging this variation in time horizons to measure changes in impact indicators of research quality will be important when thinking about the overall effects of NORHED projects.

89. Measuring the quality of graduates is even more challenging. In educational systems such as those in Colombia, Brazil, Jordan, Mexico and South Korea, college graduates are required to take a college exit test that in some cases is field specific. This information is usually public and thus valuable for assessing graduate quality: test results are available at the student level and can be tracked over time for different

programs. Unfortunately, very few other educational systems have such college exit tests.

90. Internationally, the only relevant experience is OECD's AHELO initiative. AHELO aims at measuring comparatively the acquisition of generic competencies and professional skills. In a feasibility study conducted in 2011-2012 in the areas of economics and engineering, OECD collected test data from a sample of 23,000 undergraduate students from 258 higher education institutions in 17 countries that volunteered to take part, including five Norwegian universities.⁶ However, it has been hard to interpret the data from this pilot. To date, there is no published report on the findings, suggesting that comparability might be of concern.

91. In the absence of national college exit exams, surveys of graduates and tracer studies have been employed to examine labor market outcomes as proxies of quality and relevance. A common problem of this approach, however, is that response rates are usually quite low and selected.

92. In recent years, more subjective measures have been used, such as employers' feedback. For instance, the World Economic Forum's annual competitiveness ranking uses the results of employers' surveys to compare the quality of education and higher education across countries.

93. An alternative approach is the use of student assessment surveys. One example of this approach is the National Survey of Student Engagement (NSSE), which collects information on key dimensions of the students' academic experience for student in 640 universities and colleges in the United States and 76 Canadian universities. Other examples include the Irish Survey of Student Engagement and similar surveys in Australia and New Zealand.

94. To properly establish attribution of effects in NORHED programs, ideally indicators would need to be collected for treated as well as for potential comparison units. Doing so would be relatively easier and cheaper for indicators based on existing administrative data sources than for impact indicators based on survey data. When comparison units are not feasible quantitative and qualitative indicators would be the base of study to assess that the identified causal links in the theory of change are followed.

95. In this respect, we should distinguish between the indicators that help us monitoring that projects are in the good path for success and indicators that would be the base for assessment of effects of the programs. Monitoring should be based on indicators of inputs and outputs while evaluations of effectiveness should be based on information about outcomes and impact measures.

96. With these considerations in mind, the remainder of this section has five aims. The first aim is to operationalize the theory of change of NORHED activities from section III into causal pathways linking inputs, outputs, outcomes and impacts. We do so separately for the goal of producing more and better graduates and for producing more and better research. As noted earlier, the causal pathway is the starting point for conducting any rigorous theory-based impact evaluation.

97. The second aim is to propose possible feasible indicators to measure each chain in the corresponding causal pathway. We articulate these two aims in a matrix format (tables 2 and 3) to aid with organization and visual appeal. It is worth noting that these indicators are generic in the sense that they map into the implicit theory of change for NORHED and causal pathway linking inputs to impacts. Project specifics might impose refinements on the causal pathways and relevant indicators. Similarly, data availability and costs of collecting it will vary from context to context, and such factors will need to be accounted for in deciding which indicators to measure. For these reasons, at this stage we are agnostic

⁶ All three reports of the AHELO feasibility study are available at <http://www.oecd.org/edu/skills-beyond-school/testingstudentanduniversityperformancegloballyoecdahelo.htm>

about favoring some indicators over others. At a later stage in the evaluation timeline and for each specific program, desirability will need to be weighed against cost.

98. For developing potential indicators we considered NORHED program proposed indicators as developed by the projects and Norad as our starting point. However, we classified indicators as inputs, outputs, outcomes and impact, added some more objective indicators based on national and institutional administrative records, and added indicators for certain dimensions we find relevant. For instance, we added indicators to evaluate relevance for labor market needs beyond self-reports, added more explicit measures of gender balance, among others.

99. The third aim is to briefly discuss issues of potential attribution of effects. We highlight that evaluation of NORHED activities is in most cases a small-n problem with non-random selection. This context rules out experimental methods and quasi-experimental approaches, such as regression discontinuity designs, and may favor the use of factual analyses that we describe in detail below.

100. The fourth aim of the section is to enumerate and briefly discuss potential data sources to measure different indicators. It is worth noting that it is unlikely that any single data source will contain information for all indicators. All data sources we propose have advantages and disadvantages, and we strive to highlight these accordingly. Some data sources might be context-specific and unavailable in other settings, limiting comparability. Ultimately, the choice of data to evaluate each specific NORHED project will require additional contextual knowledge that we currently lack. For these reasons, we are also agnostic about favoring some data sources over others at this stage. Having said that, we anticipate that a good evaluation would require the use of all available indicators based on administrative data, any relevant statistics available at the national and institutional level, and interviews, focus groups and/or tracer studies and other kinds of surveys to complement the available information.

101. The fifth and final aim is to provide some basic general recommendations for deciding which evaluation approach and data sources might be most suitable in assessing the impact of each of NORHED's grants.

Causal Pathways and Indicators to Measure NORHED Effects

102. To assess the potential effects of NORHED projects on the quality of graduates and research, we operationalize NORHED's implicit theory of change (Figure 2) by means of two causal pathways—one for producing more and better graduates and one for producing more and better research. In each causal pathway we logically link inputs, outputs, outcomes and impacts. Under the implicit theory of change, it is posited that if projects lead to desired effects, they could potentially strengthen the capacity of higher education institutions in developing countries to promote growth, sustainable development and to reduce poverty.

103. Table 2 presents the causal pathway for the goal of producing more and better graduates. Table 3 presents the causal pathway for the goal of producing more and better research. Note, however, that these generic pathways will need to be refined and further developed when we adapt them to each specific NORHED grant.

Table 2: Causal pathways and indicators to measure the production of more and better graduates

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
1	Develop new or revise existing educational programs to be more relevant for local, national, regional and labor market needs	New and Revised programs in operation	Educational programs more relevant for local, national, regional and labor market needs	Increase the skill fit between supply and demand of graduates
<i>Indicators</i>	<p>Number of new programs created with more relevance for the local, national, regional labor market in mind (e.g. programs that conducted their own or commissioned a survey of the relevant labor market or consulted with relevant industry and government stakeholders)</p> <p>Number of existing programs with revised curriculum, teacher program design and/or course materials towards increasing relevance for the local, national, regional labor market</p> <p>Number of programs created or revised that received input from industry or government stakeholders</p>	<p>Number of new students enrolled in new and revised programs (total and by gender)</p> <p>Number of graduates from new and revised programs (total and by gender)</p> <p>Whether programs continue to conduct their own or commissioned surveys of the relevant labor market or consulted with relevant industry and government stakeholders. If so, frequency</p> <p>Evaluation of relevance for labor market needs of curricula, teacher program design and course material selection</p> <p>Exit interviews with new graduates to assess readiness to enter job market</p>	<p>Number of graduates finding jobs in relevant local labor markets (total and by gender)</p> <p>Average number of job offers a graduate receives before accepting a job (total and by gender)</p> <p>Number of graduates starting up businesses in relevant local labor markets (total and by gender)</p> <p>Focus groups with select local industry, government and NGO leaders to assess program effectiveness and demand for program graduates</p>	<p>Skill gap (labor supply and demand gap) in relevant industries for new and revised programs</p> <p>Time to first employment for graduates of new and revised programs (total and by gender)</p>

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
2	Establish diversity balanced recruitment procedures	Diversity balanced recruitment procedures operating	Higher proportion of female faculty, staff and students	Greater gender equality
<i>Indicators</i>	Number of procedures for balanced recruitment established	Review and rating of Bachelor/Master/PhD programs modules to assess dimensions such as whether or not they practice diversity-based hiring with preferential recruitment; the level of admissions/HR staff buy-in; whether decision-makers follow and adjust to trends over time or only address immediate shortfalls with respect to gender	<p>Fraction of female students, faculty and staff</p> <p>Extent to which the academic workplace is gender-neutral; interviews with faculty, staff and students</p> <p>Document review (organizational charts, committee rosters) to determine representation of women in academic unit governance</p>	<p>Number and percentage of female graduates</p> <p>Gender gap in wages at the institutional level</p> <p>Gender gap in number of tenured and tenure track faculty positions</p>

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
3	Strengthen the qualifications of academic staff	More and Better prepared academic staff	Increase exposure of students to highly qualified teachers	Increase quality of graduates
<i>Indicators</i>	<p>Recruitment of academic staff with Master or PhD qualifications by relevant unit (total and by gender)</p> <p>Master or PhD training of existing academic staff by relevant unit (total and by gender)</p> <p>Evaluation of the quality of the obtained degrees by looking at rankings of the colleges and universities where available</p>	<p>Number of academic staff with strengthened qualifications (Master/PhD) by relevant unit (total and by gender)</p> <p>Retention rates of academic staff with master and PhD in relevant unit (total and by gender)</p> <p>Interviews with newly hired staff to assess quality of orientation and mentoring programs within the academic unit</p>	<p>Ratio of qualified academic staff with Master or PhD degrees to students in relevant unit</p> <p>Document review (schedules, course assignments) to assess effectiveness of deployment of most qualified teachers</p>	<p>Average wages of graduates from relevant unit (total and by gender)</p> <p>Time to first employment for graduates of relevant unit (total and by gender)</p> <p>Number of graduates becoming entrepreneurs</p>

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
4	Improve Linkages with Industry and Community	Bigger support of Industry and Community	Educational programs more relevant for local, national, regional and labor market needs	Increase the skill fit between supply and demand of graduates
<i>Indicators</i>	<p>Number of events directed to reach industry and community members</p> <p>Extent to which the institution has industry-friendly technology transfer and commercialization policies and what the level of institutional support is for translational science (for example, if they have a medical campus, do they support faculty in taking their discoveries from the laboratory to bedside)</p>	<p>Funds (e.g. scholarships) provided by industry in the area</p> <p>Number Collaboration agreements with industries</p> <p>Number of Internships for students</p> <p>Availability of graduate placement programs</p> <p>Interviews with academic unit administrators to assess their outreach (talks, memberships, meetings) to industry and local community</p>	<p>Number of graduates finding jobs in relevant local labor markets (total and by gender)</p>	<p>Skill gap (labor supply and demand gap) in relevant industries for new and revised programs</p> <p>Time to first employment for graduates of new and revised programs (total and by gender)</p>

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
5	Improve Academic Networks	Increase collaboration with other academic institutions	More knowledge transfers between South-South and South-North networks	Increase quality of graduates
<i>Indicators</i>	Number of events directed to increase collaboration with other institutions (regional versus South-North vs. international professional networks, face to face versus online networks)	Participation in events directed to increase collaboration with other institutions (regional versus South-North vs. international professional networks, face to face versus online networks) Quality index based on college and university rankings when available to assess the quality of existing collaborations	Knowledge transfers within South-South and South-North networks in partnerships (e.g. inform curricula, new teaching methods, etc.) Document review to assess a) frequency of citation of South-led academic work by academics from Northern institutions, and b) vice versa; c) ratio of a to b	Average wages of graduates from relevant unit (total and by gender) Time to first employment for graduates of relevant unit (total and by gender)

Table 3: Causal pathways and indicators to measure the production of more and better research

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
1	Strengthen the qualifications of academic staff	More and Better prepared academic staff	Increase quality and relevance of research activities	Increase quantity and quality of research
<i>Indicators</i>	Recruitment of academic staff with Master or PhD qualifications by relevant unit (total and by gender) Master or PhD training of existing academic staff by relevant unit (total and by gender) Evaluation of the quality of the obtained degrees by looking at rankings of the colleges and universities where available Availability of mentoring programs for junior academic staff	Number of academic staff with strengthened qualifications (Master/PhD) by relevant unit (total and by gender) Retention rates of academic staff with master and PhD in relevant unit (total and by gender)	Number of new research seminars and workshops in relevant unit Number of submitted research grant applications by academic staff in relevant unit (total and by gender) Number of awarded research grant applications by academic staff in relevant unit (total and by gender) Number of scientific manuscript submitted to publication outlets by academic staff in relevant unit (total and by gender) Interviews with senior administrators to assess comparative social and intellectual impact of research within the academic unit Number of policy related reports and briefs (total and by gender) Number and type of other dissemination activities (e.g. media mentions) (total and by gender)	Number of presentations in academic conferences (total, by gender, by North or South origin of researcher presenting) Number of scientific publications (total, by gender and by North or South origin of lead author when applicable) Number of scientific publications in collaboration with Norwegian scientists (total and by gender) Quality of scientific publications (citation index) (total, by gender and by North or South origin of lead author when applicable) Number of patents (total, by gender and by leading institution)

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
2	Improve Research Infrastructure	Increased access to more and better research resources	Increase quality and relevance of research activities	Increase quality of research
<i>Indicators</i>	<p>Number of academic books purchased</p> <p>Number of new subscription to journals</p> <p>Number of new or upgraded laboratories</p> <p>Number of new computers</p> <p>Number of new software licenses</p> <p>Number of new servers</p> <p>Number of new printers</p> <p>Availability of institution sponsored research (e.g. seed funding for research), availability of Contracts and Grants staff, availability of workshops and training in proposal preparation</p>	Use of research infrastructure by academic staff and students	<p>Number of submitted research grant applications by academic staff in relevant unit</p> <p>Number of awarded research grant applications by academic staff in relevant unit</p> <p>Number of scientific manuscript submitted to publication outlets by academic staff in relevant unit</p> <p>Number of ongoing research projects</p> <p>Interviews with academic staff to assess availability of seed funding, sponsored research administrator access, help with budget preparation, time off for proposal preparation, post-award and other services</p> <p>Assessment of comprehensiveness and ease of access to information about sponsored research through review of available online documents</p>	<p>Number of presentations in academic conferences (total, by gender, by North or South origin of researcher presenting)</p> <p>Number of scientific publications (total, by gender and by North or South origin of lead author when applicable)</p> <p>Quality of scientific publications (citation index) (total, by gender and by North or South origin of lead author when applicable)</p> <p>Number of patents (total, by gender and by leading institution)</p>

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
3	Improve Linkages with Industry and Community	Better collaboration with and bigger support of Industry and Community	Increase relevance of research activities	Increase quality of research
<i>Indicators</i>	<p>Number of events directed to reach industry and community members (e.g. informational sessions, fund raiser events)</p> <p>Number of policy briefs</p> <p>Extent to which the institution has industry-friendly technology transfer and commercialization policies and what the level of institutional support is for translational science (for example, if they have a medical campus, do they support faculty in taking their discoveries from the laboratory to bedside)</p> <p>Interviews with research project leaders to discover factors leading to successful partnerships with industry and satisfaction with level of institutional support for collaborative external partnerships</p>	<p>Number of sponsorship provided by industry in the area</p> <p>Number of collaboration agreements and contracts for applied research with industries</p>	<p>Uptake/influence of research in public policies</p> <p>Uptake/influence of research findings and products by local communities/civil society/private sector</p>	<p>Quality of scientific publications (citation index) (total, by gender and by North or South origin of lead author when applicable)</p> <p>Number of patents (total, by gender and by leading institution)</p>

	<i>Inputs</i>	<i>Outputs</i>	<i>Outcome</i>	<i>Impact</i>
4	Improve Research Networks	Increase collaboration with other academic institutions	More knowledge transfers between South-South and South-North networks	Increase quantity and quality of research
<i>Indicators</i>	Number of events directed to increase collaboration with other institutions (regional versus South-North vs. international professional networks, face to face versus online networks)	Participation in events directed to increase collaboration with other institutions (regional versus South-North vs. international professional networks, face to face versus online networks) Quality index based on college and university rankings when available to assess the quality of existing collaborations	Number of scientific manuscript submitted to publication outlets in collaboration with other institutions (total and separately for regional institutions and North-South collaborations). Quality index when available Number of collaborative research grants submitted in collaboration with other institutions (total and separately for regional institutions and North-South collaborations). Quality index when available	Number of presentations in academic conferences (total, by gender, by North or South origin of researcher presenting) Number of scientific publications (total, by gender and by North or South origin of lead author when applicable) Number of scientific publications in collaboration with Norwegian scientists (total and by gender) Quality of scientific publications (citation index) (total, by gender and by North or South origin of lead author when applicable) Number of patents (total, by gender and by leading institution)

Methods and Challenges for Establishing Attribution of Potential Effects

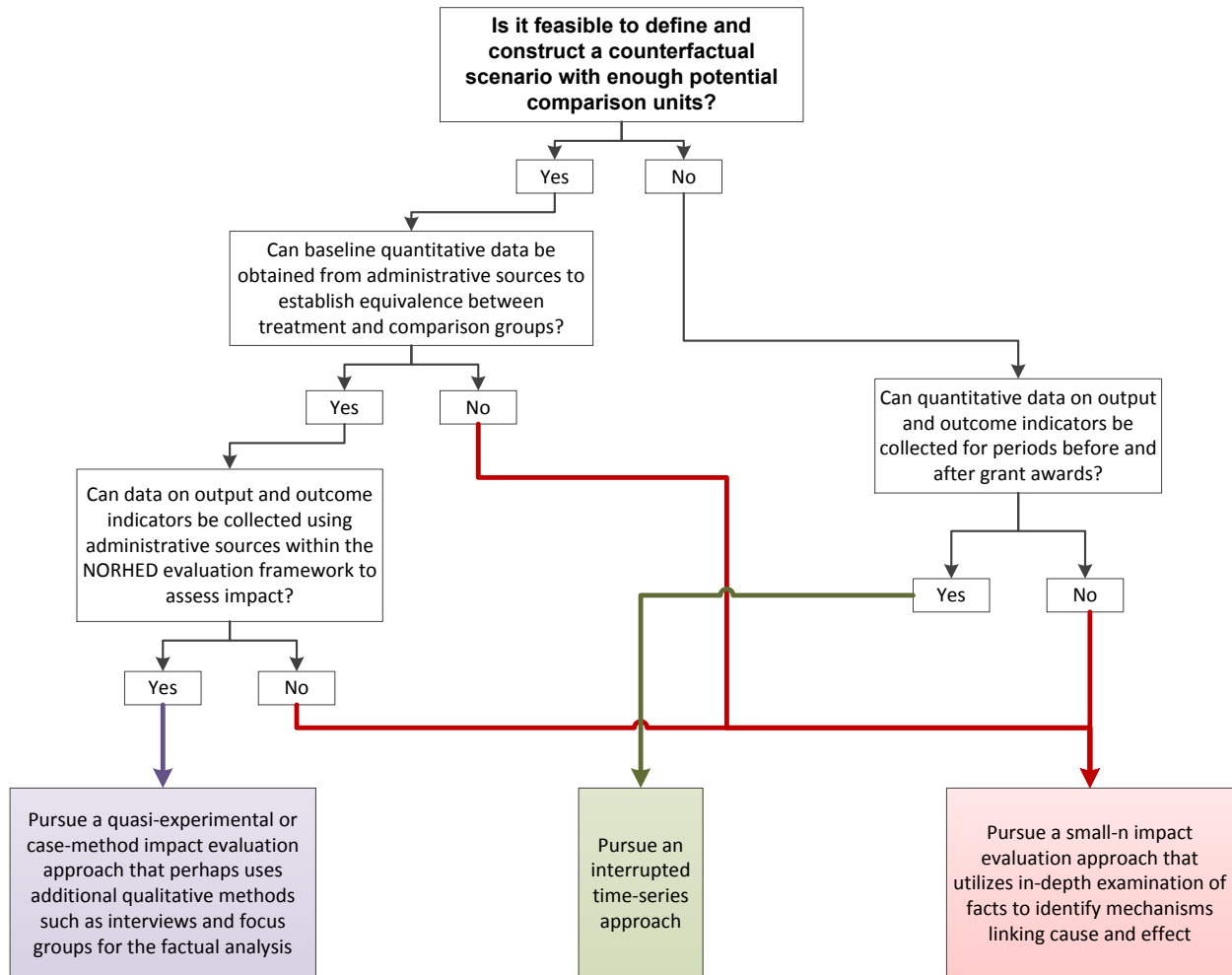
104. In all cases, we will employ a theory-based approach to evaluate the impact of NORHED grants. By gaining a more specific understanding of the implementation context of each grant we will be able to map the causal pathway for each project by fine-tuning the generic pathways outlined previously.

105. Having a detailed causal pathway and a solid understanding of the context of implementation, our attention will turn to choosing the most appropriate methodological approach for attributing potential effects. In choosing the most appropriate methodology, we propose to follow a decision tree like the one depicted in Figure 3.

106. If data on indicators are available before, during and after the introduction of the grant but no counterfactual scenario is feasible, one might be able to use an interrupted time series design (see e.g. Hinrichs, 2012) to examine whether trends in the indicators change before and after the introduction of each grant program. This analysis would allow for a study of timing of effects: whether they occur immediately or only with a delay, whether the trend continues over time, stabilizes, or reverses after a certain time. The disadvantage of an interrupted time-series design, however, is that it can confound other external factors that might be driving observed changes.

107. Ideally, one would like to be able to establish counterfactual conditions to treated units by finding an observationally similar control group to which we compare the results in indicators. Since the allocation of grants was not done at random, the best one can hope for is for the potential comparison group to be similar along observed dimensions. Without random assignment, one cannot guarantee that both units will be comparable along unobserved dimensions. Moreover, if there are knowledge or research spillovers, some of the obvious comparison candidates might not be valid.

Figure 3: Decision tree to guide methodological choices for establishing attribution



108. This caveat notwithstanding, when a counterfactual scenario is feasible and relevant baseline and follow-up quantitative data are available to pursue a quasi-experimental impact evaluation approach, the grant selection mechanism might be used to circumvent some of the biases stemming from unobserved differences between treatment and comparison units. Specifically, if the grant selection process is entirely based on observed characteristics of grant applicants and information on these characteristics is available for all grant applicants, then the selection mechanism is *unconfounded*: conditional on observables, treatment assignment is as good as random. If grant selection is unconfounded, then quasi-experimental approaches such as propensity-score matching that eliminate the biases in observed characteristics that determine grant assignment yield causal effects of grant awards. If selection is only unconfounded conditional on observables and time-invariant unobservables, one could potentially combine propensity score methods with a difference-in-difference methodology in a semi-parametric approach like the one proposed by Abadie (2005).

109. In general, possible candidates to serve as comparison units to NORHED grants might be:

- Other institutional units (e.g. departments) within the same organization that do not receive NORHED funding. The advantage of this approach is that these potential comparison units share the same institutional context. The major disadvantage is that if there are spillovers (see for

example, Romer, 1990), they are more likely to occur within the same organization (e.g. infrastructure improvements) and lead to underestimating the effects of NORHED grants.

- The same institutional unit (e.g. department) in a different organization that does not receive NORHED funding. The major advantage to using such comparison unit is that it limits the extent to which spillovers might occur. The major disadvantage is that the comparison will be affected by differences in institutional environments, some of which will be unobserved.
- Synthetic control group analysis (see Abadie et al., 2007) by which each NORHED treated unit is compared to a convex combination of other untreated institutional units in which more similar units are given more weight in the group. The main advantage is that by pooling multiple possible comparison units into a synthetic control group, comparability will more likely be achieved. The main disadvantage, however, is that this procedure requires extensive amounts of data, particularly for pre-treatment conditions. (Comparability between the treated and synthetic unit is assessed based on pre-treatment similarities.) Therefore, this approach might not be operational for many grants that seek to fund new (as opposed to existing) programs.

110. Note that even when a counterfactual scenario can be identified and there are data to pursue a quasi-experimental approach—exploiting, for instance, grant selection rules, care needs to be taken to properly take into account the correct level of variation. For example, if the assignment unit is a university department but the analysis unit is the researcher, not accounting for the fact that there is clustering in outcomes across researchers within the same department leads to erroneously underestimating the standard errors of the treatment effect. This problem is known as the Moulton problem and to account for it one needs to appropriately cluster standard errors in all analyses (see, for example, Angrist and Pischke 2009).

111. If we determine that data availability, the selection rules, and the availability of a counterfactual condition permit a quasi-experimental design, we will endeavor to guarantee that such an approach will provide meaningful results by, among others, minimizing the probability of Type II errors: not rejecting the null hypothesis of no program impact when the null is in reality false. To do so, we will carry out diagnostic power calculations that employ all information available, including variation in outcomes, the predictive power of baseline covariates, and clustering.

112. If interrupted time-series or quasi-experimental approaches are not a feasible strategy to establish attribution, we propose to pursue small-n impact evaluation approaches based on an in-depth examination of facts to identify mechanisms linking cause and effect. These approaches include the general elimination methodology, process tracing, and contribution analysis. The goal of these approaches is to explain what changed and how by seeking evidence to substantiate the causal pathway linking inputs to outcomes (White and Phillips 2012).

113. These small-n approaches are typically of a mixed-method nature. They aim to establish causality by building credible evidence in favor of a program’s causal pathway. As such they make use of the best sources of data available, whether quantitative or qualitative. In practice, most of the evidence used in these evaluations stems from qualitative data sources such as interviews and documental analysis, which we discuss below.

114. One final challenge that we anticipate when undertaking the proposed theory-based evaluation approach concerns the timing of effects. Changes in inputs and outputs described above should be observed during the NORHED time frame of 2013-2018. In some cases, changes in outcomes might also be observed, although we anticipate that would be less common.

Possible Data Sources

115. The set of indicators presented above represents an ideal list of possible indicators. However, we acknowledge the fact that having access to the sources of data needed to collect all this information can be challenging, especially in some of the countries where NORHED funds activities.

116. In addition, there is not a single source of data that will contain all information to track the indicators we propose to evaluate effects. For this reason, we propose combining multiple data sources when available. We discuss in this section potential data sources and data collection methods to arrive at measures of the proposed indicators. Each potential data source has advantages and disadvantages, and we are explicit about these.

117. One important consideration for quantitative data sources is comparability over time. Factors of maturation, learning, or history may intervene between measurement intervals that affect the comparison of differences at any two points in time. For some indicators measurement reliability may not be problematic, but for others based on subjective estimates, content analysis or classification pre-post comparisons may prove misleading and small effects may be hard to detect. If third party data are used, it will be extremely important to check for each data set that the same operational definitions are in place and the same collection methods used at each of the multiple points in the time series.

118. Qualitative data sources such as interviews and content analysis of documents may provide rich, nuanced and potentially heuristic sets of supplemental data to assist in identifying comparison groups. As Shadish, Cook, and Campbell (2002) note, qualitative and quantitative methods complement and enrich one another when applied within a common set of epistemological assumptions. For example, interviews prior to implementation of an intervention may not only enlist recommendations on counterfactual candidates from credible local informants, but may help to: i) discover important mediators and moderators that were not initially considered as part of the causal chain; ii) identify additional molecular components of intermediate outcome indicators that reveal themselves only over time; and iii) expose ways in which effects are likely to be highly specific to the local circumstances. Interviews in conjunction with survey methods and social network analysis may be useful in determining whether or not treatment diffusion is an issue, by examining the local and regional networks of scientists and students participating in an intervention to see if they have shared its benefits with peers outside of the study setting. If there are knowledge or research spillovers, some of the obvious comparison candidates might not be valid.

119. For qualitative data sources, the most important data challenge will be to minimize respondent and evaluator biases. Doing so requires a careful planning of interviews to ensure adequate stakeholder mapping and sampling as well as adequate preparation, recording and documentation (White and Phillips 2012).

Qualitative data

120. Interviews with key stakeholders may be useful in determining if there are any threats to internal validity of a proposed intervention, such as unplanned events occurring between the intervention and outcome measurement, or threats to construct validity, such as intervention demand characteristics or novelty effects (Shadish, Cook, and Campbell, 2002). Review of documents related to the study protocol, including the proposal, CVs of personnel, consent forms, and instrumentation, along with interim written materials such as progress reports and email correspondence, where available, can provide useful information for understanding the implementation process or for documenting whether key elements or outcomes in the intervention's theory of change have in fact occurred.

121. For some projects there may be related online documents including web pages, video and audio

materials, and legacy and social media comment. Materials for content analysis do not have to be naturally-occurring; diaries can be solicited from program participants (Davies & Dart, 2005), or before-after narratives from knowledgeable observers. Established category schemes based on attribution theory are available for classifying explanatory discourse, as well as themes drawn from theories of change; these should prove useful in understanding how stakeholders characterize and account for relationships between interventions and outcomes. Integrated solutions for capturing, organizing and analyzing qualitative data (for example, visualizing semantic networks) are available through tools like NVivo and DiscoverText. As with all content analyses, immersion in a subset of the data corpus and relevant literature is required to develop appropriate procedures for unitizing and categorizing the data, and for testing those categories against as yet unexamined examples. A second data coder to resolve inconsistencies and establish reliability is desirable.

122. Where translation of documents is required prior to coding, document review may be affected by a lack of semantic and conceptual equivalence across languages. Further problems may develop if translation is required to facilitate interviews, as normative expectations for the interviewee and interviewer roles may differ considerably between the parties, with cultural-level variation in self-presentational concerns including depth and breadth of disclosure (Behling & Law, 2000).

Administrative data

123. The next possible data source for indicators concerning program features, student body and academic staff would be university-level administrative data. Administrative data may prove very useful for measuring indicators such as number of new students and graduates, recruitment of staff with Masters or PhD qualifications, qualified teacher to student ratios, books, laboratories and computers, among others. There are many advantages to using administrative data. These include: i) data are potentially comprehensive in that they potentially contain information on many proposed indicators; ii) data are likely to be comparable over time and across units within a given university; iii) response rates for units within a university are usually close to universal; iv) collection costs are minimal in that administrative data leverage existing organizational protocols for reporting; and v) there is little incentive for differential reporting or misreporting across units in response to NORHED activities.

124. However, administrative data also have some disadvantages. These include: i) data might not be available for some critical indicators; ii) data may not be disaggregated to the relevant organizational units; iii) data might simply not be available in many settings due to poor organizational protocols; iv) even if they exist, data may not be comparable across universities and countries; v) reporting format and content may vary over time; and vi) administrative data are not a good source for some proposed qualitative indicators, such as the relevance of new curricula.

Survey of organizational units

125. To complement the strengths of administrative data and to circumvent some of their weaknesses, one could use surveys of organizational units (awarded funding and control). These surveys could be carried out by enumerators from the research team. For example, it is unlikely that university-level administrative data contain information on publications, grant proposals submitted and awarded, conference presentations or collaboration with other regional partners or with researchers from the North.

126. The main advantages of using a survey of organizational units are: i) questions can be tailored to collect data on specific indicators; ii) design of the survey can ensure comparability across units over time within and between universities; and iii) questions can combine quantitative and qualitative (e.g. perceptions) information.

127. There are, however, disadvantages to using surveys of organizational units. These include that: i) response rates are likely to be low overall; ii) response rates may differ between units awarded funding from NORHED and potential comparison units, given the potential differential incentives for reporting; iii) survey data collection is costly and time-consuming; iv) data quality is difficult to validate externally; v) data quality is heavily dependent on interviewee knowledge and willingness, and it is hard to ensure that the same administrative role responds to survey questions across units and over time; and vi) even if appropriately collected, organizational survey data will not suffice to provide all relevant information for the proposed indicators (e.g. average wages of graduates or industry-level skill gaps).

Surveys of graduates and tracer studies

128. It is unlikely that university-level administrative data or the survey of organizational units will contain information to measure many outcome and impact indicators. For example, these data sources are not likely to include the number of graduates finding jobs in relevant local labor markets, average wages of graduates, gender gaps in wages and time to first employment. For these reasons, one alternative data source would be surveys of graduates or tracer studies. The advantages are similar to those of surveys of organizational units. They allow tailoring of the survey to ensure comparability across units and over time and the possibility of combining quantitative and qualitative information. The main disadvantages are that response rates among these types of surveys are usually low, especially when the survey is not directly administered by the organization, and that data collection is costly and time-consuming.

129. Tracer studies can be very useful for educational institutions. These studies let institutions study their graduates' employment and determine whether or not the curriculum and training programs provided are equipping graduates with the necessary skills to be successful in the local job market. Tracer studies can then help institutions identify deficiencies in their programs and make appropriate revisions. Noko and Ngulube (2013) conducted tracer studies of records and archives management graduates in Zimbabwe. They note that such studies can also help academic units develop alumni networks that will be highly useful in identifying job opportunities for their graduates. Superior tracer studies will likely employ a combination of qualitative and quantitative methods (e.g. Mammo, 2007). Effective triangulation would involve survey data collected from graduates and key informants in the relevant industry, government and NGO sectors; interviews with program administrators; and document review including study of curricula, program descriptions, marketing materials, and recruitment postings in legacy and new media. Interactive features of career center web sites may be leveraged to collect anecdotal data from alumni posting in forums as well as to survey graduates and prospective employers on an ongoing basis. Analysis of traffic logs of career services web sites may provide useful data about what information graduates and employers are looking for (Renny et al., 2013).

130. Survey data of graduates, collected annually, may provide the most immediately useful feedback. However, institutions may not have reliable contact information on their alumni, and even among those who are successfully contacted, the response rate may be low (although response rates of 40% or better are not uncommon in the tracer study literature) (Noko and Ngulube, 2013). Low response rates raise issues of the representativeness of the respondents with respect to the graduate population as a whole. For example, are there systematic ways in which respondents and non-respondents differ with respect to their satisfaction with training and their career outcomes? Do graduates with whom the alumni or career office has lost contact differ in significant ways with respect to these key indicators? Careful attention to these issues can increase confidence in the validity of tracer survey findings.

Surveys of firms

131. None of the data sources described above is well suited to accurately measure skills gaps between the supply and the demand of graduates or the relevance of educational programs for local, regional and

national labor markets. To capture this kind of information, researchers often rely on survey data on firms. These surveys are suitable to elicit employer perceptions on the quality of graduates from different universities, whether the supply of graduates is adequate for current and future demands and whether the training is relevant for industry needs.

132. As with other forms of survey data, surveys of firms are costly and time-consuming. Moreover, it is often difficult to ensure that the relevant officer responds to questions, and it is not always clear who the relevant officer should be as in many cases this varies from employer to employer.

Newspaper job vacancy data

133. An alternative way to measure skills gaps is to collect newspaper data on vacancies by industry and to contrast those vacancy-data with data on numbers of graduates. The advantage of this approach is that newspaper data are a readily available source of data to measure the demand for skills by region, industry and occupation.

134. The main disadvantages, however, of using newspaper data include: i) it is costly to code these data; ii) job postings often do not contain all relevant information; iii) newspaper job postings might not be the natural sources of information on vacancies for university-trained prospective workers; and iv) it is unclear how job posting offers translate into actual positions filled.

135. In addition to job postings in legacy media (print newspapers), additional sources of information to be reviewed could include postings on the corporate or organizational web sites of representative employers from the industry sector, the university's own career center postings or listserv notices, job listings on regional employment networking sites; and postings in trade publications, both print and online.

National surveys

136. Researchers have often used already-collected nationally representative survey data to measure skill gaps by industry and occupation. The main advantages of using nationally representative survey data include: i) data are readily available, and usage is typically free of charge; ii) nationally representative survey data often contains numerous and detailed questions on employment and qualifications; iii) data usually exist for many countries including developing ones; and iv) surveys are done annually or biennially, which enables tracking indicators over time.

137. There are, however, disadvantages of using nationally representative survey data, especially for measuring skill gaps relative to NORHED programs. These include: i) national surveys do not typically identify which university workers graduate from or their major which makes it difficult to construct unit-specific skills gaps; and ii) national surveys are not always representative at the regional or local level.

Academic electronic databases

138. For indicators related to the quantity and quality of research, potentially fruitful data sources are academic electronic databases. These databases include citation indexes such as those in, Google Scholar, Social Science Research Network, EconLit, IDEAS, Eldis, Francis, Embase, Inter-Science, Thompson Reuters Web of Science, EBSCO and WHOLIS, Harzing's Publish or Perish, and WASS-Sense global rankings among others. The main advantages to using these data sources include: i) databases are usually comprehensive, up to date and have worldwide coverage; ii) sources identify authors and their university affiliations; and iii) data are typically available for free.

139. There are some disadvantages to using these databases, however, to track outcomes and impacts of

NORHED grants. These include: i) data coding is time-consuming; ii) it is not clear whether these databases are useful for tracking working papers in some academic fields; and iii) these databases are not useful for tracking research grant submissions, paper submissions or patents.

140. Information on research grant submissions, paper submissions or patents is often available from the contracts and grants offices of the individual universities, and patents from their technology transfer offices, if these exist. Patent filing data by country from the WIPO (World Intellectual Property Organization) database might be also a useful backdrop against which to compare self-reported patent activity by any given institution.

Academic staff websites

141. In the U.S. and Europe, it is common for researchers affiliated with universities to list their publications, working papers and grants on their websites or their CVs, which are usually available in personal websites. It might be possible to obtain, therefore, grant and working paper information from researcher websites. The main advantage of this data source is that information is free and directly accessible. The main disadvantage, however, is that the cultural norm of making publications, working papers and grants public through websites might be specific to Northern countries and not as commonplace in developing country-settings. An additional disadvantage is that the availability of data hinges on the availability of suitable technological platforms, which might not be available in many NORHED programs.

142. We conclude this data section by highlighting the importance of obtaining reliable data for all processes linking inputs, outputs and outcomes in the causal pathway to obtain robust conclusions about NORHED's impact. Many of the data sources that we propose—particularly those that we propose for measuring quantitative indicators—will likely be more relevant for measuring outcomes. In some settings, these data sources may also serve to measure outputs. In many cases, we will leverage the power of qualitative data sources such as those outlined above to measure input and output indicators. In all cases, however, we will ensure that we employ the right mix of data to measure all key indicators to properly identify hypothesized mechanisms linking cause and effect within the NORHED evaluation timeframe.

VII. Conclusions and Recommendations

Conclusions

143. The team developed a general theory of change in capacity development for higher education based on international research findings. It also explicated the implicit theory of change underlying the NORHED program, based on reviews of NORHED documents and the logical frameworks that were constructed for the overall program and for all 46 projects.

144. Comparisons of the general and NORHED-implicit theories of change indicate that the theory of change implicit in the NORHED program and projects is broadly consistent with the general theory of change. As a collective program, NORHED has most of the necessary conditions in place for becoming a successful program, provided that the enabling conditions are met.

145. However, the program is most vulnerable in terms of the enabling conditions that grantees require to use NORHED funding effectively. Individual projects are implementing in highly variable local conditions in a large number of lower capacity countries that span three global regions. It is only realistic to expect that at least some, if not most, participating institutions will fail to meet one or more of the enabling conditions key to success, such as a good governance environment, adequate financing, or meritocratic, transparent and fair staff recruitment procedures. Thus, the individual projects may be at risk of implementation problems that can affect the likelihood that they can achieve their objectives.

146. Reform at the higher education level is complex, and NORHED seeks reform in difficult contexts. The inescapable conclusion is that NORHED and the grantees must constantly monitor inputs and outputs and track progress toward the projects' intended objectives. This investment, if acted on, can yield huge benefits. It catches projects that are failing in time to be able to get them back on track.

147. A thorough analysis of the options for attributing outcomes to the NORHED program is summarized in a decision tree on page 34 that can be used to guide the choice of approach. The most robust approaches to establishing attribution will probably not be available for the NORHED program.

148. All options for determining the effects of the NORHED program and being able to attribute these with even modest confidence to the NORHED program require baseline data for the outcomes sought. They also require the systematic monitoring of inputs and outputs and the tracking of progress toward achieving the intended outcomes during implementation.

149. The team concluded that it was impossible to overstate the importance of measuring inputs, outputs, and outcomes before implementation (baseline data) and during implementation. Without baseline data and monitoring and evaluation during implementation, Norad will not know what was achieved, whether any observed achievements can be attributed to the NORHED program, whether a project is failing when there is still time to get it back on track, or why a project might be succeeding or failing.

150. Norad's recent and thoughtful report (*Can We Demonstrate the Difference that Norwegian Aid Makes?*) found that "none of the reports on grants that were evaluated could reach firm conclusions about the results being achieved. Reports showed well what money was being spent and what direct activities or services were being delivered. But critical questions about whether those services gave rise to real benefits for poor people and other target groups proved elusive." (p.xv)

151. NORHED and Norad will not be able to learn which aspects of the program worked or did not work and why unless projects have baseline data and systematic monitoring and evaluation of implementation

progress. The NORHED program will become just one more Norad program for which firm conclusions about the results being achieved cannot be determined.

152. Tables 2 and 3 in the report identify the causal pathways (inputs, outputs, outcomes, and impacts) and indicators for measuring the production of more and better graduates and the production of more and better research. The report also identifies a variety of data sources for measuring these indicators and discusses the advantages and disadvantages of each--for example, the likelihood of bias, the ease and cost of data collection.

Recommendations

153. All recommendations are relevant to Norad's Evaluation Department and NORHED's program management, but recommendations 1 and 5 are especially relevant to the Evaluation Department; recommendations 2, 3, and 4, to NORHED's program management.

1. Use the general theory of change as the theoretical framework for evaluating the NORHED program. Although NORHED's design is implicitly and broadly consistent with the general theory of change, the general theory of change has the advantage of being closely linked to international research findings that expand the options for interpreting data from assessments of NORHED's investments.
2. Ensure that all projects:
 - a. Have adequate baseline data for inputs, outputs, and outcomes; and
 - b. Systematically monitor inputs and outputs and track progress toward objectives during implementation.
3. As part of the measurement initiative, a limited set of outcome indicators needs to be selected that can be used across the projects. The indicators selected should be measurable with data that are commonly available across grantees, easy and reasonably inexpensive to collect, and minimally subject to bias and reliability problems.
4. Monitor the status and effects of the enabling conditions on how well grantees can use NORHED-financed inputs to achieve their objectives. As noted, NORHED has most of the necessary conditions in place for becoming a successful program if the enabling conditions are met. However, the NORHED program is operating in environments where these enabling conditions are at risk. If the enabling conditions are often found to be absent and crippling to the best efforts of grantees, it may be advisable to adjust the NORHED program design by collaborating with national and international partners to take concerted action to get the necessary conditions in place. In the meantime, NORHED can focus attention on these conditions through dialogue with national and university-level leaders.
5. Select the approach to be used for assessing the effectiveness of the NORHED program. DPMG recommends that Norad draw interim conclusions about the effectiveness of the NORHED program by evaluating a sample of projects that have been implementing for a sufficient number of months to allow progress and problems to surface. In the spirit of the formative evaluation tradition, DPMG also advises that projects be evaluated a sufficient number of months prior to closing to give grantees and NORHED a chance to get failing projects them back on track. DPMG understands that Norad may wish to separate the evaluation of the NORHED program from its operation. However, it has found that the formative model adds more value in terms of salvaged projects and in terms of learning by program staff and grantees.

References

- Abadie, A. (2005) 'Semi parametric difference-in-differences estimators', *Review of Economic Studies* 72, 1-19.
- Abadie, A., Diamond, A., and Hainmueller, J. (2010) 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program', *Journal of the American Statistical Association*, 105(490). pp. 493-505.
- Barth, M. (2013). 'Many roads lead to sustainability: A process-oriented analysis of change in higher education', *International Journal of Sustainability in Higher Education*, 14(2), 160–175.
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks, CA: Sage.
- Altbach, P., and Salmi, J. (2011) *The Road to Academic Excellence: the Making of World-Class Research Universities*. Washington, D.C.: World Bank.
- Angrist, J. and Pischke, J.S. (2009). *Mostly harmless econometrics*. Princeton, N.J., Princeton University Press.
- Baser, H., and Morgan, P. (2008) *Capacity, Change and Performance*, European Centre for Development Policy Management (ECDPM), Working Paper No. 59B.
- Becker, W. E., and Kennedy, P. E. (2005) 'Does Teaching Enhance Research in Economics?', *Perspectives on Research and Teaching in Economics*, 95(2). pp. 172–176.
- Boeren, A., and Holtland G. (eds.) (2005) *A Changing Landscape: Making Support to Higher Education and Research in Developing Countries More Effective*. The Hague, Netherlands: Nuffic.
- Boyer, E. L. (1990). *Scholarship Reconsidered: Priorities of the Professoriate*. New Jersey: Carnegie Foundation for the Advancement of Teaching.
- Clark, B.R. (1995). *Places of Inquiry: Research and Advanced Education in Modern Universities*. Berkeley: University of California Press.
- Cloete, N., and Bunting I. (2013) *Strengthening Knowledge Production in Universities: Five South African Case Studies*. Paris: OECD.
- Cloete, N., Fehnel, R., Maassen, P., Moja T., Perold H., and T. Gibbon (eds.) (2002). *Transformation in Higher Education: Global Pressures and Local Realities in South Africa*. Lansdowne: Juta and Company.
- Davies, R., & Dart, J. (2005). *The 'most significant change' (MSC) technique. A guide to its use*. On-line resource.
- Douglass, J.A., Thomson, G., and Zhao, C. M. (2012). 'The Holy Grail of Learning Outcomes', *University World News* [Online]. 4 March 2012. Available from:

<http://www.universityworldnews.com/article.php?story=20120301093145248&query=AHELO>.
[Accessed 4 March 2014].

European Centre for Development Policy Management (ECDPM). (2008). *Capacity Change and Performance Insights and Implications for Development Cooperation*, European Centre for Development Policy Management, Working Paper No. 21.

Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences*. New York: Basic.

Gauthier, E. (1998). *Bibliometric Analysis of Scientific and Technological Research: a User's Guide to the Methodology*. Ottawa: Statistics Canada.

Halliwell, J. (2008). *The Nexus of Teaching and Research: Evidence and Insights from the Literature*. Toronto: Higher Education Quality Council of Ontario.

Hanushek, E. A. and Woessmann L. (2009). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation*, National Bureau of Economic Research, Working Paper 14633. Cambridge, Massachusetts: National Bureau of Economic Research.

Hinrichs P. (2012). 'The Effect of Affirmative Action Bans on College Enrollment, Educational Attainment and the Demographic Composition of Universities', *Review of Economics and Statistics*, 94(3). pp. 712-722.

Hoidn, S. and Kärkkäinen K. (2014). *Promoting Skills for Innovation in Higher Education: A Literature Review on the Effectiveness of Problem-Based Learning and of Teaching Behaviours*, Education Working Paper No. 100. Paris: OECD.

Kim, H. and Lalancette D. (2013). *Literature Review on the Value-Added Measurement in Higher Education*. Paris: OECD.

Knobel, M. (2014). "Brazil's shortage of professors", *Inside Higher Education*. 1 June 2014.
<http://www.insidehighered.com/blogs/world-view/brazils-shortage-professors#sthash.pkqx8Bg2.T8VYovn0.dpbs>

Kotecha, P. (2012). *Higher Education in the Southern African Region: Current trends, challenges, and recommendations*. Johannesburg, South Africa: Southern African Regional Universities Association (SARUA).

Lindow, M. (2009, February 20). 'Africa's New Crisis: a Dearth of Professors', *Chronicle of Higher Education*, A27.

Mammo, W. (2007). 'Demise, Renaissance or Existence of LIS Education in Ethiopia: Curriculum, Employers' Expectations and Professionals' Dreams', *International Information and Library Review*, 39 (1). pp. 145-57.

McKensie, K. and Schweitzer R. (2001). 'Who Succeeds at University? Factors Predicting Academic Performance in First Year Australian University Students', *Higher Education Research & Development*, 20. pp. 21-33.

McKinsey and Company. (2010). *How the World's Most Improved School Systems Keep Getting Better*. [Online] Available from: http://mckinseysociety.com/downloads/reports/Education/How-the-Worlds-Most-Improved-School-Systems-Keep-Getting-Better_Download-version_Final.pdf

Miller, R. K. (2007). *Beyond Study Abroad: Preparing Engineers for the New Global Economy*, Archived Speeches and White Papers, Olin College of Engineering. Available from: http://www.olin.edu/about/presidents_office/archive.

Morgan, P. (2006). *The Concept of Capacity*. Maastricht, The Netherlands: European Centre for Development Policy Management (ECDPM).

Mouton, J. (2007). 'Post-Graduate Studies in South Africa: Myths, Misconceptions and Challenges', *South African Journal of Higher Education*, 21(8). pp. 1078–1090.

Nickson, S. C. (2012). 'Developing faculty for the 21st century in South Africa: Building capacity through collaborative preparation programs.' In J. E. Groccia, M. A. Al-Sudairy, & W. Buskist (Eds.), *Handbook of college and university teaching: A global perspective* (pp. 108–124). Thousand Oaks, CA: SAGE.

Noko, P., and Ngulube, P. (2013). 'A Vital Feedback Loop in Educating and Training Archival Professionals: A Tracer Study of Records and Archives Management Graduates in Zimbabwe', *Information Development*. DOI: 10.1177/0266666913510308

Norad (2009). *Evaluation of the Norwegian Program for Development, Research and Education (NUFU) and of Norad's Program for Master Studies (NOMA)*. Oslo: Norad.

Nuffic (2012). *Final report evaluation NPT and NICHE*. [Online] Available from: <http://www.nuffic.nl/en/library/final-report-evaluation-of-npt-and-niche.zip/view>.

Nuffic (2013). *Integrated Capacity Development*. [Online] Available from: www.nuffic.nl/en/capacity-building/niche/cross-cutting-themes/integrated-capacity-development. Accessed: March 23, 2014.

OECD (2013). *PISA 2012 Results: Creative Problem Solving Students' Skills in Tackling Real-Life Problems Volume V*. Paris: OECD.

Olsson, A., Meek, L., and Cooke, N. (2013). *Research and Innovation Management*, Program on Innovation, Higher Education and Research for Development (IHERD) Policy Brief. Paris: OECD.

Qualters, D. (1995). 'A Quantum Leap in Faculty Development: Beyond Reflective Practice', *To Improve the Academy*, Paper 341. [Online] Available from: <http://digitalcommons.unl.edu/podimproveacad/341>

Qualters, D. (2009). 'Creating a Pathway for Teacher Change', *The Journal of Faculty Development*, 23(1). pp. 5-13.

Renny, Chandra R., Ruhama, S., & Sarjono, M. W. (2013). 'Exploring Tracer Study Service in Career Center Web Site of Indonesia Higher Education', *(IJCSIS) International Journal of Computer Science and Information Security*, 11(3). pp. 36-39.

- Rogers, P.J. (2008). 'Using Program Theory for Complicated and Complex Programs', *Evaluation*, 14(1). pp. 29-48.
- Romer P. (1990). 'Endogenous Technological Change', *Journal of Political Economy*, 98. pp. 71-102.
- Salmi, J. (2011). 'The Road to Academic Excellence: Lessons of Experience.' In Altbach, P. and Salmi J. (eds). (2011). *The Road to Academic Excellence: the Making of World-Class Research Universities*. Washington, D.C.: World Bank.
- Salmi, J. (2010). 'Ensuring the Successful Implementation of Financing Reforms', *Financing Higher Education in Africa*. Washington DC: The World Bank, pp. 157-170.
- Salmi, J., and Hauptman A. M. (2006). *Innovations in Tertiary Education Financing: A Comparative Evaluation of Allocation Mechanisms*, Education Working Paper Series Number 4. Washington, D.C.: World Bank.
- Salmi, J., Kärkkäinen, K., Mukherjee H., and Ushihakala J. (Forthcoming). *Evaluation of Finland's Support to Higher Education Institutions: NSS and HEI ICI Programs*. Helsinki: Ministry for Foreign Affairs.
- Scott, I. (2009). 'Towards an Agenda for SoTL in Africa?', *International Journal for the Scholarship of Teaching & Learning*, 3(1). pp. 1-8.
- Sethi, G., Andharia, J., and Nupur. (2012). *Understanding Capacity-building Needs: Current Models for Excluded Communities*. New Delhi, India: UNICEF India.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin Co.
- Shongwe, M., and Ocholla, D. N. (2011). 'A Tracer Study of LIS Graduates at the University of Zululand, 2000-2009', *Mousaion*, 29(2). pp. 227-245.
- Simister, N., and Smith, R. (2010). *Monitoring and Evaluating Capacity Building: Is it really that difficult?* International NGO Training and Research Centre (INTRC) Working Paper No. 23.
- Thomson Reuters. (2008). *Using Bibliometrics: a Guide to Evaluating Research Performance with Citation Data*. New York City: Thomson Reuters.
- Ubels, J., Acquaye-Baddoo, N.-A., and Fowler, A. (eds.). (2010). *Capacity Development in Practice*. London, England: Earthscan. [Online] Available from: <http://www.snvworld.org/en/publications/capacity-development-in-practice-complete-publication>.
- Van Deuren, R. (2013). *Capacity Development in Higher Education Institutions in Developing Countries*, Maastricht School of Management, Working Paper No. 2013/30. [Online] Available from: <http://ideas.repec.org/p/msm/wpaper/2013-30.html>.

- Vogel, I. (2012). *Review of the Use of 'Theory of Change' in International Development*. London: UK Department of International Development.
- Wang, Y. (2012). *Education in a Changing World : Flexibility, Skills, and Employability*, Education Working Paper No. 69104. Washington, D.C.: World Bank.
- White, H. (2011). *Theory-based impact evaluation: principles and practice*. International Initiative for Impact Evaluation, Working Paper # 3.
- White, H. and Phillips, D. (2012). *Addressing attribution of cause and effect in small n impact evaluations: towards and integrated framework*. International Initiative for Impact Evaluation, Working Paper # 15.
- World Bank. (2002). *Constructing Knowledge Societies: New Challenges for Tertiary Education*. Washington DC: the World Bank.
- World Bank. (2012). *Putting Higher Education to Work: Skills and Research for Growth in East Asia*. Washington, D.C.: World Bank.
- Yizengaw, T. (2008). *Challenges Facing Higher Education in Africa & Lessons from Experience: A Synthesis Report for the Africa – US Higher Education Initiative*. Washington, D.C.: African-U.S. Higher Education Initiative.
- Yusuf, S. and Nabeshima K. (2007). *How Universities Promote Economic Growth*. Washington, D.C.: World Bank.
- Zeegers, P. (2004). 'Student learning in higher education: A path analysis of academic achievement in science', *Higher Education Research & Development*, 23(1). pp. 35-56.

Preparatory study:
Development of a theory of change and suggesting methods for evaluating
NORHED (call-off A1)

Terms of References (15 January 2014)

1. Background

On the 04th of November 2013 Norad's Evaluation Department awarded the framework agreement for a series of evaluation assignments of the *Norwegian Program for Capacity Development in Higher Education* (NORHED) to the Development Portfolio Management Group at the University of Southern California. This is the first call-off under the framework agreement, an agreement which took effect on the 22nd of November 2013.

NORHED is operated by the Norwegian Agency for Development Cooperation (Norad) and has an annual budget of approximately 130 million NOK. The objective of NORHED is to contribute to development in the recipient countries by strengthening capacity in higher education institutions. *“Strengthening of higher education institutions refers to: 1. Producing more and better research relevant to the identified areas/sub-programs [and] 2. Producing more and better qualified graduates, men and women, relevant to the identified areas/sub-programs.”* (Norad 2013:4).

In the first call-off NORHED received 173 applications⁷ for project support, out of which 46 will receive funding. The projects are owned by 26 higher education institutions from 16 low- and middle income countries⁸, and 10 higher education institutions from Norway ([Norad 2013](#)). The higher education institutions have been awarded project funding for 750 million NOK over five years. Most projects are situated in Uganda, Ethiopia, Malawi and South Sudan. NORHED has developed an overall results framework and all project applicants have developed independent results framework in accordance with the overarching framework. The majority of institutions are required to submit a revised version of this framework by the 31st of March 2014. The first disbursements are made in December 2013 and the first projects are expected to start in January 2014.

2. Purpose and scope

This preparatory study⁹ is a first in a line of evaluation-assignments undertaken to learn about what works and why in NORHED's support to capacity development of higher education institutions in low- and medium income countries. The purpose of this particular study is to contribute to a discussion with the NORHED administration and other stakeholders about how the program can make a difference, and to

⁷ As stated in the TOR for the framework agreement: all applicants have accepted to be part of an evaluation, including those who did not receive funding.

⁸ 11 of these countries are located in Sub-Saharan Africa.

⁹ The overall purpose of the framework agreement is *“to enable learning by relevant stakeholders within the field higher education and development, so that future investments can become more effective in building capacity in higher education institutions as a contribution to development.”* Terms of References for the framework agreement, page 1.

provide a basis for future evaluations of NORHED strategies. This will be done through the development of a testable theoretical framework, including a proposal of methods and indicators, with input from, and verified through stakeholder consultation. It is expected that a discussion of NORHED strategies will both contribute to learning among stakeholders and ensure that future evaluation assignments are relevant.

The focus of this call-off will be on the effects of capacity development at the level of institutions in the countries where projects are based. With respect to time horizon, the theoretical framework shall take into account NORHED's long term perspective, while being explicit on what is testable in the short term, i.e. a three year period.

3. Objectives

Develop a theoretical framework (theory of change) to test capacity development

The first objective of this call-off is to develop a testable theoretical framework in adherence with the principles of a theory-based impact evaluation, discussed below and laid out, for example in White (2009).

A theory of change is a theoretical framework that maps out the causal chain from inputs to outputs, outcomes and impacts (White 2009:274-275). In simple terms, the theory of change explains what needs to occur to go from A (intervention) to Z (objectives), and what the underlying assumptions are for this to occur. Many of these causal links are already embodied in the NORHED program documents (either as program theory or in a program log frame), with underlying assumptions discussed as risks.

Nevertheless, the benefit of getting such a theory spelled out by external higher education experts with substantial contextual experience is that key assumptions in the underlying theory of change can be externally verified. Furthermore, a clearly stated theory of change will allow for a comparison of NORHED strategies with existing literature on capacity building of higher education institutions and it may allow for empirical testing of some assumptions.

The theoretical framework shall be developed to answer the following evaluation questions:

To what extent and how does NORHED build capacity in higher education institutions in terms of “*more and better research [and] (...) more and better qualified graduates*” (Norad 2013:4), in comparison to the counterfactuals:

- No NORHED-funding
- Other types of donor funding to higher education institutions

In addition the preparatory study shall discuss whether it is possible and useful to also assess NORHED funding in comparison to

- NUFU and NOMA funding (NORHED's predecessors)
- No-funding

The theoretical framework should develop a model for the entire results chain for NORHED, including the funding mechanism, and the collaboration between institutions beyond financial support. The model should take into account the importance of the context within which NORHED operates; however, since this pre-

study does not involve empirical research beyond document reviews and some interviews, at this stage the context will not be described in any detail. The theoretical framework should include intermediate factors that need to be present for capacity development to take place. For example, if a concentration of talent (researchers, students and internationalization) is important to produce highly qualified graduates, then the framework should map out how NORHED funding can help higher education institution achieve this (as suggested by Altbach and Salmi (2011)¹⁰). Similarly, if good governance in higher education institutions is a pre-condition, the framework should discuss whether and how NORHED can achieve this, and so on. Other intermediate factors may also be equally important. In addition, the discussion accompanying a theoretical framework should discuss the evidence base of current NORHED strategies, and comparing these with other capacity building strategies that are not part of the NORHED program. It is essential that underlying assumptions crucial for the program's success are made explicit. To the extent that multiple theoretical approaches to capacity development exist, the team should justify its choice.

While NORHED's program staff and documents should be consulted, it is the responsibility of the evaluation team and not the NORHED administration to ensure that the theoretical framework is developed. Furthermore, the theoretical framework developed in this preparatory study may deviate from NORHED's own program theory, in which case the discussion should explain why.

The theoretical framework developed under this call-off is likely to form the basis for future evaluations of the NORHED program.

Propose indicators

Based on this theoretical framework the evaluation team shall propose indicators which can be used to test how and why NORHED has/does not have an effect on the production of graduates and research. These indicators shall be firmly rooted in the theoretical framework for capacity development. The indicators may, but need not be similar to indicators developed by the project owners themselves as part of NORHED's own results framework. One reason for such a discrepancy is that the evaluation team should measure all institutions on the same indicators, while higher education institutions are allowed to develop separate indicators. Indicators along the entire results' chain should be proposed.

Discuss and propose methods

The team shall discuss how to capture a change in these indicators and propose methods to investigate whether changes can be attributed to the program using quasi-experimental methods, in line with the International Initiative for Impact Evaluation (3ie)-principles for a theory-based impact evaluation (White 2009)¹¹. All suggestions shall be well-founded and the extent to which a change can be detected within a period of 3 years or earlier shall be discussed. The team shall discuss whether the selection mechanism for funding can be utilized to create a comparison group for institutions that have received funding. The purpose of using a comparison group would be to attribute changes to the program.

¹⁰ As specified for the TOR in the framework agreement, this is an example, and other equally relevant approaches may exist.

¹¹ See also guidelines and policies: <http://www.3ieimpact.org/en/about/strategy-documents-reports/> .

Acknowledging that a large-n impact evaluation may not be feasible due to the size of the sample and heterogeneity of institutions, the team shall propose a small-n impact evaluation strategy. A small-n impact evaluation shall include an explicit strategy to address attribution. See for instance White and Philips (2009) for a discussion of challenges related to assessing attribution in a small-n impact evaluation.

While attributing a change in indicators to NORHED is important, explaining why changes occur or fail to occur is essential for learning. Qualitative methods, such as focus group discussions, in-depth interviews or similar may be essential to explain how and why changes in indicators occur.

Presentation of an early version of the theoretical framework in Addis on the 14th of March¹²

To present the evaluation and allow for stakeholder input, one or two representatives from the evaluation team shall present an early version of the theoretical framework in Addis in Ethiopia on March the 14th. The presentation will be held during a workshop organised by the NORHED administration.

Verification of theory of change at a stakeholder workshop in Oslo

One or two representatives from the evaluation team shall present the draft preparatory study that contains the theory of change and proposed indicators at a workshop in Oslo towards the end of April. The workshop will be organized by the evaluation department. The evaluation department will invite the NORHED-administration to this meeting, and may also invite other stakeholders. The workshop shall facilitate a discussion of the proposed theoretical framework and shall allow for a verification of the developed theoretical framework and indicators. The evaluation team shall consider and respond to proposals of changes or additions to the framework.

4. Methods and competence

It is expected that qualitative methods such as interviews with the NORHED and other stakeholders, literature and document review, are the most important methods employed in this preparatory study. However, knowledge of quasi-experimental methods is crucial and must be covered by the proposed team as the preparatory study shall discuss these methods, even though analysis is not expected at this stage.

The team should at least cover the following competencies:

- Capacity building of higher education institutions
- Experience in using qualitative methods
- Experience with quasi-experimental methods

5. Deliverables and timeline

- a) A plan for how the task should be solved (approach and methodology), including a budget specifying who should do what and when, should be submitted by e-mail, at the latest by the 22nd

¹² Provided that NORHED projects documents are accessible early enough for the evaluation team.

of January 2014. The budget should specify all costs, including travel. The plan and budget are subject to the approval of the Evaluation Department.

- b) A report of maximum 10 000 words, excluding annexes, responding to these Terms of References, and prepared in accordance with Annex 4 to the framework agreement. The report is subject to the Evaluation Department's approval. The development of the report has several sub-deliverables:
 - a. Oral presentation of the theoretical framework and if possible also indicators at a stakeholder workshop in Addis Ababa in Ethiopia on the 14th of March 2014. One or two researchers who have been central in writing the preparatory study should attend. The content of this presentation will depend on how far the evaluation team has come in the development of the theoretical framework.
 - b. Presentation of a written draft of the preparatory study at a workshop in Oslo in late April 2014. One or two researchers who have been central in writing the preparatory study should attend. A draft shall be handed in for consultation with the Evaluation Department prior to the workshop. The Evaluation department will distribute the draft report to stakeholders electronically. The Evaluation Department may request that the team proposes a program for the workshop.
 - c. A final draft of the preparatory study should be handed in, not later than the 20th of May, 2014. The preparatory study will be published online when approved by the Evaluation Department.
- c) A four page brief laying out the theory of change, suggested methods and headline indicators should be developed after the workshop in Oslo and handed in together with the final draft on the 20th of May. The target group for this brief is higher education institutions.

General comment to deliverables: The Evaluation department shall be consulted throughout the process and shall be alerted to any important deviance from the work plan.

Timeframe

The time frame for the first call-off is eleven person-weeks in total.

6. Budget

The total budget for this first call-off is maximum 665 000 NOK, which includes all expenses related to the call-off, including hours and travel budget. Once the budget is agreed upon and approved, expenses exceeding the budget will not be covered.

References:

Altbach, P. G. and J. Salmi (2011). The Road to Academic Excellence: The Making of World-Class Research Universities, World Bank Publications.

Norad (2013). A presentation of NORHED: The Norwegian Program for Capacity Development in Higher Education and Research for Development. e. a. r. Department for global health. Oslo
Norad

Norad (2013). "46 projects on higher education and research to receive funding." NORHED
Retrieved 11.12.2013, 2013, from <http://www.norad.no/en/support/norhed/46-projects-on-higher-education-and-research-to-receive-funding>.

White, H. (2009). "Theory-based impact evaluation: principles and practice." Journal of development effectiveness 1(3): 271-284.

NORHED's Reconstituted LogFrame

NORHED Strategic LogFrame	Indicators	Monitoring (who)	Assumptions (which)
Norway contributes to sustainable development	<ul style="list-style-type: none"> • Gender equality and human rights improvements • Reduced corruption in partner countries • Mitigation of Climate Change • Sustainable economic and environmental development 	<ul style="list-style-type: none"> • OECD • Statistics Norway • Eurostat 	<ul style="list-style-type: none"> • UN rules respected • Good governance and democracy
Mission (Development Objective): NORHED contributes to evidence-based policies and decisions that enhance sustainable economic, social and environmental development		<ul style="list-style-type: none"> • Norad • Evaluation Dpt. • Embassies of Norway 	<ul style="list-style-type: none"> • Graduates employed in priority economic sectors • WHO health indices improved
Vision (Impact): NORHED builds capacity in higher education and research in selected low and middle income countries and in selected economic and societal sectors	<ul style="list-style-type: none"> • Research environment <ul style="list-style-type: none"> ○ Empowerment of individual researchers in terms of freedom of choice for research objectives and of publication strategy ○ Balanced diversity in staff ○ Symmetry in international research networks, percentage of co-publication, PI status etc. • Study environment <ul style="list-style-type: none"> ○ Improved staff student ratios ○ Improved gender balance in teaching staff ○ Student safety on campus and in student hostels ○ Availability of health services and food program for students ○ Availability of internships and social programs • Internationalization <ul style="list-style-type: none"> ○ Diversity of student body ○ Percentage of foreign graduate students • Staff development <ul style="list-style-type: none"> ○ Percentage of graduate degree holders ○ Merit based career structure 	<ul style="list-style-type: none"> • Norad • Participating universities • Embassies of Norway • Ministries of partner countries 	<p>Internal assumptions:</p> <ul style="list-style-type: none"> • Staff recruitment procedures are meritocratic, transparent and fair • Gender and diversity sensitive rules and regulations • Rigor in research methodology is kept at international standard • Staff motivated and supportive institutional environment <p>External assumptions:</p> <ul style="list-style-type: none"> • Harassment of staff and students does not take place • Freedom of speech and other human rights are respected • Supportive and enabling institutional environment

NORHED Strategic LogFrame	Indicators	Monitoring (who)	Assumptions (which)
	<ul style="list-style-type: none"> ○ Tenure track and full tenure for qualified staff ● Institutions have procedures for developing their profile (strategy), based on lessons learned from NORHED and similar projects ● Institutional good governance 		
<p>NORHED's goals (2018):</p> <ul style="list-style-type: none"> ● Improved human capital (capacity) (more and better qualified male and female graduates) in six priority areas/sub-programs: (i) Education and Training, (ii) Health, (iii) Natural Resource Management, Climate Change and Environment, (iv) Democratic and Economic Governance, (v) Humanities, Culture, Media and Communication, and (vi) Capacity Development in South Sudan ● Strengthened research in the same six priority areas/sub-programs 	<ul style="list-style-type: none"> ● institutional governance (within participating units) <ul style="list-style-type: none"> ○ Meritocratic leadership recruitment in place, annual performance review ○ Autonomy and accountability. Resources used in accordance with purpose ○ Independent auditing system ● Quality and relevance of Graduates <ul style="list-style-type: none"> ○ Increased percentage of Study programs accredited ○ Graduation rates improved ○ Increasing percentage of graduates employed six month upon graduation ● Quality and relevance of Research <ul style="list-style-type: none"> ○ Higher publication rates ○ Increased demand for services from unit ○ More citations in local news media 	<ul style="list-style-type: none"> ● NORHED team ● Norad ● Evaluation Dept. 	<p>Internal Assumptions:</p> <ul style="list-style-type: none"> ● NORHED partnerships and networks are effective ● Research base contributes to educational and outreach programs ● Labour market or social demand for graduates in priority areas ● Zero tolerance for corruption under Norad's budget ● Conflict sensitivity <p>External Assumptions:</p> <ul style="list-style-type: none"> ● Sufficient institutional autonomy ● Climate and the environment taken into account ● Institutional strategy includes priority fields both in S&N partner institutions ● NORHED complements Norad's research support through Norway's research council ● NORHED complements Norad's master program in Petroleum and Energy at NTNU
<p>NORHED's actions</p> <p>Individual projects:</p> <ul style="list-style-type: none"> ● Establish a results framework, and adequate administrative procedures <ul style="list-style-type: none"> ○ Establish diversity balanced recruitment procedures ○ Fair selection procedures for fellowship holders ○ Knowledge management systems ○ Data collection protocols <p>Teaching:</p> <ul style="list-style-type: none"> ● Study program, academic level and structure <ul style="list-style-type: none"> ○ Master program design 	<p>NORHED's inputs:</p> <ul style="list-style-type: none"> ● Competitive selection mechanism <ul style="list-style-type: none"> ○ Expert panels recommends 46 out of 173 proposals ○ Results Frameworks established (0 – due end of March 2014) ● Contracts signed (36), specific budgets developed and first tranche of funds transferred by primo 2014 ● NORHED budget available ● Technical assistance to project development 	<ul style="list-style-type: none"> ● NORHED team ● Universities ● Project teams 	<p>Internal assumptions:</p> <ul style="list-style-type: none"> ● NORHED's program guidelines well developed, and builds on experience from NUFU and NOMA ● Transparent communication on program guidelines with target institutions in partner countries and Norway ● Adequate timeline for project submission, selection and contract development ● Sufficient administrative support for setting up lines of communication, procedures for transfer of funds, rules for transparent spending of funds

NORHED Strategic LogFrame	Indicators	Monitoring (who)	Assumptions (which)
<ul style="list-style-type: none"> ○ Sandwich-master programs ○ Joint degrees ● Study program development <ul style="list-style-type: none"> ○ Needs assessment ○ Teaching program design ○ Curriculum development ○ Selection of course material ○ Study program specific workshops for teachers ● Enrolment system and procedures <ul style="list-style-type: none"> ○ Recruitment of students and gender balanced student support ○ Enrolment of first student cohort Staff recruitment and development <ul style="list-style-type: none"> ● Good HR practices <ul style="list-style-type: none"> ○ Mainstreaming gender perspectives ○ Staff up-grading programs ● In-service teacher training programs <ul style="list-style-type: none"> ○ Short courses and training packages in general teaching methodologies ○ E-teaching and e-courseware development ● Formal academic qualification up-grading <ul style="list-style-type: none"> ○ Master and Ph.D. scholarships ○ Travel grants Joint Research <ul style="list-style-type: none"> ● Institutional cooperative actions <ul style="list-style-type: none"> ○ Formal network of institutions ● Research methods and tools <ul style="list-style-type: none"> ○ Comparative studies ○ Action research ○ Field work ○ Baseline surveys ○ Shared databases ● Research support <ul style="list-style-type: none"> ○ Supervision of younger researchers ○ Post doc grants ○ Travel grants ● Research outreach <ul style="list-style-type: none"> ○ Organization of research conferences and workshops ○ Research publication and dissemination Institutional capacity building 			<ul style="list-style-type: none"> ● Agreed salary structure, fellowships, tuition fees, bench fees, per diem etc. established ● Fellowship holders contributes to institutional development ● Best resource base in participation institutions are mobilized <p>External assumptions:</p> <ul style="list-style-type: none"> ● Institutional support ● Adequate infrastructure ● Sufficient core staff ● Personal safety for staff and students ● Quality Assurance Systems <ul style="list-style-type: none"> ○ QA systems for teaching capacity, curriculum, teaching material etc. ○ Accreditation ● Managerial and leadership capacity at institutional level <ul style="list-style-type: none"> ○ Learning Management Systems ○ Change program support ○ Gender sensitivity programs ○ Mentor programs ○ Management training and recruitment programs ○ Women's Leadership training ○ Ethics review board

NORHED Strategic LogFrame	Indicators	Monitoring (who)	Assumptions (which)
<ul style="list-style-type: none"> • Physical capacity <ul style="list-style-type: none"> ○ E-learning platforms ○ Video labs ○ Library facilities ○ Laboratories and community laboratories ○ Minor construction and equipment for training, research laboratories, and field stations Outreach and translational research <ul style="list-style-type: none"> • Building knowledge communities <ul style="list-style-type: none"> ○ Capacity building for external stakeholders ○ Business, society, academia network activities • Knowledge exchange <ul style="list-style-type: none"> ○ Translation of research findings into policy recommendations ○ Communication and dissemination of results • Student and graduate placement program Program review <ul style="list-style-type: none"> • Procedures for annual project performance reviews 			

EVALUATION REPORTS

- 10.00 Taken for Granted? An Evaluation of Norway's Special Grant for the Environment
- 1.01 Evaluation of the Norwegian Human Rights Fund
- 2.01 Economic Impacts on the Least Developed Countries of the Elimination of Import Tariffs on their Products
- 3.01 Evaluation of the Public Support to the Norwegian NGOs Working in Nicaragua 1994–1999
- 3A.01 Evaluación del Apoyo Público a las ONGs Noruegas que Trabajan en Nicaragua 1994–1999
- 4.01 The International Monetary Fund and the World Bank Cooperation on Poverty Reduction
- 5.01 Evaluation of Development Co-operation between Bangladesh and Norway, 1995–2000
- 6.01 Can democratisation prevent conflicts? Lessons from sub-Saharan Africa
- 7.01 Reconciliation Among Young People in the Balkans An Evaluation of the Post Pessimist Network
- 1.02 Evaluation of the Norwegian Resource Bank for Democracy and Human Rights (NORDEM)
- 2.02 Evaluation of the International Humanitarian Assistance of the Norwegian Red Cross
- 3.02 Evaluation of ACOPAM An ILO program for "Cooperative and Organizational Support to Grassroots Initiatives" in Western Africa 1978 – 1999
- 3A.02 Évaluation du programme ACOPAM Un programme du BIT sur l'« Appui associatif et coopératif aux Initiatives de Développement à la Base » en Afrique de l'Ouest de 1978 à 1999
- 4.02 Legal Aid Against the Odds Evaluation of the Civil Rights Project (CRP) of the Norwegian Refugee Council in former Yugoslavia
- 1.03 Evaluation of the Norwegian Investment Fund for Developing Countries (Norfund)
- 2.03 Evaluation of the Norwegian Education Trust Fund for African the World Bank
- 3.03 Evaluering av Bistandstorgets Evalueringsnettverk
- 1.04 Towards Strategic Framework for Peace-building: Getting Their Act Together. Overview Report of the Joint Utstein Study of the Peacebuilding.
- 2.04 Norwegian Peace-building policies: Lessons Learnt and Challenges Ahead
- 3.04 Evaluation of CESAR's activities in the Middle East Funded by Norway
- 4.04 Evaluering av ordningen med støtte gjennom paraplyorganisasjoner. Eksempifisert ved støtte til Norsk Misjons Bistandsnemda og Atlas-alliansen
- 5.04 Study of the impact of the work of FORUT in Sri Lanka: Building Civil Society
- 6.04 Study of the impact of the work of Save the Children Norway in Ethiopia: Building Civil Society
- 1.05 –Study: Study of the impact of the work of FORUT in Sri Lanka and Save the Children Norway in Ethiopia: Building Civil Society
- 1.05 –Evaluation: Evaluation of the Norad Fellowship Programme
- 2.05 –Evaluation: Women Can Do It – an evaluation of the WCIDI programme in the Western Balkans
- 3.05 Gender and Development – a review of evaluation report 1997–2004
- 4.05 Evaluation of the Framework Agreement between the Government of Norway and the United Nations Environment Programme (UNEP)
- 5.05 Evaluation of the "Strategy for Women and Gender Equality in Development Cooperation (1997–2005)"
- 1.06 Inter-Ministerial Cooperation. An Effective Model for Capacity Development?
- 2.06 Evaluation of Fredskorpset
- 1.06 – Synthesis Report: Lessons from Evaluations of Women and Gender Equality in Development Cooperation
- 1.07 Evaluation of the Norwegian Petroleum-Related Assistance
- 1.07 – Synteserapport: Humanitær innsats ved naturkatastrofer: En syntese av evalueringsfunn
- 1.07 – Study: The Norwegian International Effort against Female Genital Mutilation
- 2.07 Evaluation of Norwegian Power-related Assistance
- 2.07 – Study Development Cooperation through Norwegian NGOs in South America
- 3.07 Evaluation of the Effects of the using M-621 Cargo Trucks in Humanitarian Transport Operations
- 4.07 Evaluation of Norwegian Development Support to Zambia (1991 – 2005)
- 5.07 Evaluation of the Development Cooperation to Norwegian NGOs in Guatemala
- 1.08 Evaluation: Evaluation of the Norwegian Emergency Preparedness System (NOREPS)
- 1.08 Study: The challenge of Assessing Aid Impact: A review of Norwegian Evaluation Practise
- 1.08 Synthesis Study: On Best Practise and Innovative Approaches to Capacity Development in Low Income African Countries
- 2.08 Evaluation: Joint Evaluation of the Trust Fund for Environmentally and Socially Sustainable Development (TFESSD)
- 2.08 Synthesis Study: Cash Transfers Contributing to Social Protection: A Synthesis of Evaluation Findings
- 2.08 Study: Anti- Corruption Approaches. A Literature Review
- 3.08 Evaluation: Mid-term Evaluation of the EEA Grants
- 4.08 Evaluation: Evaluation of Norwegian HIV/AIDS Responses
- 5.08 Evaluation: Evaluation of the Norwegian Research and Development Activities in Conflict Prevention and Peace-building
- 6.08 Evaluation: Evaluation of Norwegian Development Cooperation in the Fisheries Sector
- 1.09 Evaluation: Joint Evaluation of Nepal's Education for All 2004-2009 Sector Programme
- 1.09 Study Report: Global Aid Architecture and the Health Millennium Development Goals
- 2.09 Evaluation: Mid-Term Evaluation of the Joint Donor Team in Juba, Sudan
- 2.09 Study Report: A synthesis of Evaluations of Environment Assistance by Multilateral Organisations
- 3.09 Evaluation: Evaluation of Norwegian Development Cooperation through Norwegian Non-Governmental Organisations in Northern Uganda (2003-2007)
- 3.09 Study Report: Evaluation of Norwegian Business-related Assistance Sri Lanka Case Study
- 4.09 Evaluation: Evaluation of Norwegian Support to the Protection of Cultural Heritage
- 4.09 Study Report: Norwegian Environmental Action Plan
- 5.09 Evaluation: Evaluation of Norwegian Support to Peacebuilding in Haiti 1998–2008
- 6.09 Evaluation: Evaluation of the Humanitarian Mine Action Activities of Norwegian People's Aid
- 7.09 Evaluation: Evaluation of the Norwegian Programme for Development, Research and Education (NUFU) and of Norad's Programme for Master Studies (NOMA)
- 1.10 Evaluation: Evaluation of the Norwegian Centre for Democracy Support 2002–2009
- 2.10 Synthesis Study: Support to Legislatures
- 3.10 Synthesis Main Report: Evaluation of Norwegian Business-related Assistance
- 4.10 Study: Evaluation of Norwegian Business-related Assistance South Africa Case Study
- 5.10 Study: Evaluation of Norwegian Business-related Assistance Bangladesh Case Study
- 6.10 Study: Evaluation of Norwegian Business-related Assistance Uganda Case Study
- 7.10 Evaluation: Evaluation of Norwegian Development Cooperation with the Western Balkans
- 8.10 Evaluation: Evaluation of Transparency International
- 9.10 Study: Evaluability Study of Partnership Initiatives
- 10.10 Evaluation: Democracy Support through the United Nations
- 11.10 Evaluation: Evaluation of the International Organization for Migration and its Efforts to Combat Human Trafficking
- 12.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative (NICFI)
- 13.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Brasil
- 14.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Democratic Republic of Congo
- 15.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Guyana
- 16.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Indonesia
- 17.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative. Country Report: Tanzania
- 18.10 Evaluation: Real-Time Evaluation of Norway's International Climate and Forest Initiative
- 1.11 Evaluation: Results of Development Cooperation through Norwegian NGO's in East Africa
- 2.11 Evaluation: Evaluation of Research on Norwegian Development Assistance
- 3.11 Evaluation: Evaluation of the Strategy for Norway's Culture and Sports Cooperation with Countries in the South
- 4.11 Study: Contextual Choices in Fighting Corruption: Lessons Learned
- 5.11 Pawns of Peace. Evaluation of Norwegian peace efforts in Sri Lanka, 1997-2009
- 6.11 Joint Evaluation of Support to Anti-Corruption Efforts, 2002-2009
- 7.11 Evaluation: Evaluation of Norwegian Development Cooperation to Promote Human Rights
- 8.11 Norway's Trade Related Assistance through Multilateral Organizations: A Synthesis Study
- 9.11 Activity-Based Financial Flows in UN System: A study of Select UN Organisations Volume 1 Synthesis Volume 2 Case Studies
- 10.11 Evaluation of Norwegian Health Sector Support to Botswana
- 1.12 Mainstreaming disability in the new development paradigm. Evaluation of Norwegian support to promote the rights of persons with disabilities.
- 2.12 Hunting for Per Diem. The uses and Abuses of Travel Compensation in Three Developing Countries
- 3.12 Evaluation of Norwegian Development Cooperation with Afghanistan 2001-2011
- 4.12 Evaluation of the Health Results Innovation Trust Fund
- 5.12 Real-Time Evaluation of Norway's International Climate and Forest Initiative. Lessons Learned from Support to Civil Society Organisations.
- 6.12 Facing the Resource Curse: Norway's Oil for Development Program
- 7.12 A Study of Monitoring and Evaluation in Six Norwegian Civil Society Organisations
- 8.12 Use of Evaluations in the Norwegian Development Cooperation System
- 9.12 Evaluation of Norway's Bilateral Agricultural Support to Food Security
- 1.13 A Framework for Analysing Participation in Development
- 2.13 Local Perceptions, Participation and Accountability in Malawi's Health Sector
- 3.13 Evaluation of the Norwegian India Partnership Initiative
- 4.13 Evaluation of Five Humanitarian Programmes of the Norwegian Refugee Council (NRC) and of the Standby Roster NORCAP
- 5.13 Real-Time Evaluation of Norway's International Climate and Forest Initiative Contribution to Measurement, Reporting and Verification
- 1.14 Can We Demonstrate the Difference that Norwegian Aid Makes? Evaluation of results measurement and how this can be improved
- 2.14 Unintended Effects in Evaluations of Norwegian Aid
- 3.14 Real-Time Evaluation of Norway's International Climate and Forest Initiative
- 4.14 Evaluation Series of NORHED Higher Education and Research for Development. Theory of Change and Evaluation Methods.

Norad
Norwegian Agency for
Development Cooperation

Postal address
P.O. Box 8034 Dep. NO-0030 OSLO
Visiting address
Ruseløkkveien 26, Oslo, Norway

Phone: +47 23 98 00 00
Fax: +47 23 98 00 99

postmottak@norad.no
www.norad.no

