EVALUATION METHODOLOGIES OF RESEARCH INFRASTRUCTURES: AN EVALUATION MODEL PROPOSAL FOR TURKEY Ö. FURKAN SİRKECİ 2018

T.C.
ANKARA YILDIRIM BEYAZIT UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF MANAGEMENT AND ORGANIZATION

EVALUATION METHODOLOGIES OF RESEARCH INFRASTRUCTURES: AN EVALUATION MODEL PROPOSAL FOR TURKEY

MASTER'S THESIS

ÖMER FURKAN SİRKECİ

MAY 2018

EVALUATION METHODOLOGIES OF RESEARCH INFRASTRUCTURES:

AN EVALUATION MODEL PROPOSAL FOR TURKEY

A THESIS SUBMITTED TO

THE INSTITUTE OF SOCIAL SCIENCES OF

ANKARA YILDIRIM BEYAZIT UNIVERSITY

BY

ÖMER FURKAN SİRKECİ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR DEGREE OF MASTER OF SCIENCE

THE DEPARTMENT OF MANAGEMENT AND ORGANIZATION

MAY 2018

Approval of the Institute of Social Sciences

Doç, Dr. Seyfullah YILDIRIM

Manager of Institute

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science of Banking and Finance.

Prof. Dr. İsmail EROL

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science of Management and Organization.

Doç. Dr. Nilay ALÜFTEKİN SAKARYA

Supervisor

<u>Examining Committee Members</u>

Doç. Dr. Nilay ALÜFTEKİN SAKARYA (AYBU)	-------------------------
Doç. Dr. Kerim ÖZCAN (AYBU)	-------------------------
Doç. Dr. Ayşegül TAŞ (Çankaya Üniversitesi)	-------------------------

PLAGIARISM

1

I hereby declare that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work; otherwise I accept all legal responsibility.

Name, Last Name: Ömer Furkan SİRKECİ

Signature:

ABSTRACT

EVALUATION METHODOLOGIES OF RESEARCH INFRASTRUCTURES:

AN EVALUATION MODEL PROPOSAL FOR TURKEY

Sirkeci, Ömer Furkan

MSc., Department of Management and Organization

Supervisor: Doç. Dr. Nilay ALÜFTEKİN SAKARYA

May 2018, 95 Pages

This study examines evaluation methodologies of research infrastructures, elaborates international practices and suggests an evaluation model for assessments conducted in Turkey. Firstly, research infrastructure concept is presented and major research evaluation methodologies with research impacts and related impact indicators in literature are analyzed. After that, five country examples of research infrastructure evaluation are scrutinized in terms of methodologies and indicators employed. Consequently, the model proposal is made which is based on a detailed analysis of international practices while taking into consideration the applicability of methodologies and tools in the context of the Turkish Research Infrastructure ecosystem. Proposed evaluation methodology sets the basis for a new research evaluation and provides a concrete model for its implementation.

Keywords: Research Infrastructures, Research Evaluation, Impact, Indicators

ÖZET

ARAŞTIRMA ALTYAPILARININ DEĞERLENDİRME YÖNTEMLERİ VE TÜRKİYE İÇİN MODEL ÖNERİSİ

Sirkeci, Ömer Furkan

Yüksek Lisans, Yönetim ve Organizasyon Bölümü

Danışman: Doç. Dr. Nilay ALÜFTEKİN SAKARYA

Mayıs 2018, 95 Sayfa

Bu çalışma kapsamında araştırma altyapılarının değerlendirme yöntemleri ile bu alandaki uluslararası uygulamalar incelenmiş olup sonuç olarak Türkiye'de yapılan değerlendirmeler için bir model önerisi yapılmıştır. Öncelikle, araştırma altyapısı kavramı hakkında bilgilendirme yapılmakta olup, literatürde yer alan başlıca araştırma değerlendirme metodolojileri araştırma etkileri ve ilgili etki göstergeleri ile birlikte ele alınmaktadır. Daha sonra araştırma altyapısı değerlendirmesine ilişkin beş ülke örneği, kullanılan yöntemler ve göstergeler açısından incelenmiştir. Sonuç olarak uluslararası uygulamaların ayrıntılı bir analizine dayanan ve aynı zamanda Türk araştırma altyapısı ekosisteminde yöntemlerin ve araçların uygulanabilirliğini de dikkate alan bir model önerisi yapılmaktadır. Önerilen değerlendirme yöntemi ile yeni bir araştırma değerlendirmesinin temeli belirlenmekte ve uygulamaya yönelik somut bir model ortaya konmaktadır.

Anahtar Kelimeler: Araştırma Altyapıları, Araştırma Değerlendirmesi, Etki, Göstergeler

Dedication

To My Family

ACKNOWLEDGMENTS

I wish to express my deepest gratitude to my supervisor Doç. Dr. Nilay ALUFTEKİN SAKARYA for her guidance, advice, criticism, encouragements, and insight throughout the research.

I also would like to thank Doç. Dr. Ayşegül TAŞ and Doç. Dr. Kerim ÖZCAN for their suggestions and comments.

TABLE OF CONTENTS

PLAGIARISM ____ iii

ABSTRACT ____ iv

ÖZET ____ v

DEDICATION ____ vi

ACKNOWLEDGMENTS ____ vii

LIST OF TABLES ____ xi

LIST OF FIGURES ____ xii

LIST OF ABBREVIATIONS ____ xiii

1 INTRODUCTION ____ 1

1.1. Background ____ 1

1.2. Problem Discussion ____ 2

1.3. Purpose ____ 3

1.4. Research Questions ____ 3

1.5. Limitations ____ 3

2 RESEARCH INFRASTRUCTURES ____ 4

2.1. Research Infrastructure Concept ____ 4

2.2. Research Infrastructures in Turkey ____ 6

2.3. Legislation about Research Infrastructures in Turkey ____ 7

3 EVALUATION METHODOLOGIES OF RESEARCH INFRASTRUCTURES ____ 11

3.1. Roots of Research Evaluation ____ 11

3.2. Research Evaluation and Purposes ____ 12

3.3. Research Evaluation Challenges ____ 13

3.4. What is evaluated: Inputs, Outputs, Outcomes, and Impacts? ____ 14

3.5. How to Evaluate: Qualitative and Quantitative Techniques ____ 17

3.5.1. Main Qualitative Evaluation Tools 17
3.5.1.1. Peer Review 17
3.5.1.2. Interview 22
3.5.1.3. Site Visit 24
3.5.1.4. Case Study 26
3.5.2. Main Quantitative Evaluation Tools 28
3.5.2.1. Bibliometrics 28
3.5.2.2. Economic Analysis 30
3.5.2.3. Data Mining 32
3.5.2.4. Survey 33
3.5.3. Outline of Evaluation Tools 35
3.6. Impacts of Research Infrastructures 36
3.7. Main Indicators of Research Infrastructure Impact Types 39
3.7.1. Scientific Impact Indicators 39
3.7.2. Economic Impact Indicators 40
3.7.3. Social Impact Indicators 41
3.7.4. Educational Impact Indicators 42
3.7.5. Technological Impact Indicators 42
4 INTERNATIONAL RESEARCH INFRASTRUCTURE EVALUATION PRACTICES 44
4.1. France High Council for the Evaluation of Research and Higher Education (HCERES) 46
4.1.1. Overall Evaluation Process 46
4.1.2. Evaluation Indicators 47
4.1.3. First Stage of Evaluation 48
4.1.4. Second Stage of Evaluation 49
4.2. Portugal National Research Council (FCT) 50
4.2.1. Overall Evaluation Process 50
4.2.2. Evaluation Indicators 51
4.2.3. First Stage of Evaluation 52
4.2.4. Second Stage of Evaluation 54
4.3. The German Council of Science and Humanities (WR) 55

4.3.1. Overall Evaluation Process 56
4.3.2. Evaluation Indicators 57
4.3.3. Individual Evaluations 58
4.3.4. Comparative Evaluations 58
4.4. Australian Research Council (ARC) 59
4.4.1. Overall Evaluation Process 59
4.4.2. Evaluation Indicators 60
4.4.3. First Stage of Evaluation 61
4.4.4. Second Stage of Evaluation 62
4.5. Czech Ministry of Education, Youth, and Sports (MSMT) 63
4.5.1. Overall Evaluation Process 64
4.5.2. Evaluation Indicators 64
4.5.3. First Stage of Evaluation 66
4.5.4. Second Stage of Evaluation 66
5 DISCUSSION FOR MODEL PROPOSAL AND CONCLUSION 68
5.1. Brief Analysis of World Examples 68
5.2. Discussions for Suggested Indicator Set of Turkey 70
5.3. Discussions for Developing Methodology for Turkey 74
5.4. Suggested Overall Evaluation Process 77
5.4.1. First Stage of the Proposed Evaluation 77
5.4.2. Second Stage of the Proposed Evaluation 78
5.5. Conclusion 80
REFERENCES 83

LIST OF TABLES

Table 1 Outline of the Features of the Evaluation Tools Reviewed....................................35
Table 2 Main Impacts of Research Infrastructures..38
Table 3 Main Stakeholders' Interest from Impacts of Research Infrastructures....................39
Table 4 Main Indicators of Research Infrastructure Impact Types.......................................43
Table 5 HCERES Evaluation Indicators 47
Table 6 FCT Evaluation Indicators 51
Table 7 FCT Qualitative Overall Grading...53
Table 8 WR Evaluation Indicators 57
Table 9 ARC Evaluation Indicators 61
Table 10 MSMT Evaluation Indicators..65
Table 11 Methodologies of World Examples...68
Table 12 Suggested Evaluation Indicators for Research Infrastructures...............................71

LIST OF FIGURES

Figure 1 WR Overall Evaluation Process..56
Figure 2 Suggested Evaluation Model for Turkey 77

LIST OF ABBREVIATIONS

ARC	Australian Research Council
CERN	European Organization for Nuclear Research
FCT	Portugal National Research Council
HCERES	High Council for the Evaluation of Research and Higher Education
MSMT	Ministry of Education, Youth, and Sports
OECD	Organisation for Economic Co-operation and Development
R&D	Research & Development
RIC	Research Infrastructures Council
TUBITAK	The Scientific and Technological Research Council of Turkey
WR	The German Council of Science and Humanities

1 INTRODUCTION

1.1. Background

Science, technology, and innovation capabilities are one of the most important factors of competitiveness and socio-economic development. In Turkey, major advances in the area of science, technology, and innovation have been obtained in recent years. Gross domestic expenditure on research and development (R&D) has increased from 1.8 billion TL to 24.6 billion between 2002 and 2016. Gross domestic expenditure on R&D as a percentage of the gross domestic product was 0,53 percent in 2002 and as of 2016, this ratio is 0,94 percent (Turkstat, 2017).

Despite these developments, the share of gross domestic expenditure on R&D at the gross domestic product is low when compared with developed countries. By its very nature, R&D activities are economically risky and for this reason, private sector R&D expenditures are limited to some extent. High and unbearable research costs and long-term return on investment for private sector cause market failures. For these reasons, public sector undertakes some risks from the production of knowledge to the commercialization process. In this context, governments establish research infrastructures for technology and innovation development and basic research activities with high risk.

Whilst research infrastructures are basically founded for research purposes, impacts and interactions of these institutions reach much more wide areas. Equipped with up to date technological facilities and qualified human resources, research infrastructures promote innovation, foster economic activities, open new market opportunities by technology transfer activities, contribute educational activities and affect social, the cultural structure of its environment (Rizutto, 2012).

Important steps have been taken to strengthen the research infrastructure ecosystem in Turkey so far. However, there are some still inadequacies in operation, management and R&D activities of research infrastructures in Turkey. Therefore, the contribution of these infrastructures to economic and social development of Turkey remains limited. For this reason, recently "Law on Supporting Research Infrastructures" No. 6550 has been adopted in 2014. This law brings new evaluation system for these infrastructures and includes reforms for efficient use and sustainability of research infrastructures.

1.2. Problem Discussion

Research Infrastructures play a crucial role in scientific research and its dissemination. They are dynamically established and used in almost every research domains, and let for many important and advance research findings and discoveries. Research infrastructures are not only focused on basic research, they have also been constructed recently to deliver direct research and development support for the overcoming major challenges that humanity faces.

Accordingly, constructing and operating research infrastructures needs a rising amount of government funding, and public institutions are thus more and more concerned with the money invested and the benefit that these infrastructures provide. These concerns are increased pressure on public budgets. Due to accountability issues, return on investment in research infrastructures is crucial. It is not convenient to conceptualize this issue in terms of traditional financial methods. Research infrastructure investments have a very wide span of benefits which reach almost all society beyond just only economic returns. Statistics are not enough to measure these benefits and cannot show them in terms of development and usage of research infrastructures. Thus, there is a need for well-organized and distinctive approaches for measuring these benefits on science, economy, and society as a whole. Today there is not integrated and agreed framework for research infrastructure evaluation in the world.

As research infrastructures have now become policy targets, they should also subject to evaluation practices. Specifically, there are growing understandings that research infrastructure evaluation needs to track more rough and systematic processes mobilizing consistent and reliable methods. There is heavy pressure to perfectly evaluate research infrastructure impacts (Pancotti et al., 2014). According to Organisation for Economic Co-operation and Development (OECD, 2016), there is a need for reliable and credible methodologies, indicators, and best practices for evaluating research infrastructures.

In Turkey, research infrastructure ecosystem is quite young and performances of the research infrastructures have not measured systematically. As stated before, "Law on Supporting Research Infrastructures" No. 6550 has been published on this issue. The evaluation methodology is now constituted by The Scientific and Technological Research Council of Turkey (TUBITAK) and further developments are still possible and continuing.

Taking into consideration of this global need and current research infrastructure ecosystem and regulations in Turkey, this thesis focuses on suggesting an evaluation methodology for Turkish research infrastructures in the light of current methodologies in literature, existing country practices and Turkey’s dynamics.

1.3. Purpose

The main motivation for this research is a need for methodology and indicators to evaluate research infrastructures in Turkey. Main research and research infrastructure evaluation methodologies need to be scrutinized and crucial country examples on the evaluation of research infrastructures need to be analyzed.

1.4. Research Questions

To achieve this purpose, the thesis will try to answer the following research questions:

- RQ1: What are the main evaluation techniques for research evaluations?
- RQ2: What are the main impacts of research infrastructures?
- RQ3: What are the main indicators for research infrastructure evaluation?
- RQ4: Which evaluation techniques and indicators are used by different countries?
- RQ5: What is the suitable evaluation methodology and indicator set for Turkey?

1.5. Limitations

This study is concentrating only on the research infrastructures which are included in Law No 6550 and public research infrastructures, private research infrastructures, and central laboratories are not included in this thesis. The study is also limited to the country examples where national evaluations are performed.

2 RESEARCH INFRASTRUCTURES

In this chapter, brief information about research infrastructures will be introduced first. Research infrastructure definitions will be provided from all over the world. After that, Turkish research infrastructure ecosystem will be elaborated with current landscape and limitations. At the end of the chapter, legislation about Turkish research infrastructures, which is also one of the main motives of this work, will be introduced.

2.1. Research Infrastructure Concept

The Oxford English dictionary defines infrastructure as “basic structural foundations of a society or enterprise”. So, research infrastructures can be simply defined as fundamental structural foundations of research (Sumathipala, 2014). In early definitions, National Research Council (1998) states that research infrastructures are crucial elements of the research enterprise. For that reason elaborating different definitions of research infrastructures used by the global research community is valuable.

In Mapping of the European Research Infrastructure Landscape Project, European Commission and European Science Foundation (2013) defines research infrastructure as a facility or platform that provides the research community with resources and services to perform top-level research in their relevant areas. As stated by National Research Infrastructure Council of the Australian Department of Innovation and Industry, Science and Innovation (2011) research infrastructures contain facilities, properties, and services which support science across the innovation ecosystem and which maintain the capacity of scientist to perform excellent research and provide innovation results.

According to German Federal Ministry of Education and Research (2011), definition research infrastructures are scientific collections and objects form an indispensable basis for research in various scientific areas. In Promoting African European Research Infrastructure Partnership project, research infrastructures defined as facilities, resources, and services used by researchers for conducting cutting-edge research for the generation, dissemination and conservation of knowledge (Sumathipala, 2014). Last but not least, from the United States, University of Minnesota Research Infrastructure Task Force (2006) adopted the definition of research infrastructures as the facilities (e.g., laboratories and clinics) and services (e.g., computing services and libraries) needed to produce novel and leading academic output (e.g., publications).

After introducing these definitions from all over the world, it can be claimed that research infrastructures are producing knowledge through research, disseminating new knowledge through education and contributing to the development of innovation. With these characteristics, research infrastructures are at the center of research, education and innovation activities. The term research infrastructure is generally used for the machine-equipment which science communities use in their advanced studies and for the buildings in which these all exist.

This term includes not only machine and equipment but also information based capitals like researchers, technicians, information and communication technology facilities and the methods of study. Large-scale research units (e.g. telescopes and accelerators), physically distributed research units and e-infrastructures (e.g. high-performance computing and information nets) are all included in research infrastructure concept (Türkiye Büyük Millet Meclisi, 2014).

Research Infrastructures which perform services on an international level, serve the society as the places where advanced scientific research activities take place. In these places, research findings are reported to the society and private sector needs are satisfied together with research development and innovation activities. These places are also equipped with state of the art tools and they bring critical mass together. European Organization for Nuclear Research (CERN) is an important example of research infrastructure and this infrastructure is used by more than 70 countries and 12.000 researchers from 105 different nationalities. CERN has hosted and contributed a lot of discovery and innovation activities (CERN, 2016).

Research infrastructures can be defined as sub-systems of the national innovation system. Research infrastructures can be set up at public institutions, the private sector, universities and research institutes. These infrastructures are acting for guiding and regulating processes of knowledge production and transferring of knowledge into new products and technologies (Laranja, 2009).

As technology develops and boundaries of research expand, research infrastructures comprise complex systems and become a more expensive investment. Instead of single-user investments, these investments need to be used effectively and efficiently by common users. For these reasons, research infrastructures are mainly funded by governments all over the world (Türkiye Büyük Millet Meclisi, 2014).

2.2. Research Infrastructures in Turkey

Research infrastructures are built in universities or public institutions and focused on national and regional priorities by taking into account public and private sector needs in Turkey. Establishment and development of research infrastructures are supported within the scope of investment programs by Ministry of Development in Turkey since the 2000s. Ministry of Development provides research infrastructure support with investment program for government organizations and universities by considering national and regional priorities as well as needs of the public and private sectors. For this purpose, approximately 6 billion TL based on 2017 prices has been allocated for these infrastructures in last 14 years. This investment includes both public and university research infrastructure investments (Kalkınma Bakanlığı, 2017).

Within the context of investment programs, two types of research infrastructure supports are provided: "thematic research laboratories" and "central research laboratories".

Thematic research laboratories are specialized research units in a particular scientific field and have the capacity to conduct research activities at the national and regional level. Via these research infrastructures; creating advanced research opportunities, advancing quantity, and quality of human resources, contributing to the development of university-industry cooperation is aimed. By the end of 2016, 99 thematic university research laboratory projects have completed and 62 projects are still under development. In terms of public research infrastructures, 33 projects are completed and 45 projects are still continued. These infrastructures are predominantly operating in health, aerospace, space, Information and communications, defense and nano-technology fields (Kalkınma Bakanlığı, 2017).

Central research laboratories support is provided for research needs of universities different units'. These infrastructures are not under the control of researchers or units, and open to all researchers. Main aims of these infrastructures are attracting more qualified personnel to contribute research culture and contribute to the development of concrete projects in university-private sector cooperation. By the end of 2016, 57 central research laboratory projects have completed and 38 projects are still continued (Kalkınma Bakanlığı, 2017).

2.3. Legislation about Research Infrastructures in Turkey

Although huge investment is made on research infrastructures, expected impact and contribution has not been harvested from these infrastructures yet. There are several problems and reasons specified by the highest science and technology policy council of Turkey which is Supreme Council for Science and Technology. Main problems can be outlined as follows (Bilim ve Teknoloji Yüksek Kurulu, 2010):

- Lack of institutional, effective and sustainable management
- The uncertainty of status, the absence of budget and personnel regime
- Due to personnel regime and wage policy, inability to attract qualified researchers from within or outside the country
- The lack of supporting elements for partnership and cooperation with the stakeholders, especially with the private sector
- Financing problems related to the investment and current expenditures needed after the initial investment

These problems on the sustainable and efficient operation of research infrastructures were discussed at 22nd meeting of Supreme Council for Science and Technology and the decree 2010/203 "Administrative and Legal Regulations on Research Infrastructures Roadmap and Research Centers" was accepted.

In this decree, the following opinions and suggestions are listed (Bilim ve Teknoloji Yüksek Kurulu, 2010):

- Elimination of uncertainties regarding the legal status and governance model of research infrastructures
- Creation of sustainable financial structures for research infrastructures
- Solution of the problems on employment of qualified personnel at research infrastructures
- Creation of research infrastructure network
- Evaluation and monitoring of performance of research infrastructures
- Providing support to the principal investigators during the installation phase of the research infrastructures
- Promotion of cooperation between research infrastructures and the SMEs
- Realization of ISO standards and accreditation work at research infrastructures

Towards these opinions and suggestions, in order to ensure sustainability and efficient use of infrastructures; administrative and legal arrangements are planned on governance, sustainable finance, staffing and accessibility of research infrastructures, cooperation between research infrastructures and relevant private, public institutions and performance evaluation, monitoring of research infrastructures.

"Law on Supporting Research Infrastructures" No. 6550 is enacted on 10 July 2014 and secondary level legislation is enacted on 28.08.2015. The main purpose of this Act is to support research infrastructures in order to ensure sustainability and more effective use of research infrastructures. In this act, research infrastructures are defined as units performing R&D activities which have qualified human resources and machinery, equipment, hardware and software based on modern technologies.

The Law's one of the main targets is the provision of a more competent and flexible managerial capacity to the research infrastructures that are established in Turkish universities. By this law, Research Infrastructures Council (RIC) is constituted by three ministers: Minister of Development, Minister of Science, Industry, and Technology and Minister of Education as the supreme governing body. RIC is authorized to decide research infrastructures that could function under the Law No 6550, set policies that are connected with research infrastructures and direct financial issues about research infrastructures such as budgeting. Under this council Monitoring and Proficiency Evaluation Committee is constituted. This committee is mainly responsible for preparing a recommendation on proficiency decision of research infrastructures for RIC.

In Law No 6550, 3 managerial bodies are established within the research infrastructures. The first body is a board of directors. Board of directors is the decision-making body of the research infrastructure and has all the administrative and financial responsibilities related to infrastructure. The second body is the advisory board. Advisory board is established for advising on strategic targets and action plans of research infrastructure. The third body is the infrastructure manager. The manager is responsible for ensuring that the activities in infrastructures are carried out in accordance with the strategic objectives and action plan. The manager is also responsible for financial issues, hiring of personnel and research agenda of the research infrastructure. The manager can also employ researchers from outside the country as well. These employment contracts will be bound to civil laws other than the administrative law which will keep these infrastructures more effective in managing their research team with enhanced conditions raised on win-win approach attracting high-qualified researchers with more salaries.

In addition to structural and financial reforms, a new evaluation process will be introduced by Law No 6550. After evaluation which will be conducted by Ministry of Development and TUBITAK, eligible research infrastructures will acquire legal entity and proficiency status. The law defines proficiency, coordination, monitoring, and evaluation in the Article 15 of the Law No 6550. According to this article, research infrastructures to be subject to the assessment are determined by the Ministry of Development and the evaluation will be carried out by the TUBITAK. After the evaluation process, eligible infrastructures will acquire proficiency certification which will be valid for 5 years.

The new research infrastructure law targets increase in economic and social benefits of infrastructures besides financial and managerial sustainability. In order to achieve this, evaluation methodology and indicators need to be compliant with this target. The mission of developing the evaluation system is assigned to TUBITAK and system design and developments are still continuing in the evaluation methodology and indicator set. Research and research infrastructure evaluation processes and methodologies are dynamic processes and continuous improvement is needed for qualified and effective measurement. For these reasons, in the third part, research and research infrastructure evaluation methodologies will be scrutinized in order to understand the general environment of the evaluation processes and to inform about current methodologies.

3 EVALUATION METHODOLOGIES OF RESEARCH INFRASTRUCTURES

Understanding and conceptualizing research infrastructure evaluation requires sufficient information about research evaluation. For that reason, firstly research evaluation concept will be introduced in this chapter. Main purposes and challenges of research evaluation will be scrutinized. After this brief introduction, answers to what is evaluated in research evaluations and how to evaluate research activities questions will be elaborated. In the last part of the chapter, impacts of research infrastructures and related indicators will be argued with current studies in the literature.

3.1. Roots of Research Evaluation

Research evaluation is a relatively new topic and has quickly developed within the short time period. Early works on research evaluation and impact assessment of research can be found in the 1970s. Science - society relationships and social responsibility of the researchers were elaborated in the literature in those times (Rose et al, 1971, Nelkin, 1979). Meanwhile, some other researchers such as Werskey (1971) had argued whether research activities and their results were being used for the benefit of the society as well. The 1970s were also the years of the economic crisis which leaded drastic decreases in research funding due to cuts in research expenditures. Limited capital leaded funding agencies to be more selective in research activities to be funded.

On the other hand, during that period, funding and other relevant decisions about research were taken in a secret manner and with only peer review. Because of some limitations in peer review methodology, it was debated that decision making processes needed to be transparent and accountable (Irvine and Ben, 1984). For that reason, data was required in inputs, outputs, outcomes, and impacts of the research activities. Science Policy Research Unit and Centre for Science and Technology Studies groups in Leiden University were amid the very first in designing policy tools for research evaluation in the late 1970s (Moed et al, 1985, Martin and Irvine, 1983). These initial trials faced some resistance from researchers who were worried about foreigners violating their area and challenging the authority of peer review as the main decision-making methodology (Martin, 1997).

Progressively, on the other hand, under severe economic and political pressure, funding agencies all over the world started to implement more systematic approaches for evaluation of research and today these evaluations are performed all over the world.

3.2. Research Evaluation and Purposes

Evaluation is an important instrument for both policy development and decision making processes so it should be planned, managed and conducted in an appropriate manner. In detail, this means defining a clear set of aims, well-structured framework, and systematic methodologies (Ruegg and Feller, 2003).

Research evaluation is a tool used for funding, enhancing, directing and managing the operations of research institutions. Due to the accountability of public bodies for spending on research institutions, research evaluation methodologies have been studied and enhanced over recent years by funding and policy bodies.

According to Rand Corporation's report which is prepared for the Association of American Medical Colleges to analyze methodologies and indicators used to evaluate research, there are four main purposes to conduct research evaluation: accountability, advocacy, allocation, and analysis (Gutrihe et al, 2013).

Accountability is the main driver for research evaluation as mentioned before. By accountability, the main target is to demonstrate whether money spent on research is used in an effective and efficient manner. Also for this purpose, researchers become accountable for their research activities. Secondly, by advocacy, mainly benefits of research are tried to be explained to the public. Developing the understanding and implications of research in the eyes of the public is also aimed at this concept.

Allocation is another important motivation for the research evaluation. Budgetary pressures and limited resources force governments to be selective in research funding. In order to allocate a limited amount of fund to research organizations, public bodies are evaluating scientific activities. Last motivation for research evaluation is analysis. The analysis is mainly performed for assessing the effectiveness of research and for providing a base for strategy and management of research.

Allocation and advocacy terms may be confused in terms of funding issues. By allocation mainly past performance is evaluated in order to allocate the funds available while by advocacy increasing overall amount of the funding of research is targeted.

3.3. Research Evaluation Challenges

One of the very early and impressing reports about research evaluation challenges "The Evaluation of Federal Research Programs" was written by Salasin et al, (1980). There are 4 characteristic challenges identified in research evaluation in this report. The first challenge is the absence of clear understanding of effectiveness. It is very unsound and complicated to evaluate and measure research without a clear definition of effective research. The second problem is competing objectives of research which makes measuring very complicated. Research objectives are now multiplied and diversified even more when compared to 1980s and assessing competing objectives such as assisting economic development, training researchers etc. become a more challenging issue.

The third difficulty is gathering results of the research. Results of research activities are not only scientific results and outputs. More diversified and complicated socio-economic impacts also exist. The last difficulty is related to the third difficulty which is compromising political and scientific measures of achievement. Beyond these difficulties, the report is also arguing the possibility and effectiveness of evaluating research based on only concrete outputs of research such as articles. Difficulty list is still up to date and challenges about research evaluation debated in the report are still being discussed today in the literature.

Another prevalent challenge argued in the literature is contribution problem. Most of the research activities are conducted by a group of scientists across research bodies operating in similar areas. It is complicated to specify which certain part of research activity contributed to a specific impact (Georghiou, 2002). Duplication of impact may occur where there are various and joint contributions of research activities. For concrete outputs like articles and patents, it is comparatively straightforward to think of ways in which contribution can be specified whereas for less concrete outputs such as contributions made in policy changes or establishment of spin-off firms etc. it is hard to specify exact contributions to the relevant fields. In these cases, various parts of research are included.

As interdisciplinary research gains importance, evaluating the results needs a wide variety of knowledge and expertise.

Time lags are another crucial challenge in research evaluation. Especially in basic research activities and also in other research activities, the substantial time period is needed for impacts to be actualized. In the medical research area, this time lag is calculated for about 17 years which is the period between publication and practice (Morris et al, 2011). For this reason, it is complicated to assess the impacts of the research activities in the short run soundly. Waiting decades for assessing research is also not feasible due to accountability issues.

Research evaluation also needs some classification of research activities such as basic/applied research, excellent/good/poor research, or transformative/incremental research etc. especially in scoring or comparing phases of the evaluation. Hence, it is a great challenge to place the research activities which fall within boundaries of classifications. It is especially important that these classifications have a great impact on researchers and institutions being assessed. This is why evaluation systems need to take into account research activities which fall within boundaries between classifications (Guthrie et al, 2013).

All types of evaluation need data to conduct. Quality of available data is the most important indicator of sound and quality based evaluation. Data sources should be accurate, valid and reliable in order to conduct effective evaluation campaign. In order to guarantee these qualifications, verification and confirmation are needed from related institutions. Some research data may also be confidential and almost impossible to reach. Researchers or research institutions are not willing to share these kinds of data such as private research infrastructure-industry collaborations.

3.4. What is evaluated: Inputs, Outputs, Outcomes, and Impacts?

Evaluation measures can be taken in different phases of the research process. Generally, four measures are specified in every process of the research.

- Input measures are utilized for resources used in the operation.

- Output measures are products and services produced after the implementation of research.
- Outcome measures are showing first and short-term impact of the research conducted.
- Impact measures are reflecting long-term effects of research (Brutscher et al, 2008, Frank et al, 2009).

Input measures are especially important in terms of accountability issues. Therefore, inputs used in the research need to be clarified. The amount of resources used is a crucial base for research measurement. Input measures could be in economic terms such as the amount of money that is used for building the infrastructure as well as human resources terms such as the number of doctorate holders at the infrastructure.

The main aim of the research activities is the progression of knowledge and this progression is generally provided with research results or outputs. These outputs can be articles, books, scientific reports, seminar papers etc. There are also “second rank outputs” which are closely tied to the research outputs specified above. These can be in forms of patents, licenses, and software etc. These outputs can be easily quantified and used in the evaluations. There is also a measurement of the process which is realized between input and output phases. Process measures are generally used for analysis objectives.

Beyond these countable outputs, there are also outcome measures. Outcomes may be in form of high-quality graduates, technological innovations, the added value created by contact research activities etc. There is a high concern for adding more and more outcome measures to the evaluation systems even though it is difficult or sometimes not possible to define them in quantitative terms.

Impact measures are long-term measures that cannot be easily assessed. Quantifying these measures is more complicated than outcome measures and measurement is much harder to implement. Examples of impact measures are increasing the productivity and employment, improvement of the well-being etc.

Generally, input-output-outcome-impact measures are seen as a linear process. The linear process suggests that research flows to the one way only; basic research transfers knowledge into the applied research, applied research results in innovations which further leads to impact. This linear model is not reflecting the reality that much because sometimes impact relates to final outcome. It is almost impossible to recognize the time that outcome is final. Instead of linear process and framework given before, it is preferable to define impact as a wide influence on society.

Although the linear process may be helpful to the conduct analysis and understand the main dynamics, there are no evident separating lines between especially output-outcome and impact measures (European Science Foundation, 2012). This linear model has been renovated into a more complicated consideration of the interaction between inputs, outputs, outcomes, and impacts. Representations of this are the alleged as Payback Framework (Buxton and Hanney, 1996). This framework combines both academic results and social benefits and can be analyzed in two parts. In first part, the research and diffusion processes are broken into particular modules within which the paybacks of research activities can be examined. In second part, a multi-dimensional classification structure is created into which the numerous outputs and impacts can be located (Hanney and Gonzalez Block 2011).

Measures used for evaluation is also related to the purposes of the evaluation stated before. Accountability and advocacy purposes are generally related to upstream measures which are outputs and some outcomes, while steering and learning purposes are related to downstream measures such as outcomes and impacts.

Evaluation can further be separated into ex-ante and ex-post forms. Ex-ante evaluation is performed before research to measure potential and chance of success. In general, ex-ante evaluation is performed with input measures and usually based on prospects and estimations. Whereas ex-post evaluation is conducted after research has been performed and output, outcome, and impact are evaluated (Geuna and Martin, 2003).

Furthermore, mid-term evaluation is also used while research activities are still continuing. By mid-term evaluations, ex-ante evaluation results are reviewed and some outputs and outcomes are measured. Although ex-post evaluation is mostly preferred at the evaluation systems, the ex-ante evaluation also gains importance due to time lag challenge explained before as time is needed for the appearance of research results. Even though it is assumed that research output, outcome or impact will appear once the research completed, actually it can take decades (Lu, 1981). So, ex-ante evaluation is also a valuable tool for research evaluation.

3.5. How to Evaluate: Qualitative and Quantitative Techniques

For every research area and level of units being evaluated throughout the world, research evaluation depends on two main corresponding techniques namely qualitative and quantitative techniques. Some countries have used the former technique, in some countries the latter and they sometimes change from one to another.

The first technique, qualitative evaluation includes evaluation of the assessed unit by researchers in the same scientific area. As the name implies the quality of the research is tried to be grasped in that technique. Main tools that are used for qualitative evaluation are namely; peer review, interviews, site visits and case studies.

The second technique, quantitative evaluation mainly measures the metrics. In order to measure, this technique produces indicators which enable comparison and evaluation among evaluated units. Quantitative evaluation selects the indicator information and uses numbers to grasp evaluation results. Bibliometrics, economic analysis, data mining, and surveys can be given as highlighted examples which are employed tools under the quantitative evaluation. Both qualitative and quantitative evaluation tools will be introduced in the following sections.

3.5.1. Main Qualitative Evaluation Tools

Qualitative evaluation tools are mainly inclusive, formative, flexible and suitable for cross and multi-disciplinary research evaluations. In this section; peer review, interviews, site visits, and case studies will be scrutinized in terms of when and how to use them in the research evaluation campaigns. Also, some advantages and limitations of the tools will be argued.

3.5.1.1. Peer Review

Peer review has a conventional and fundamental role in research evaluation. It is the main tool of decision-making in article review, promotion of researchers, allocation of resources and even designation of scientific programs. Peer review is employed by all actors such as public bodies, industry, and academia. According to Organisation for Economic Co-operation and Development peer review also has gained importance in the research institution evaluation processes and in ex-post evaluations (OECD, 2011a).

Peer review processes have been part of research system from the time when the first scientific journal: “Philosophical Transactions” with the first officially well-defined peer review process being that of the “Medical Essays and Observations” article was published in 1731 by the Royal Society of Edinburg (Kronick, 1990). Therefore, peer review is a quite ancient tool in qualitative evaluation.

Peer review’s legitimacy comes from the quality of reviewers selected. This reveals the specialization of scientific areas and the impracticality for strangers to evaluate research. From that perspective, peer review reinforces autonomy of the academic world, leaving the designation and implementation of evaluation to the researchers themselves. Taken these into the mind, the generally accepted definition of peer review can be outlined as follows. Peer review is an evaluation of scientific excellence by other researchers working in, or close to the research area in evaluation (OECD, 2011a). Peer review is based on the assumption that a decision about certain features of research is an expert judgment eligible for being accomplished only by those who are adequately informed about the cognitive development of the area questioned, its research agenda, and the researchers within it (Gibbons and Georghiou, 1987).

Recently in developed countries, research activities have been widely involved in innovation and entrepreneurship activities. As a result of this, peer review is used in new areas such as process evaluation, ex-post evaluation etc. This situation increases pressures upon adjustment and consideration of socio-economic measures. Moreover, conventional notions of peer review, specified before such as the autonomy of academic world and scientific excellence, are considered deficient. The main reason for this is accountability to governments and society as a whole. For that reason, evaluation of research activities are not only based on scientific excellence but also based on social, economic and other related measures (OECD, 2011b). In order to adjust to these changes, new definitions of peer review are written. In 2004, The Office of Energy Efficiency and Renewable Energy in the American Department of Energy defines peer review as rough, standardized and formal evaluation process which employs objective measures and qualified objective peers to evaluate the technical, scientific and business excellence, their results and the management efficiency of projects.

Peer review process has three main phases namely pre-review, review and decision (Guthrie et al, 2013). Pre-review stage is designed for easing review process. In this stage, quality evaluation is not performed; instead only application or documents are inspected for minimum requirements. By elimination of inadequate applications, review time and volume are decreased.

In review stage, as the name implies review is performed by researchers or panels to evaluate quality or other dimensions depending on the purpose of the evaluation. In panel evaluation, generally, review documents send in advance to reviewers in order to inform them. After this period, reviewers are gathered for collective evaluation. The advantage of this method is allowing discussions and clarifying confusions between reviewers which results in reaching consensus easily. The main pitfall of this method is the possibility of domination of the panel reviews by one or couple peers. Good moderation needed to overcome this problem. Moreover, travel costs and meeting costs of reviewers are other disadvantages of the panel review. On the other hand, individual peer review process does not require travel costs. More specific, independent reviews and comments can be collected from reviewers. This could be seen as an advantage but on the other hand, it can make it difficult to reach a consensus among reviewers.

Decision stage is the final stage at which latest evaluation results are gathered and outcome of the peer review communicated with the evaluated parts or decision makers. The outcome of the evaluation may be success/failure decision or some sort of ranking or score for comparison. Besides the outcome also some recommendation and feedback are also shared with the evaluated parts.

Quality of the peer review processes is quite important and argued in the literature. Different sets of measures are available for evaluation of the peer review processes. For example, Chubin (1994) suggested seven necessities for the qualified peer review process:

- *Effectiveness* in the management of the resources.
- *Efficiency* in the exploitation of resources such as money and time.

- *Accountability* of the process to researchers and the society.
- *Responsiveness* in assisting new ways for innovation and new research areas.
- *Rationality* by transparent procedures and functioning.
- *Fairness* by equal and objective evaluation of independent experts.
- *Validity* by obtaining the pretty much same results from repeated evaluation which can be also named as replicability and removing contingencies in the process.

These necessities include most of the current tensions and some trade-offs may exist when employing peer review processes.

The main advantage of the peer review is the trustworthiness and credibility of the tool among researchers. Peer review has been commonly employed and verified in research evaluation activities, for that reason favor for peer review among researchers is very high (Ware, 2008). Perhaps the second most significant advantage is the strong confidence and dependence to the peer review system by funding agencies and policy institutions (Research Councils of United Kingdom, 2006). These advantages are crucial in credibility and recognition of the decisions after evaluation.

Furthermore, Scott (2007) states that peer review pervades research activities from start to endpoint. From this perspective involvement in peer review processes is routine for the researchers. Sometimes researchers are even not paid for the review processes but they have the opportunity to see latest research activities which are valuable for their research vision.

Independent peer reviewers or experts have just exposed to the evaluated unit before evaluation campaigns for that reason they can easily recognize strengths, weaknesses, opportunities, threads, and ways to improvement (Bozeman, 1993). This is also an important benefit for grasping quality throughout the evaluation process. Moreover learning ways of improvement is also crucial plus of the peer review which is not that much available in other evaluation tools especially in quantitative evaluation tools.

One of the main disadvantages of the peer review tool is high costs for both reviewers and the applicants in terms of money and time (Abramo et al, 2011). First important cost is expenditure on payment of reviewers and use of review colleges. For example in the

United Kingdom 2008 Research Assessment Exercise, mainly uses peer review process, costs about £60 million to perform (Macilwain, 2009).

Another precious cost is time and it is becoming more and more valuable for high-quality researchers in this competitive research environment. For peer review processes, in order to assure qualified evaluation, high-quality researchers are employed and this creates opportunity cost in terms of time for the researchers. Peer review processes tend to be time-consuming, even in some cases, it can take over a year. In addition to this, time is also valuable for the evaluated parties; they have to wait for all the process to be completed (OECD, 2011a).

Besides to efficiency criticisms, there are also effectiveness debates about peer review. There is no clear evidence of peer review accomplishes the targets of the review. On the other hand, it is also not clear that peer review process is ineffective (Brutscher et al, 2008). This situation is mainly due to the dominance of the peer review in research evaluation practices. For this reason, there is very limited opportunity to compare peer review with other tools available.

Another criticism about peer review is about the intrinsic conservatism of the peer review. Innovative approaches to the research activities which are not supported by preceding literature might be underprivileged in the peer review process. Reviewers from academia are more likely to support tried and verified approaches that are recognized in the area questioned (Braben, 2004). Likewise, new researchers or new research units in the field might be disadvantaged because they do not have previous research history to show their strength to the reviewers.

Transparency is also argued as a challenge for the peer review process in the literature. Reviewers are usually anonymous throughout the process and the decisions taken in the panels or committees are generally opaque to the audience. Even with this possible challenge, acceptance of the peer review is pretty high as specified before.

When evaluation goes beyond the scientific excellence and looks for socio-economic impacts or other impacts, it becomes questionable whether academic reviewers are eligible enough to conduct this kind of evaluations. In order to overcome this challenge recently "Extended Peer Review" process is introduced. By extended peer review, besides scientific peers, there are also non-scientific peers such as from industry or experts from government etc. are employed in the evaluation campaigns (Pereira and Funtowicz, 2005). By extending the peer composition, more widely and comprehensive evaluation is targeted. Furthermore, some researchers go beyond in the literature and present “Extended Peer Communities”. These communities contain not only relevant experts from the relevant scientific area but of all who desire to take part in the determination of specified issue (Funtowicz and Ravetz, 2003).

In addition to “Extended Peer Review”, “Informed Peer Review” is also argued nowadays. In informed peer review, evaluation outcomes are the product of expert appraisal informed by bibliometric information and other quantitative data (Abramo and D’Angelo, 2011). For both cost concerns and thought of the potential of bibliometrics to contribute to quality evaluations, many governments have applied a composite method in which the peer review method is informed by bibliometric analyses (Jonkers and Zacharewicz, 2016). Furthermore, Hicks (2012) claims that debatably institutional or area level performance-based evaluations systems employing peer review based on bibliometric indicators characterize the present state of the art. This methodology has been applied in many countries in recent years.

The subjectivity of the peers is a must for fair evaluation in the peer review evaluations. There should not be any conflict of interest between reviewers and the evaluated parties. This is especially difficult to ensure in developing countries in terms of scientific activities where there are not enough qualified peers in the specific evaluation area. In order to overcome this problem, international reviewers are needed in these countries. Moreover, developed countries are also employing international reviewers in order to catch international quality standards and to compare their research performance with other developed countries’ performance.

3.5.1.2. Interview

Interviews are fundamentally just conversations (Flinders, 1997). Interviews are designed and used for collecting information from evaluated parties by asking those questions directly (Holm-Hansen, 2007). By interviews, evaluation parties can get extra and in-depth information about the topic questioned. Interviews can be conducted by in-person or by telephone, however, research evaluation interviews are usually performed by in-person. Interviews employed for evaluation may stimulate understanding and change however importance is on intellectual understanding instead of producing personal change (Flinders, 1997).

Interviews could be used just before the announcement of the results of the evaluation process. After conducting quantitative or qualitative evaluations in order to gather further information as well as clarify missing points throughout the evaluation process, interviews may be conducted. Also evaluated parties views and arguments about the evaluations can be acquired.

Basically, there are two types of interviews which are used for research evaluation; structured interviews and in-depth interviews (Westat, 2002). Structured interviews are conducted with cautiously designed and worded questionnaire. Emphasis on this type of interview is on getting answers to carefully formed questions. Patton (2015) describes these interviews as a fixed response interview and states that this interview type is valuable when data need to be collected and compared in a short time period.

On the other hand, in-depth interviews are not rigid as structured interviews and prearranged questions sets are not asked in order to stay open and adaptable to grasp full reaction and understanding of the respondent which is a very looked-for approach in qualitative data collection. Evaluators can use this kind of interviews when the topic is very complicated or controversial and preparing questions about the topic is hard. Also unstructured interviews are useful for illustrating data or explaining quantitative results (Tracy, 2013).

The main advantage of the interview technique is that it yields more details, new understandings, and insights. Interview technique allows evaluators to clarify questions which in turn increase chances of valuable answers (Westat, 2002). Furthermore, interview technique lets participants emphasize meaningful and crucial points with their own words in spite of being limited to predetermined questions, so participants may feel more comfortable with this technique.

Most featured limitation of interview technique is high costs in terms of money and time. Travel costs of respondents, the time needed for interviews and educating interview personnel are important burdens of interview technique (Westat, 2002). In addition to this, a lot of disorganized data is gathered which make analysis and evaluation activities difficult and time-consuming as well. Also, interview is quite subjective because questions asked and topics covered are all depending on evaluator's opinions and decisions.

3.5.1.3. Site Visit

Site visits are designed for interaction with individuals of the research unit or institution being assessed. Site visits generally include presentations by the evaluated institution, observations by experts, document review before and/or during the site visit and interviews with stakeholders of the evaluated institution throughout one or couple days (Lawrenz et al, 2012). According to Lawrenz et al. (2003) site visit occurs when individuals with specific expertise and preparation go to a site for a restricted time period and gather information about assessment object either by their own experience or through the described experiences of other experts to write down an evaluation report addressing the objective of the site visit.

Evaluated body usually arranges and introduces extensive documents about their activities both in paper and oral form during presentations. These forms and information are shared with experts before the evaluation so that site visit experts examine these documents and prepare some questions. Site visit experts may question evaluated body about the documentation provided, presentations performed and about other issues relevant to the evaluation. Generally, site visit experts are chosen by agencies that are accountable for evaluation or funding. Usually, site visits are organized by evaluation agencies.

In his study, Hansson (2010) indicates that conducting site visits to research institutions are quite common and trend practice. Site visits are employed in evaluations all over the world from the USA to European countries. Usually, site visits are performed by expert groups who deliver a connoisseurship approach to evaluation (Eisner, 1998). For this reason, selection of experts for the site visits is a vital process. Site visit experts need to be knowledgeable enough in the questioned area and capable of deciding what kind of information is needed. In addition to these, experts also should be skilled in how to gather information from the evaluated personnel (Lawrenz et al., 2003). Evaluation generally depends on expert's individual answers to the evaluation criteria based on the information they gather throughout the site visit and filtered through their understandings and experiences (Rons et al, 2011).

Usually, site visits require some sort of initial reporting such as self-evaluation reports which contains beneficial information about the evaluated body. These reports are also very helpful for the experts in discussions at the site visit. After individual assessments, members of the expert panel generally discuss their personal findings and conclusions at closed-door meetings. As a result of this meeting generally, the final report is prepared which contains conclusions of the expert panel, advice for evaluated body’s improvement as well as evaluation of the current situation. Scoring may also be part of the final report of the site visit in some cases.

The most important advantage of the site visit is that it provides a chance for evaluators to communicate with all types of stakeholders at different levels of the evaluated unit. This is also one of the main drivers of the site visits because it gives the opportunity to interact with individuals who would not normally be accessed in other evaluation techniques (Guthrie et al, 2013). Moreover, by that way site visits allow wide variety viewpoints and involvements to be taken into consideration in the evaluation process. This is also addressing another advantage of the site visit which is interactive nature of the process. Site visit enables evaluated bodies to understand and contribute to the evaluation process. This makes the site visit evaluation process more open and two-way process instead of being externally imposed an assessment.

The main limitation of the site visit is that it is expensive in terms of time and money. Sending experts to each institution for evaluation for several days is costly. These costs include travel and accommodation costs of experts. In addition to monetary costs, time costs for qualified experts are also high. These experts have tough agenda so employing them for several days in site visits is quite difficult (Patton, 2015). For that reason, it is difficult to use site visit technique across a great number of evaluated institutions. Moreover, site visits can be troublesome for experts because different stakeholders may be involved in the process. Also, the time needed for preparation for the site visit is another difficulty for the experts.

Evaluated bodies can also have some difficulties in terms of preparation for the site visits. Long time needed for short visit such as preparing presentations and informative documents for a site visit, motivating researchers and other personnel etc. Furthermore, site visits are not suitable for repetition because of its nature and cost implications (Guthrie et al, 2013). For that reason, site visits are repeated on average in every 5 years. Also in each visit, time is limited to several days which is only enough to talk with a limited group of personnel. Another limitation is about the transparency of the site visit. Site visit process is not transparent to other institutions involved in the evaluation process and comparably subjective because it rests on the communication between experts and the personnel.

3.5.1.4. Case Study

According to United States Government Accountability Office (1990), case study as an evaluation technique is a means of learning about a complicated case based on a complete understanding of that case gained through broad description and total analysis of that case and in its context. Therefore main objectives of the case study approach are to clarify, discover or describe an activity. A case study is a descriptive technique which provides in-depth information about a context with the purpose of generating a detailed understanding of that area instead of general conclusions.

Case studies can be employed in a wide variety of means in many diverse contexts, so it is complicated to categorize specific modes of operation that can be briefly described. In literature, some classifications are described as follows (Yin, 1993, Stake, 1995):

- Illustrative case studies: This type is useful for causative and descriptive evaluations which are used to illustrate a specific, unfamiliar topic. They do not have to be demonstrative and it is not appreciated to have many case studies for this purpose.
- Exploratory case studies: These case studies allow users to gather more information before developing research questions. It is generally used as an entry to beyond, more detailed research.
- Descriptive case studies: This type involves descriptive theory, then subjects are observed and the data gathered through study is compared with the existing theory.

- Instrumental case studies: In this type, focus group or individual let researchers understand more than what is primarily open to researchers.
- Cumulative case studies: As the name implies, existing data from different sources are gathered to produce a new set of case studies of this type.
- Critical instance case studies: It interested in one or more special themes which aim to understand certain situations of interest rather than reaching generalizable conclusions.

From the research evaluation perspective, case studies can belong to one or more categories specified above. This range of applicability correspondingly makes it difficult to define a strict methodology for producing case studies (Guthrie et al, 2013). Case studies are not usually conducted for comparing in research evaluation; instead, they are employed for illustrative necessities. Case studies can be used for advocacy purposes to illustrate examples of some categories of research activities or outcomes.

According to United States Agency for International Development (2013), case studies main advantages for evaluators can be outlined as the flexibility of its use, efficiency, addressing context and details. Case study approach is so flexible and broad that it is difficult to specify its use. Case studies can focus on the specific research community to access greater detail as well as research funding mechanisms. Case study method has low cost when compared with other techniques. Usually, case studies are used to introduce examples within a context in a couple of cases and combined with other tools which decrease the costs. However relative cost per subject may be high because of detail required in each case. Case studies are examining activities with its surrounding context and have the wider vision. Moreover, they have the ability to reach greater detail and context. When used in right way, case studies can provide in-depth contextual information which enables full understanding of specific cases. To sum up, the advantages of case studies have an advantage over other techniques when “why” and “how” questions are asked. Their strength is delivering context and detail to support generalizations produced by other means (Yin, 2009).

The main limitation of the case study is the specificity of its nature. According to United States Agency for International Development (2013), it is difficult to generalize any results grasped through the case studies because results can be only relevant to that specific context. Another important limitation of the case study is relative subjectivity of the technique. Even though case studies are using real data, generally outputs they rely on personal comments or inferences and it is hard to assess its validity of results. Lastly, as specified before case studies need high investment per subject.

3.5.2. Main Quantitative Evaluation Tools

Quantitative evaluation tools are producing numerical results and outputs. They do not need interpretation or judgment, for that reason they are considered objective compared to qualitative techniques. In this section; bibliometrics, economic analysis, data mining and surveys will be examined and some advantages and limitations of the tools specified will be discussed.

3.5.2.1. Bibliometrics

The term bibliometrics has been used from the beginning of the twentieth century when it was used by librarians as a tool for journal collection. Pritchard (1969) used bibliometrics as a method for the first time as follows: enlighten the process of written communication and of the nature and development of a discipline by counting and analyzing the various aspects of written communication. However, bibliography became a method that could be applied to evaluate patterns of academic communication and research outputs only after the creation of the Science Citation Index in the early 1960s (Haustein and Laiviere, 2014).

According to OECD (2008), bibliometrics utilizes data on numbers and authors of scientific publications, articles, citations and patents to evaluate the output of individual researchers, research teams, scientific organizations, and even countries, to identify national and international networks, and to map the development of new areas of science and technology.

Bibliometrics introduce insights throughout the activity, knowledge transfer and linkage measurements. Activity measurement deals with a number of scientific articles as a measure of the size and volume of the outputs and the scientific activity of the specific research area. Knowledge transfer measurement is related to citation process which shows the communication of knowledge inside the scientific community and introduces indirect measurement of quality. Linkage measurement maps the linkages between individual researchers and research areas in order to reveal the development of the cognitive and social networks of scientific research. Bibliometrics technique needs databases for these purposes. The three most featured bibliographic databases are Scopus, Web of Science and Google Scholar (Wouters et al, 2015).

One of the very important and recent principles lists about bibliometrics based research evaluation is published as the Leiden Manifesto. According to this manifesto, there are 10 principles of best practice in bibliometrics based evaluation as follows (Hicks et al, 2015):

- Quantitative evaluation needs to support expert, qualitative evaluation
- Research missions of the evaluated unit need to be measured
- Locally relevant research excellence needs to be protected
- Data collection and analysis processes need to be transparent, open, and simple
- Data verification and analysis need to be allowed for evaluated units
- Variation by research areas in publication and citation analysis should be taken into account
- Individual researchers' evaluation needs to be based on the qualitative judgment of their portfolio
- Misplaced concreteness and false precision need to be avoided
- Systemic effects of indicators and evaluation should be recognized
- Indicators should be assessed updated regularly

The bibliometric technique has substantial advantages, as the range and influence of indicators advances, however, limitations also need to be clearly understood. The first important advantage of bibliometric technique is considerable cost advantage and low administrative burden compared to other techniques (Oppenheim, 1997). Bibliometrics provide the advantage of using data that can be gathered by minimal interaction with the researchers themselves. This helps to decrease the administrative burden compared to other alternatives such as peer review as well as ensures that data gathered are probably to be representative as they are obtained for whole research societies and not just those researchers who give their consent (Smith, 1981).

The second crucial advantage of bibliometrics is objectivity. Bibliometrics do not rely on the subjective evaluations of individual researchers, but instead are an indication of the use which is made of the publication by the entire scientific society (Jonkers and Zacharewicz, 2016). Furthermore, bibliometrics may be perceived as objective in the sense that it reflects properties of the scientific document themselves, it is replicable and based on the practices and perceptions of numbers of researchers (White, 1990).

Bibliometrics evaluation results can be understood by various audiences because numerical results and trends are easy to perceive. Moreover, generated quantitative indicators for research evaluation by bibliometric measures can be very beneficial in the evaluation of large volumes of complex research material. Bibliometrics also allows analysis of trends including global trends. Results of the analysis are comparable globally which is very useful for evaluation bodies.

The main pitfall of the bibliometrics is that counts performed by this technique indicate quantity and productivity but not the impact or quality of the research clearly (Okubo, 1997). All publications have different importance and quality, adjustment processes of bibliometrics may not sufficiently account for differences in these aspects. Bibliometric evaluation counts and evaluates only publications and patents as research outputs and pays no attention to other outputs or long-term outcomes and impacts as well (Ruegg and Feller, 2003). Related to this limitation, bibliometrics also cannot assess future potential of the research.

In addition to these pitfalls, bibliometrics uses mostly publication data which can be seen as "indirect" peer review because journal articles are published only after peer review evaluation. For that reason, bibliometrics carries all the limitations of peer review as well according to OECD (2011b).

3.5.2.2. Economic Analysis

Economic analysis can be named as efficiency or economic evaluation as well. There are several types of economic analysis. Most commonly used types are a cost-benefit analysis, cost-effectiveness analysis and cost-utility analysis (Haddix et al, 2003).

Cost-benefit analysis as an economic analysis tool exhibit to investors (specifically those investing public money to the research) whether the research conducted is of benefit for the public's best interest. Economic impacts are evaluated by scheduled objectives and the analysis is generally conducted from the point of view of whole society. Typically cost-benefit analysis is limited to national vision; however, for research activities, this may not be suitable. The objective of the cost-benefit analysis is detection and monetization of all possible impacts of research under evaluation to determine the related benefits and costs (Clarke et al, 2013).

Contrary to cost-benefit analysis, cost-effectiveness and cost-utility analysis indirectly assume that one research activity will be supported without considering of its net benefit. Therefore cost-effectiveness analysis may result in a decision to support research activity that does not benefit itself because analysis assumes that the output is worth having and the only problem is to define the most effective way to achieve it (Drummond et al, 2005).

Different types of economic analyses provide different forms of evaluations. Canadian Treasury Board (1976) listed some of the standard and best practice steps of an economic analysis as follows:

- Examination of needs, constraints, and formulation of objectives
- Definition of options that allows the evaluators to compare them fairly
- Analysis of incremental impacts and collection of data about benefit and costs
- Expression of cost and benefit data in a valid monetary unit of assessment
- Application of deterministic model and conduction a sensitivity analysis to decide proper variables for evaluation
- Analysis of risks and factors that cannot be monetized

Economic analysis can increase the analytical capability of experts significantly. Use of economic analysis can contribute to an understanding of the relations between impacts and inputs when available data is imperfect and complex. Economic analysis can be used to generate quantitative results with great detail and to show cause-effect relations (Ruegg and Feller, 2003). Moreover, economic analysis provides fiscal accountability and aids for setting priorities when resources are scarce in research investments.

One of the disadvantages of economic analysis is that both the technique and results may be problematic to capture, replicate and communicate with public or non-experts. Moreover, all impacts cannot be grasped in these quantitative methods which are imperfect and variable in their capability to capture relations between changing technical know-how and social and economic impacts (Ruegg and Feller, 2003). Economic analysis main logic is depending on all inputs and outputs of activities can be converted to a common unit of impact namely dollars. However, assigning monetary values to inputs and outputs in research activities is difficult and not always appropriate to do (Weimer and Vining, 1992).

3.5.2.3. Data Mining

As research organizations gather large amounts of data in their databases, it has become quite complicated to find out how relevant information can be gathered from these data. There is more and more data stored in databases nowadays however there is not a clear understanding of these data. Data mining is a research technique to assist with that challenge (Fayyad, 2004). Usually, data mining technique is assumed to be a mechanical process of gathering data from databases. Data mining also allows grasping useful information from the raw data. This technique uses algorithms to discover relationships within big data stored in databases. The main target is to introduce existing data set in a meaningful way, which analyzes data and eases their complexity while losing as little data as possible (Fayyad et al, 2002).

Effectively employing data mining technique to discover the information entails specific steps to be taken; to arranging the data, employing the process and interpreting the information. Specific steps are further specified by Rokach and Maimon (2008) as follows:

- Development of understanding of the targets of research and its application
- Selection of databases for analysis of data, data cleansing, and transformation
- Selection of appropriate data mining tasks such as regression and clustering
- Selection and employment of algorithms for searching patterns
- Evaluation and interpretation of the mined data
- Presentation, documentation, and usage of the knowledge

Data mining could be a trustworthy and inclusive way to collect information on the inputs, outputs, and impacts of research activities. In research evaluation, data mining technique is generally employed to extract information on research costs and outputs from databases (Boyack and Jordan, 2011). Presently, data mining for research evaluation purposes likely to be used with bibliometric methods to scrutinize citation data and link articles and their academic uptake to research funding.

Most important limitation of data mining comes from the nature of databases to be extracted. Size and number of fields in datasets can surge the research space that should be analyzed which leads requirement of more complex algorithms to gather and analyze data precisely. Furthermore, data mining processes are confronted when they deal with incomplete data and when datasets were not prepared for data mining purposes. Presentation of information gathered through analysis is quite important and complicated issue. Findings should be presented in an accurate and meaningful way for expert interpretation. Effective presentation of analysis of data mining results is an ongoing debate (Fayyad et al, 2002).

Despite these limitations, data mining enables to make sense of complex and unorganized data by analyzing them. Besides, data mining extracts information from data which were initially outside human understanding. Computer-based and automated analyses of data generally result in accurate and reliable analysis results; therefore, it advances the quality of findings. Another advantage of the data mining is related to its potential to decrease the burden of data collection process. Data mining shifts the burden of research evaluation from information suppliers to data collectors (Baker, 2010).

3.5.2.4. Survey

Survey technique can be defined as the collection of data from a sample of individuals through their responses to question sets (Check and Schutt, 2012). Survey technique lets for a variety of methods to employ participants, data collection, and instrumentation. Survey technique can employ quantitative strategies such as questionnaires with numerical items, qualitative strategies such as open-ended questionnaires or both strategies at the same time (Ponto, 2015).

Interviews were already covered before in detail in qualitative evaluation side, for that reason in this section, questionnaires will be scrutinized. Survey questionnaires can be conducted in person, mailed or posted on the internet site. In order to get proper results sample design should be properly prepared. Samples may be randomized and stratified as well as longitudinal or cross-sectional (Ruegg and Feller, 2003).

Survey process generally includes following steps based on Health Communication Unit of University of Toronto (1999):

- Defining the purpose of the survey
- Searching for available resources and deciding on a method
- Developing and writing process of survey
- Pilot testing and revising as necessary
- Selecting the sample
- Collecting and processing data
- Analyzing the results

Surveys are employed in combination with other research evaluation tools and these combinations depend on the purpose of the evaluation. On the other hand surveys are likely to provide information that gives an outline and for this reason, it is generally used with in-depth qualitative techniques. For that reason, surveys can also be employed as an independent technique for research evaluation purposes.

The main advantage of the survey method is providing a cheaper way to collect a significant amount of information about research conducted and its stakeholders, even in its initial phases. In the setting of research assessment, surveys are tended to be better technique when evaluators want to collect information covering all scientists in a particular group as they let standardized information to be collected from a group. Also, surveys can be conducted in a relatively short time period, covering as many respondents as possible and collect the crucial amount of data. Surveys give detailed information about users not available through other resources. Moreover, various audiences can generally understand the results and many individual find statistical findings informative and credible (Ruegg and Feller, 2003).

The main pitfall of the survey technique is that it generally lacks detail and inadequate in terms of depth of data as specified before. Data gathered through survey are generally straightforward and even after analysis, it is hard to find details about subject evaluated. Moreover, responses to the surveys could be a pitfall as well because accuracy and validity issues may endanger the survey process. Closed-ended questions force survey participants to a constant set of answers which are proxies for an actual answer. However, generally participants real answers are difficult to grasp in simple scale set of answers (Bryman, 2015).

3.5.3. Outline of Evaluation Tools

After elaboration of evaluation tools for research evaluation, features of these tools are outlined in Table 1.

Table 1 Outline of the Features of the Evaluation Tools Reviewed

Tool	Brief Definition	Usage	Advantages and Limitations	Quantitative/ Qualitative
Peer Review	Formal evaluation process which uses objective measures and peers to evaluate excellence	Used at all kinds of research activities including article review, allocation of resources etc.	Advantages: Credibility, useful feedbacks Limitations: High costs, low transparency, intrinsic conservatism, subjectivity	Qualitative
Interview	The technique used for collecting information from evaluated parties by asking them questions directly	To get extra and in-depth information	Advantages: Provide new understandings and insights, interactive communication Limitations: High costs and subjectivity	Qualitative
Site Visit	Visit by experts to evaluated party. Generally, consist of a series of meetings with relevant stakeholders.	To get detailed and in-person information about evaluated body	Advantages: Interactive communication Limitations: High costs in terms of time and money for both evaluator and evaluated parties, subjectivity	Qualitative
Case Study	Tool for discovering complicated case based on a complete understanding gained through broad description and analysis of the context	Not used for comparisons, mostly for illustrative reasons	Advantages: Flexibility in usage, efficiency, rich in details Limitations: Very specific, hard to generalize	Qualitative

Bibliometrics	A tool that employs data on numbers of scientific publications, citations, and patents to evaluate the output of research	Measures volume, knowledge transfer and linkages	Advantages: Low cost in terms of time and money, objectivity, easy to understand Limitations: Does not cover impacts, replicate of peer review	Quantitative
Economic Analysis	Analysis for comparing costs and results which assess efficiency. Three types: cost-benefit, cost-effectiveness, and cost-utility analysis	States impacts in financial terms compare costs & benefits and outcomes between alternatives	Advantages: Boosts analytical capability of evaluators, provides financial accountability Limitations: All impacts cannot be captured, assigning monetary values to research	Quantitative
Data Mining	The mechanical technique of gathering data from databases for grasping useful information	Used for collecting information on input and output of research activities	Advantages: Useful for analyzing complex data and extracting information Limitations: Huge and incomplete databases, presentation difficulties	Quantitative
Survey	Collection of data from a sample through their responses to question sets	Provide a broad overview of research activities and comparable data	Advantages: Cheaper and faster way to get data, results are easy to grasp Limitations: Lack detail, accuracy and validity problems	Quantitative

3.6. Impacts of Research Infrastructures

Main impacts of research infrastructures which are defined in current literature will be scrutinized in this section. In the literature, there are several important studies about the impacts of research infrastructures. Zuijdam et al. (2011) examined research infrastructures' impacts in the Netherlands in detail. They also scrutinize the added value and roles of the research infrastructures by using four impact types. Those are namely; scientific, social, educational and economic impacts. In terms of impact types, mainly scientific added value and socio-economic impacts are comprised. Socio-economic impacts are defined in three dimensions. For education effects creation of networks and human capital; for economic effects, the economic value of research infrastructures, for social effects added value for society titles are used in the report.

The OECD Global Science Forum also conducted a study for investigating impacts of the CERN (European Organization for Nuclear Research) which is one of the very successful international research infrastructures in the world. In this study, five impact categories are specified as follows: scientific, economic, technological, educational and social impacts (Michalowski, 2014). By scientific impacts, pure scientific outcomes that intended to develop basic knowledge are covered. By educational impacts training of scientists, engineers, administrators, and other professionals are identified. By technological impacts, developing and diffusing technological innovations are emphasized. By social impacts, education of students and/or teachers and various forms of public outreach is highlighted.

European Commission (2010) prepared an analytical framework for impacts of research infrastructures. These impacts are scientific, technological, economic, social, political and environmental impacts. In this report, scientific impacts are defined as the new knowledge generated and the theoretical progression of science. Technological impacts cover the innovations in the creation of goods and services that come up as spin-off firms from the progress of research infrastructures. Economic impacts are defined as influences on economic growth and employment. Social impacts are basically contributions to society arising from research progress. The impact on political steadiness and unity, resulting from the operation of a research infrastructure and from the advances in science defined as political impact. Environmental impacts are specified as direct effects on physical environment such as water pollution or energy consumption.

Chinese Academy of Sciences (2007) prepared a research report on impacts of research infrastructures and provided six types of impacts namely; scientific, economic development, high technology development, national security, educating scientific talent and international cooperation impacts. Lastly Elias (2009) in his presentation on "Developing Methodologies for the Impact Analysis of Research Infrastructures on Innovation" identifies six broad impact categories of research infrastructures as follows: scientific, economic, social, environmental, political and technological impacts. After covering main studies in the literature, main impacts of research infrastructures are outlined in Table 2 provided below.

Table 2 Main Impacts of Research Infrastructures

Author	Scientific Impacts	Economic Impacts	Social Impacts	Educational Impacts	Technological Impacts
Zuijdam et al (2011)	✔	✔	✔	✔	
Michalowski (2014)	✔	✔	✔	✔	✔
European Commission (2010)	✔	✔	✔		✔
Chinese Academy of Sciences (2007)	✔	✔			✔
Elias (2009)	✔	✔	✔		✔

Main stakeholders of research infrastructures are the academic world, funding bodies, industry, and society as a whole. As stated before, research infrastructures are crucial and expensive investments which are used by academic world and industry. The academic world is supplying researchers to the infrastructures and also enjoying advances which are achieved in infrastructures. For this reason, the most important impact for academia is a scientific impact. Young researchers and students are gaining expertise and experience in infrastructures so educational impacts are also crucial for the academic world. Industry interacts with infrastructures in very different ways such as using them in their R&D activities, establishing alliances, conducting joint projects etc. These interactions result in economic and technological impacts which are most crucial impacts for the industry.

Funding bodies use and have interest in all impact categories but scientific and technological impacts are most vital impacts for them. These impacts are easy to measure and can be grasped in short-term compared to other impact categories. Lastly, the general public is another important stakeholder who pays the taxes and creates the source for these infrastructure investments. For this reason, almost all impact categories are watched closely by the general public as well. Specified stakeholders' interests in infrastructures are outlined in Table 3.

Table 3 Main Stakeholders' Interest from Impacts of Research Infrastructures

Interest From	Scientific Impacts	Economic Impacts	Social Impacts	Educational Impacts	Technological Impacts
Academic World	***	*	*	***	**
Funding Bodies	***	**	*	**	***
Industry	*	***	*	*	***
Society	**	***	***	***	***

3.7. Main Indicators of Research Infrastructure Impact Types

In this section, main indicators of impact types will be elaborated. Some indicators employed in research and research infrastructure evaluation can be included in several impact types at the same time. For example, patents are results of scientific activities which have characteristics that contain both economical and technological impacts. With the indicators elaborated in this section, both qualitative and quantitative data can be gathered so both qualitative and quantitative evaluation methodologies can use these indicators in their analysis.

3.7.1. Scientific Impact Indicators

As can be observed in Table 2, the scientific impact is the primary impact specified in evaluations. It is involved in all of the five classification systems of research infrastructure impacts in the existing literature. Scientific impact is the major and direct function of research infrastructures in terms of purposes for building and its effects determine the functional level of other impacts.

Scientific impacts of research infrastructures have not been explored comprehensively. Although impacts and characteristics of research infrastructures have been scrutinized and studied, these studies have been in the initial phase. Existing literature is focusing predominantly on the short-term and direct scientific impacts and extended long-term impacts are either unregarded or only mentioned superficially (Qiao et.al, 2016).

The research results today have an effect on the development of knowledge of future. These results can be formulated as facts, models, theories, the development and expansion of domains and disciplines. Scientific impacts generally measured with research outputs which include bibliometric indicators such as articles, citations, and patents (King, 1987). Moreover, research outputs other than articles such publications (books, working reports, theses etc.), scientific projects, prizes, and events are also included in the indicators. In addition to the quantity of these indicators, quality of the indicators is also considered in the evaluation systems (Godin and Dore, 2014).

According to OECD presentation on Socio-Economic Impact of Research Infrastructures, scientific data are regularly gathered and broadly acknowledged. In addition to indicators specified above, OECD also specifies collaboration of infrastructure researchers and users, new frameworks derived from infrastructure and PhD and Post-doc applications and positions as scientific impact indicators (Alagna, 2015).

3.7.2. Economic Impact Indicators

Economic impact and its indicators are probably the second crucial element in research infrastructure evaluation due to accountability issues. As in scientific impact case, economic impact is also included in all elaborated literature. Research infrastructure investments are quite expensive and risky investments. Return on investment is strictly needed and evaluated by the authorities by economic impact indicators. Economic impacts of research infrastructures can be defined as the short-term effects on economic growth, employment arising from the activities of a research infrastructure as well as medium to long-term contributions to the economy regarding technological improvements made by the operations of a research infrastructure (European Commission 2010).

Main indicators of economic impact evaluation include standard economic turnover data, technological innovation, spin-off firm formations, procurements, new jobs created, value and number of access, patents and licensing (Alagna, 2015). Segers (2015) further stated industry collaborations and industry participation in infrastructure events, co-funding and training events as indicators of economic impact. The volume of exports and high technology commerce are also emphasized in literature as indicators (Godin and Dore, 2014).

3.7.3. Social Impact Indicators

The wider impacts that research infrastructure has on society is the least studied area because such impacts are hard to trace and measure. While direct relationships are almost impossible to establish, social science methods can be employed to estimate the potential contribution of research infrastructures (Griniece et al., 2015).

Social impact discusses the effect of knowledge has on well-being, and on the behaviors, activities and actions of society. For individuals, social impact means welfare and life quality. For a group of individuals' social impact can contribute to shifting understanding of society or assist modernize their way of conducting business (Godin & Dore, 2014). The social impact of research infrastructures can be also stated as the contribution to common welfare stem from advancement made in research, which arises from the research process and its contribution to improving the quality of life of the society (European Commission, 2010).

In Social Impact Assessment Methods through Productive Interactions project (Spaapen et al, 2011), which is one of the leading studies in social impact evaluation of research, social impacts are identified and evaluated by tracking productive interactions. Productive interactions are defined in the study as interactions between scientists and societal figures in networks in which knowledge is formed and valued that are at the same time scientifically and socially related and robust. In this study, main focus of social impact evaluation is on interactions between researchers and stakeholders rather than results and outputs of research activities.

Due to difficulties in measuring social impacts, indicators in this area are wide and diverse from contribution to overcoming grand challenges to scientific dissemination and education. Some concrete examples are a number of open visits to research infrastructures, training events to all layers of society including high school students to industry participants, public relations activities of research infrastructures and general public awareness about research performed and results of research (Alagna, 2015, Segers, 2015).

According to Evaluating Research in Context Project (Van den Brink et al, 2010), aiming developing methodologies and indicators for assessing the societal relevance of research; main aspects of societal relevance are dissemination of knowledge, interest of stakeholders and use of results. Under these aspects, for different research areas, non-academic publications, PhD’s in industry advisory and consultancy activities, mobility of graduates, collaborations with societal stakeholders, income from use of results, spin-offs with industry contacts are counted as indicators of social impact.

3.7.4. Educational Impact Indicators

The educational impact could have been positioned under the scientific impact but due to its importance in missions of research infrastructures, it has been placed independently. Research infrastructures are magnets for qualified researchers. They have a considerably high effect on mobility issues.

In order to capture the impacts on education aspect, it is crucial to understand structural moves in organizational behavior that arise thanks to the research infrastructure. Monitoring indicators on the number of researcher and students attracted, particularly from abroad, can show the dynamics of knowledge exchange that specific research infrastructure is enabling (Griniece et al., 2015). Further, it is important to capture how training on research infrastructures is carried out and what are the practices for informal knowledge transfer. Educational activities structured for both researchers and society needs to be tracked under this impact type as well.

3.7.5. Technological Impact Indicators

Product, process innovations and technical knowledge are forms of technological impact of research infrastructure activities (Godin and Dore, 2014). Technological impacts of research infrastructures are also defined by European Commission (2010) as the innovations in the production of goods and services that emerge as spin-offs from the development of research infrastructures or the benefits arising from the developments in scientific know-how that originated from their operations. In the light of these concepts it can be claimed that technological impact indicators have strong relationships with scientific and economic impact indicators.

Main indicators of technological impacts of research infrastructures are all intellectual rights, especially patents, spin-off firms and links to private sector such as industry projects (European Commission, 2010, Godin & Dore, 2014).

After analyzing all impact indicators, main indicators of research infrastructure impact types are outlined in Table 4 below.

Table 4 Main Indicators of Research Infrastructure Impact Types

Impact/ Indicator	Scientific Impacts	Economic Impacts	Social Impacts	Educational Impacts	Technological Impacts
Main Indicators	Publications, Citations, Patents, Scientific Projects, Prizes, Scientific Events …	Spin-off Firms, Private Sector Collaborations, Licenses, New Jobs Created, Value and Number of Access …	Number of Open Visits, Training Events, Non-academic Publications, Advisory and Consultancy Activities …	Number of Researcher Attracted, Training Activities, Informal Knowledge Transfer Mechanisms …	Intellectual Rights, Spin-off Firms, Links to Private Sector ...

4 INTERNATIONAL RESEARCH INFRASTRUCTURE EVALUATION PRACTICES

The global competitiveness of a country's research community is gradually more influenced by the quality and status of its research infrastructures (Hungary National Research Development and Innovation Office, 2014). For that reason, national research infrastructures' performance and quality are measured by the national authorities.

Evaluation systems for research infrastructures differ in world examples. In some countries, there are systems in which each institution has assessed its own infrastructure, while in some other countries there is an evaluation system that includes all infrastructures. Key element for decision-making in evaluation design is whether the evaluation should be performed at the institutional or national level. This guides the choice of methodologies and indicators, depending on needs and objectives (Mahieu and Arnold, 2015).

Since the implied model in our Law No. 6550 addresses all the research infrastructures within Turkey, the country systems described in detail in this section have been selected among national level evaluation practices rather than institutional level evaluation practices. For example, in United States ministries and public institutions are evaluating their own research infrastructures and therefore they are out of scope for that reason.

Selected countries include France, Germany, Australia, Portugal, and the Czech Republic. The first three mentioned countries are developed countries. Czechia and Portugal have been included in this study due to their relative similarities with Turkey in terms of research landscape and development level.

The institution, which is responsible for the evaluation of research infrastructures in France, aims to be the best in the world in its own field and performs evaluations in other countries as well. There is a single set of criteria in the evaluation system, but each infrastructure is not expected to be successful at each criterion. Mainly informed peer review and site visit tools are employed in evaluations. Quality orientation is really essential in this evaluation system. All infrastructures are covered in the evaluation. At the end of the evaluation, no grades are given to the related infrastructure; only the report is written and transmitted to the authority. There is an interactive system in both the selection of the experts and the preparation of the report processes. In both processes evaluated infrastructure’s opinion is taken. Since the infrastructures required to be evaluated in Turkey have different missions, ages and from different fields, it is concluded that the French system contains elements that can address the needs of the system in our country.

The Portuguese evaluation system is a two-stage process and European Science Foundation (ESF) is also taking part in evaluations. At the end of the evaluation process, infrastructures are categorized into 36 categories according to the evaluation score (weak / medium / good / very good / excellent / extraordinary), size and intensity. In Portugal, first stage evaluation is made through document reviews and panel evaluation methods, in second stage site visits are conducted and only written report is prepared at the end of the site visit. Then this report evaluated in the field panel and a final grade is given to the infrastructure. Application details and sensitivity throughout the evaluation process is valuable and with these characteristics, this system can also be taken as an example for Turkish system.

The German evaluation system is a two-phase process that uses document reviews, panel evaluations and interviews. At the end of the evaluation process, infrastructures are graded as outstanding, very good, good, satisfactory and sufficient in each of the four evaluation indicators. While grading, infrastructures are subject to the binary comparison which is the most distinctive characteristic of this system.

The Australian system is a two-stage and elimination based process. The first stage assessment is carried out by a panel of external and local experts through the documentation. The same panel conducts interviews with infrastructure representatives in the second phase. The Australian system has elimination in every phase of the evaluation which creates a difference with other country examples.

The Czechia has passed the peer-review method in 2014 after examining world examples. The evaluation process has two stages; in the first stage, formal and contextual evaluation is performed, in second stage in-depth evaluations are conducted. In both stages, documents prepared by infrastructures are reviewed by experts. Although the chosen method is welcomed in Czechia, the task of assigning a lot of foreign specialists to the system has attracted a reaction. Stakeholders in the country's R&D ecosystem are criticizing Czech authorities for implementing their systems in other countries directly in Czechia without regard for their country's dynamics. Distinctively in Czech system, external user opinions are gathered by surveys.

4.1. France High Council for the Evaluation of Research and Higher Education (HCERES)

High Council for the Evaluation of Research and Higher Education (HCERES) is performing an evaluation of research and higher education institutions and research infrastructures in France. HCERES has an independent status by law, allowing it to perform its activities without having to be exposed to any pressure from the government, research infrastructures or any other parties. In 2015, HCERES has evaluated approximately six hundred research infrastructures with the help of over three thousand experts. Main aims of the evaluation are outlined by the institution as; helping research infrastructures on the identification of potential areas of development, providing information to funding bodies and providing information to the researchers and students about the evaluated body (Cappy, 2015).

4.1.1. Overall Evaluation Process

HCERES uses mainly qualitative evaluation methodology with a site visit and peer evaluation and these qualitative evaluations are supported by quantitative data which can be named as bibliometrics. Evaluation process starts with a declaration of which research units are to be evaluated by HCERES. After that, application forms are completed by the infrastructures and delivered to HCERES. Council shares these application forms and necessary information about the infrastructures with experts.

A site visit is conducted by the expert committee who review the documents supplied by the infrastructures. After the visit, the draft report is written by these experts without any grading. These reports are validated by HCERES after site visits and shared with research units for any objection. Evaluation reports are also published on the HCERES website. For these reasons, it can be stated that overall evaluation system is quite transparent.

4.1.2. Evaluation Indicators

HCERES uses six evaluation indicators with observable facts and quality indicators connected to each of the six indicators. These evaluation indicators and sub-indicators are shown in Table 5. Observable facts are qualitative indicators that reflect outputs and results of the research or managerial activities. Observable facts can be interpreted as bibliometric data. Even though quantitative indicators can be obtained in some research activities, these indicators are used only as a support in the evaluation process (HCERES, 2014). In order to assess these quantitative indicator data in terms of quality, HCERES specifies quality indicators as a component of a peer evaluation. Briefly, experts conduct peer evaluation by considering both observable facts and quality indicators. In this respect, HCERES utilizes both qualitative and quantitative evaluation and uses both peer review and bibliometrics. So it can be stated that HCERES uses mainly informed peer review methodology which is a hybrid model that employs both peer review and bibliometrics.

Table 5 HCERES Evaluation Indicators

HCERES Evaluation Indicators	Observable Facts	Quality Indicators
Scientific Production and Quality	Publications, Lectures, Other scientific reports, Formulation of instruments, resources and methodologies etc.	Originality, Scope, Importance, Presence, Reputation, and Selectivity of research conducted etc.
Academic Reputation and Appeal	Projects, Collaborations, Participation in networks, Researchers and Students, Prizes etc.	Role and responsibility for projects and collaborations, Reputation of prizes, High standard of researchers and students etc.
Interactions with the Social, Economic and Cultural Environment	Patents, Licenses, Prototypes, Spin-Offs, Professional or technical articles and Study reports, Software and models etc.	The originality of methods and products transferred, Quality and success of dissemination, Relationship with recent scientific knowledge etc.
Organization and Life of the Institution	Presence of objectives, Scientific coordination, Decision-making process, Human resource policy etc.	Achievement of strategic objectives, Existence of scientific coordination structure, Clarity of communication etc.
Involvement in Training Through Research	Presence of Trainees, Thesis examined, Supervision rate of students, Publications, Designed training modules etc.	Quality of supervision and thesis, Certification of training, Reputation of educational outputs, Student participation in the research unit etc.

Strategy and Research Perspectives for the Next Five Years	The existence of a scientific policy, Strategy to achieve research objectives, Articulation of policies and strategies into concrete actions etc.	Originality and coherence of research perspectives, Synergy, Credibility of strategy, Richness and openness to partnerships, Self-evaluation capabilities etc.

Resource: The HCERES Standards

4.1.3. First Stage of Evaluation

Infrastructures are supplying necessary data on both observable facts and quality indicators by the application forms supplied by HCERES. After application forms received to HCERES, on average five experts are appointed to each application. A number of experts change as the infrastructure's size changes. These experts conduct pre-evaluation on applications of the research infrastructures and prepare questions for the second stage of the evaluation. (HCERES, 2010, HCERES, 2016).

The expert selection process starts with an expert proposal by HCERES board, research institution directors, evaluated research infrastructure's directors and other evaluation bodies. Proposed experts constitute a pool of potential experts. Currently, HCERES has over ten thousand potential experts in the potential expert pool. These experts can be researchers or technicians in the field of the research unit, from the academic world or private sector. Experts may also be either French or foreigner. HCERES chooses academic experts by evaluating candidate's expertise in the field of research infrastructure, experience in the research institution and in international level. In the nomination of non-academic experts, at least 5-year experience is needed in the evaluation field (HCERES, 2010).

As mentioned before, both scientific and managerial performances are assessed and for that reason when choosing experts HCERES looks for at least one of the qualifications specified below.

- Experience at research organization management bodies
- Skills in the relevant field
- Managerial skills
- Business competence

Furthermore, at least 20% of the experts that are appointed by HCERES are working abroad. This point shows us the importance of international point of view of HCERES. When one of the experts nominated from the pool and participate in the evaluation process, he or she can be registered expert for 4 years. In order to be a registered expert, relevant HCERES department should suggest to HCERES president and approval of HCERES president is needed (HCERES, 2010).

4.1.4. Second Stage of Evaluation

Site visit stage is started after the examination of the applications of the research infrastructures by the experts. Experts prepare some questions for a site visit on the basis of suggested evaluation report templates by HCERES. After individual reviews and preparation, these experts constitute an expert committee. After the committee constituted, consultation with infrastructure’s director is arranged and opinions and concerns of the infrastructure obtained. After all these processes, HCERES appoint one chair to every expert committee within the committee members. The committee chair directs the debates and discussions in the site visits (HCERES, 2015).

Site visits start with an expert committee meeting in which evaluation details are discussed and role distribution is made. After that, the open session is organized with the infrastructure director and the necessary researchers (Guthrie et al., 2013). In this session less than one hour is required to highlight the most striking scientific facts and / or important developments achieved by the infrastructure. Next, interviews with researchers, technical teams, infrastructure board and PhD students are made (HCERES, 2015).

After the site visit, an evaluation report is constituted collectively by the expert committee. The evaluation report, which is the main outcome of the entire evaluation process, consists of two main sections: one with general evaluations and one with detailed evaluations. General evaluation section simply outlines the detailed evaluation chapters and prepared as an executive summary. Detailed evaluation chapters have a critical point of view and specify strengths, weaknesses, opportunities, and threads of the research infrastructure in each six evaluation criteria. Moreover, recommendations are also welcomed in the report (HCERES, 2016).

Because of lack of grading in the evaluation process, expert reports need to be quite detailed and informative. When evaluation reports are read which are available on HCERES webpage, evaluations reports exceed one hundred pages. Also, it can be inferred that primary purpose of the HCERES evaluations is to support the development of infrastructures and the realization of their strategic priorities.

In previous evaluation system, infrastructures are graded as A+, A, B and C in each" evaluation criteria but this process is terminated in last evaluation methodology. The main reason for termination is that grades are much more visible and evaluation reports are not taken into consideration compared to grades. Moreover, in 4 scaled grading, showing and explaining differences between great numbers of infrastructures is quite difficult. Two different infrastructures might get the same grading but there can be still some quality differences between the two. For these reasons, HCERES prepares only evaluation reports at the end of the evaluation process.

4.2. Portugal National Research Council (FCT)

National Research Council of Portugal (FCT) is the national funding agency that supports science, technology, and innovation in all scientific areas under the charge of the Ministry for Science, Technology and Higher Education of Portugal. FCT has been conducted research infrastructure evaluations since 1996. In 2013, main evaluation methodology is updated and still, this version is in use. Currently, three hundred research infrastructures are evaluated by the FCT regularly. The main aim of the evaluations conducted by the FCT is funding of the research infrastructures (FCT, 2017).

4.2.1. Overall Evaluation Process

FCT evaluation system mainly focused on funding and evaluation phases are designed accordingly. The evaluation system consists of two main stages; first stage evaluations for basic funding and second stage evaluations for strategic funding. The evaluation process begins with the filling of the online application form, which is used in both phases, by the research infrastructures. The first stage of evaluation includes evaluations of experts on the application form filled by research infrastructures.

After first stage evaluations, basic funding decisions and qualified infrastructures for the second stage are announced. In the second stage of the evaluation site, visits and interviews are conducted for strategic funding purposes. Evaluation criteria used in evaluations and details of the two main evaluation processes are given below.

4.2.2. Evaluation Indicators

In the first and second evaluation stages, infrastructures are assessed with the indicators and weights specified in Table 6. As can be seen in the weight sections, the first four indicators are employed on the first stage and all the indicators are included in the second stage. The changes in the weights of the evaluation criteria in the second phase are determined by research type of research infrastructures which are basic or applied research. For example, in an infrastructure that conducts basic research, weights of 35% in indicator A and 5% in indicator E can be used in assessments (FCT, 2013).

When indicators are analyzed in terms of impact and input considerations, by criteria A, B and E impacts are measured, by criteria B, C, D, and E generally inputs are assessed. Furthermore, FCT is also defining A, B and E criteria as performance indicators and B, C, D and E criteria as strategic program indicators.

Table 6 FCT Evaluation Indicators

FCT Evaluation Indicators	Sub-Indicators	First Stage Weight	Second Stage Weight
A. Productivity and contribution to the National Scientific and Technological System	Research outputs (Published articles, patents, models, books, conference proceedings, new products, processes etc.) and contribution to the national S&T system, dissemination of activities, training of researchers	25%	%20 - %35
B. The scientific and technological merit of the research team;	International merit, scientific productivity, competitiveness, skills, and composition of the research team	25%	20%
C. Scientific merit and innovative nature of the strategic program	Impact, originality, and relevance of the strategic program in terms of scientific, technological and cultural aspects	25%	20%

D. Feasibility of the work plan and reasonability of the requested budget;	Organization and management of objectives and resources, adequacy of the proposed budget	25%	20%
E. Impact of the scientific, technological and cultural output	Impacts on economy, knowledge & technology transfer activities and dissemination	0%	%5 - %20

Resource: FCT Evaluation Guide

4.2.3. First Stage of Evaluation

In the first stage of the evaluation, peer evaluation methodology is used. Online application forms contain information about first four criteria. These forms are filled and send to FCT by research infrastructures. FCT conducts peer review with expert panel evaluation regarding these application forms.

Evaluations are performed by four evaluation panels which are specialized in four major scientific domains and these panels have several working groups. Evaluation panels have generally 15-20 experts and every working group have 4 experts (Deem, 2015). One coordinator is assigned to each working group. These coordinators are also participating in the second stage of the evaluations. Initially, for each infrastructure application, maximum 5 individual evaluations are performed. After that, each working group assesses at least 10 individual reviews. For interdisciplinary infrastructures, more than one working group assessment is performed. After all these evaluations, working groups are preparing consensus report which includes grades, suggestions and other details. These reports are shared with infrastructures and infrastructures have right to object these evaluation results (Allspach et al, 2015).

During the first stage of evaluations, the 5-point scale is used in which 5 is the maximum and 1 is the minimum score for each indicator. After grading each indicator, the cumulative score of the first evaluation stage is calculated. According to cumulative scores, research infrastructures are subject to qualitative overall grading. Details of qualitative overall grading are shown in Table 7. Research infrastructures need to have at least 15 points from all indicators qualifying the second stage of the evaluation.

As a result of evaluation of the sustainability of activities of infrastructures 5.000 – 400.000 EUR basic funding is allocated for each infrastructure. Basic funding is allocated according to three parameters namely; performance evaluation results, laboratory intensity, and dimension of the infrastructure. According to performance evaluation results, infrastructures which are graded good or higher than good are qualified for the basic funding.

Laboratory intensity level is measured by panel members by considering laboratory equipment and dimension of the infrastructure is calculated based on the number of doctorate integrated members. As can be inferred from components of the basic funding, the main reason for the funding is supporting infrastructures so that they can continue their activities. The first stage of evaluation is also a preparation for the second stage of the evaluation and provides inputs to it (FCT, 2013).

Table 7 FCT Qualitative Overall Grading

Grade	Description	1st Stage Cumulative Score
Exceptional	R&D Unit recognized as an international reference for its scientific and technological output and exceptional contributions to its area of research	≥ 15[1]
Excellent	R&D Unit distinguished by the high quality and international merit of its scientific and technology output and with significant contributions to its area of research	≥ 15[1]
Very Good	R&D Unit with high quality and national merit and with significant contributions of international relevance in its area of research	≥ 15[1]
Good	R&D Unit with quality at the national level, reduced internationalization and some contributions to its area of research	< 15 > 12[2]
Fair	R&D Unit without significant contributions to its area of research	≤ 12 > 11[3]
Poor	R&D Unit without contributions to its area of research and with other weaknesses.	< 11

1 Additionally the application must score at least 4 points in each of the ratings of criteria A and C, and it must also score at least 3 points in each of the ratings of criteria B and D.
2 Additionally the application must score at least 3 points in any of the four evaluation criteria ratings.
3 Additionally the application must score at least 3 points in each of the ratings of criteria A and C, and it must also score at least 2 points in each of the ratings of criteria B and D.

Resource: FCT Evaluation Guide

4.2.4. Second Stage of Evaluation

Site visits and interviews are conducted in the second stage of the evaluations. Both site visits and interviews are performed by expert panels. Six disciplinary and one transdisciplinary panel are set up for evaluations in this stage. Each panel is constituted by 4 or 5 internationally recognized experts. Coordinators of the working groups are also members of these evaluation panels. This connection is providing input and knowledge to the second stage of the evaluation. Panels are performing maximum 10 site visits or interview (Allspach et al, 2015).

One of the panel experts is assigned as panel chair. He or she has the duties of coordinating and moderating the site visits and interviews, elaborating the final report and delivering the results of the evaluations. FCT also employs observers for this stage of evaluations. Observers are not attending evaluations and they are not part of a panel. In order to ensure the consistency of the evaluations, they are working closely with panel chairs (Deem, 2015).

Site visit process starts with the announcement of the first stage evaluation results. Infrastructures need to prepare themselves for the critics and comments in the first evaluation stage. Infrastructure managers are responsible for the organization of the site visit. Each infrastructure is allowed maximum three hours of communication with the panel. The first one hour is for the presentation of the infrastructure and debates on this presentation, the next one hour is for visiting the research site and the last one hour is for meeting with researchers and students (Allspach et al, 2015).

During the site visits, infrastructures are allowed for maximum 15 minutes presentation. Presentations should cover presentations of research teams, strategic program and answers to specific questions raised by panel members. After the presentation is accomplished, research facility and equipment are visited by panel members. A site visit is finished with meetings and discussions with researchers and students. Meetings with students and young researchers are performed without the presence of infrastructure manager in order to the facilitation of the students (Allspach et al, 2015).

After the site visit, panel meeting is organized for discussion and integration of the first stage evaluation results and main conclusions from site visit evaluations. Before the meeting, all documents and necessary information are supplied for panel members about the performed evaluations. The panel meets physically, discusses all applications and reaches consensus on the final report. The final report includes the grading, budget recommendations, brief comment on the application, which will be shared with research infrastructure and secret comments for FCT. Moreover, the final report needs to contain strengths, weaknesses, opportunities, and threats of each research infrastructure. This final report finalizes the evaluation campaign, provides feedback to the applicants, and also be a starting point for mid-term evaluations.

In grading, a 10-point scale is used in which 10 is the maximum and 1 is the minimum grade. The qualitative grade is based on the panel's own judgment of the general merit of each infrastructure after the site visit. This grade is not the direct output of quantitative algorithm based on the weights attributed to each evaluation criteria. If considered essential, the observer for the second stage of the evaluation has a right to organize a final meeting with all the evaluation panel chairs in order to confirm and ensure the uniformity of the qualitative grading of all the assessed research infrastructures. Evaluation results have 6-year validity and after 3 years, mid-term evaluation is performed (FCT, 2013).

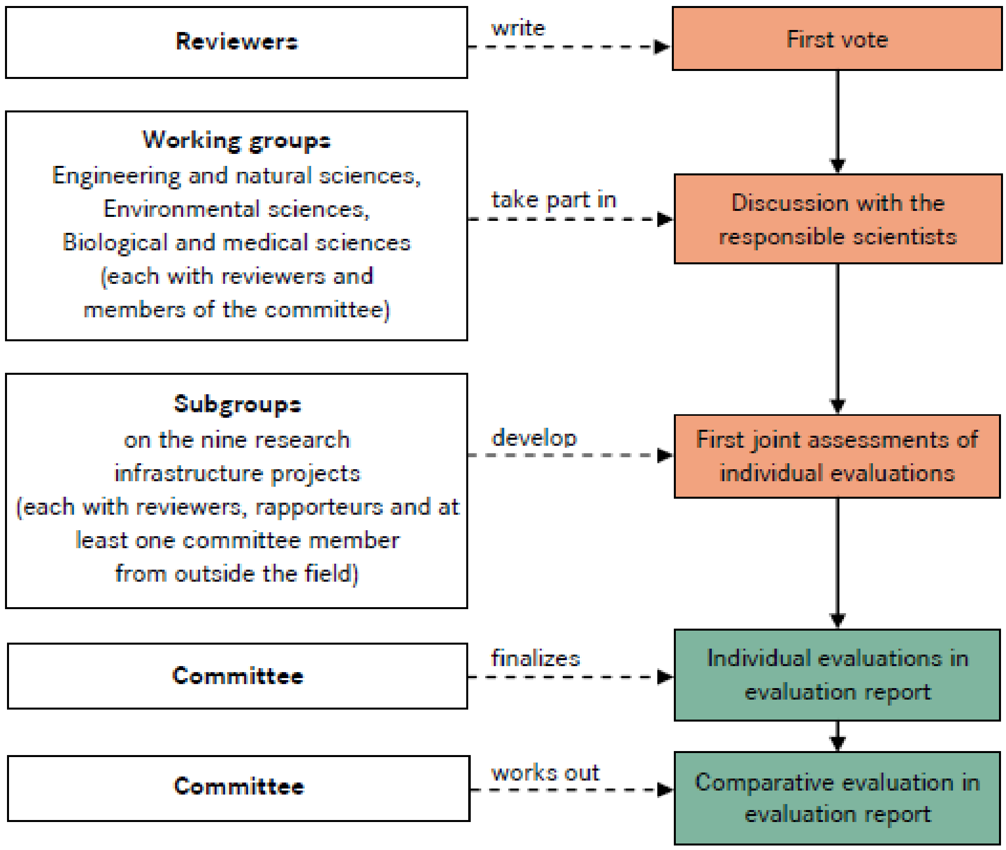
4.3. The German Council of Science and Humanities (WR)

The German Council of Science and Humanities (WR) is one of the main science policy advising councils in Germany. Council's general duties include preparing reports and suggestions on scientific institutions' performance and funding as well as reporting on general issues in science and higher education system. By depending on its duties, the council is mandated by German Education and Research Ministry for evaluation of large-scale research infrastructure projects for a national roadmap to carry out the science-driven evaluation. This evaluation model is designed in 2011 and tested immediately, and successfully completed in 2013. By this evaluation model, 9 research infrastructures which are submitted by German Education and Research Ministry are evaluated in 2013 as a pilot phase and 12 research infrastructures are evaluated in 2017 (WR, 2017).

4.3.1. Overall Evaluation Process

The scientific evaluation process is controlled by evaluation committee which is composed of council members who are appointed by WR and international experts. The pilot evaluation process is constituted by two main parts namely; individual and comparative evaluation stages. Individual evaluations are performed by international experts and comparative evaluations are performed by WR members and international experts. All evaluations are finished within 9 months (WR, 2013). The overall evaluation process is outlined in Figure 1.

Figure 1 WR Overall Evaluation Process



Resource: WR Report on the Science-Driven Evaluation

First of all, each infrastructure is subject to extensive individual evaluation by international experts on the basis of evaluation indicators. Moreover, the position of the infrastructure within its counterparts in the international area is determined. The committee then makes comparative assessments on the basis of individual evaluation reports (WR, 2017).

4.3.2. Evaluation Indicators

Table 8 shows the four indicators and related sub-indicators of evaluation in which individual and comparative assessments are made.

Table 8 WR Evaluation Indicators

WR Evaluation Indicators	Sub-Indicators
A. Scientific Potential	Impact of the research areas, Collaboration opportunities, Created innovations
B. Utilization	Origin and size of user groups, Availability of access procedures, International users' rate
C. Feasibility	Reliability of governance, Technical and personnel requirements
D. Relevance to Germany	Impact on the visibility and attractiveness of German science, Contribution to Germany and German education

Resource: WR Report on the Science-Driven Evaluation

Under the scientific potential indicator, infrastructure is assessed in terms of the impact of the opening of new research areas or the development of existing ones. Moreover potential users' quality, interdisciplinary collaboration opportunities, the importance of created innovations etc. are also questioned in this indicator. Under usage indicator; the origin and size of user groups and the availability of infrastructure access procedures are assessed. Also, international users' rate, industry's interest in infrastructure and finance of operations are also evaluated by this indicator.

Feasibility indicator considers the institutional and personnel requirements for institutions that will host the infrastructure as well as the technical requirements for the infrastructure are assessed. Also, the presence of reliable governance concept, enough qualified technical staff and researchers and employment and training of new generation researchers are questioned by this indicator. The relevance of the infrastructure to Germany indicator looks for the interests of Germany, the impact on the visibility and attractiveness of German science. Furthermore, the contribution of infrastructure to the present strong and weak fields of Germany, contribution to the education activities in Germany universities and international collaborations of German researchers are also scrutinized with this indicator (WR, 2017).

4.3.3. Individual Evaluations

First of all, for each infrastructure project, three international experts are assigned for evaluating the technical specifications of the infrastructure. On the basis of the evaluation indicators, documents are filled by the infrastructure according to a certain form and these forms are evaluated by three international experts. After that appropriate individual evaluation reports are written in a certain form.

Secondly, field-based working groups consisting of committee participants and international experts meet with scientists who are responsible for infrastructures. This view provides a more detailed understanding for the evaluators. This stage can be named as interview stage.

Finally, three international experts, a committee member (as a reporter) related to the field and a non-field committee member constitute a subgroup and write a joint evaluation report in the light of individual reports and interviews. This report is the basis for the subsequent comparative committee evaluation (WR, 2013).

4.3.4. Comparative Evaluations

Research infrastructures should not be considered to be inadequate by international experts specifically in terms of scientific potential. If they are assessed as inadequate they cannot qualify for the comparative evaluation process and for these infrastructures only individual evaluations are performed. Projects that qualify this stage are rated according to one to five-star scales for each indicator as a result of the evaluation committee's assessment. The number of stars on the scale is expressed in words (WR, 2013):

- (*****) Outstanding
- (****) Very Good
- (***) Good
- (**) Satisfactory
- (*) Sufficient

In the comparative evaluation process, infrastructures are subject to a binary comparison on a case-by-case basis. If one infrastructure takes one more star from the other, it is showing that the highly noted infrastructure is better at that indicator. However, this does not give an idea about the magnitude of the difference. At the end of the process, each infrastructure project is not given a certain total score and it is expected that the weighting of the dimensions will be made according to the policy priorities of the relevant decision maker (WR, 2017).

4.4. Australian Research Council (ARC)

Australian Research Council (ARC) is a federal agency who is responsible for consulting the Australian government in research topics. One of the main duties of ARC is the evaluation of Australian research infrastructures'. Main aims of the evaluation are stated as connecting existing Australian scientific strengths and constructing critical mass with new capacity for inter-disciplinary, cooperative methods to address the most stimulating and important scientific problems and creating infrastructures that have an impact on the wider society through collaboration with universities, public institutions, industry, and non-profit sectors.

Research infrastructures are evaluated and funded on the basis of scientific excellence and their contribution to the social, economic and cultural developments of Australia (Barros, 2013). At the end of the evaluation process, ARC sends a recommendation to the Ministry of Education and the Ministry is involved in the funding of research infrastructures' (ARC, 2015a). ARC admits research infrastructure applications and eliminates the applications throughout the evaluation process. In 2014 evaluation campaign, one hundred applications were evaluated and only twelve applications are selected and funded (ARC, 2015b).

4.4.1. Overall Evaluation Process

The evaluation process consists of competitive two stages. In the first stage, declaration of intent is requested from the infrastructures for preliminary evaluation. These declarations are evaluated by the Selection Advisory Committees consisting of 12 local and international experts in each discipline. Infrastructures that qualify to participate in the second stage are determined after the evaluation performed in the first stage (ARC, 2015a).

In the second stage, candidates are required to make full application through the application form. Full applications are evaluated by the expert evaluators and the Selection Advisory Committees. After that, research infrastructures are invited for the interview subsequently. Interviews are held with the infrastructure officers. Finally, evaluation results and recommendations are presented to the Ministry of Education. The ARC can eliminate applications at any time during the evaluation process. Reasons for elimination include not satisfying eligibility criteria and providing deceptive or lacking information (ARC, 2015a).

4.4.2. Evaluation Indicators

ARC defines research impacts as the evident influence that research makes to the society, economy, national security, culture, services, and quality of life, beyond contributions to science. Further ARC states that these impacts are all included in objectives and selection criteria (ARC, 2015b).

Table 9 shows the ARC evaluation indicators and weights. In the preliminary evaluation phase, only the first two of the indicators are used. All indicators are employed in the second stage of the evaluation. The reason for this practice might be a need for quick and effective evaluation of the preliminary evaluation which can shorten the overall evaluation time.

While the weight of the "quality of the research program" and the "researchers" are equal in the preliminary evaluation stage in 2014, the size of the "quality of the research program" is increased to 70% in the preliminary evaluation performed in 2017 (ARC, 2017).

Table 9 ARC Evaluation Indicators

ARC Evaluation Indicators	Sub-Indicators	Preliminary Evaluation Weight	Full Evaluation Weight
A. Research Program- Quality and Innovation	Innovation activities, Collaborations	70%	20%
B. Investigators	Capacities and contributions of the researchers	30%	20%
C. Institutional Support	Host institution support and relations	0%	20%
D. Governance, Leadership, and Mentoring	Management, Strategic Plan, Mentorship to graduate students	0%	20%
E. Outcomes and Linkages	Contribution to the country, Outputs	0%	20%

Source: ARC Funding Rules for Centres of Excellence

In research program- quality and innovation indicator; innovation, conceptual framework, and collaborations of the infrastructures are questioned. In investigators indicator, capacities and contributions of the researchers who are part of an infrastructure are examined. In the aspect of institutional support, the host research institute's relations with infrastructure and issues related to its support are elaborated (ARC, 2017).

In governance, leadership and mentoring indicator, infrastructure management is questioned in various competencies (preparation of the strategic plan, mentorship to graduate students, etc.) Finally, in the indicator of outcomes and connections, the contribution of the infrastructure to the country and outputs are assessed (ARC, 2017).

4.4.3. First Stage of Evaluation

During first application process, inclusive summary of research infrastructure is introduced to the ARC by the research infrastructures. Applications should cover all the information needed for evaluation. Applications are uploaded to ARC's online research management system and in the system content, format requirements are controlled. For each application, ARC assigns members of Selection Advisory Committees and these evaluators are performing evaluations and preparing reports on the applications in light of the evaluation indicators (ARC, 2015a).

Applications are assigned to two types of evaluators in general, at least two detailed evaluators and two general evaluators. Detailed evaluators are selected from the Australian and international research community. Detailed evaluators perform comprehensive evaluations of applications and these evaluations are taken into consideration by general evaluators. General evaluators are usually selected from ARC College of Experts, which is a pool of experts of international standing drawn from the Australian research community: from academia, the private sector and public research organizations (ARC, 2015c).

Evaluator's assignments are made by research management system of ARC. The system produces word cloud of the proposals based on title and summary of the proposal and codes. After that, system generates evaluator suggestions by looking into evaluator's codes, expertise, and resume. However, assignments are not made automatically by the system. The information from the system is used as input and final decision is made by ARC officials (ARC, 2015c). Throughout the evaluation process names of the evaluators are not shared with research infrastructures (ARC, 2015a).

Evaluators are preparing reports and comments about research infrastructures and these reports and comments are shared with infrastructures. Research infrastructures have the opportunity to react to the evaluator's reports and to request for additional information about evaluations performed. After evaluations, ARC ranks each application by taking into account evaluation indicators relative other applications. Lastly, shortlist of selected applications is announced by the ARC and call for full applications are opened (ARC, 2017).

4.4.4. Second Stage of Evaluation

Full applications can be submitted only when invited by the ARC and should contain all the information needed for the second stage of the evaluation. Applications are evaluated by the Selection Advisory Committees and independent evaluators employed by the ARC. Ranking according to merit is also repeated in this stage of evaluation. After this ranking, ARC invited applicant research infrastructures for interviews. Interviews are done with infrastructure directors and key researchers. After the interview process, eligible infrastructures are specified and recommended to Ministry of Education by the ARC according to evaluation results (ARC, 2015a).

In this recommendation, ARC is using a rating scale which includes outstanding, excellent, very good, good and uncompetitive. Outstanding infrastructures are recommended unconditionally, excellent infrastructures are strongly recommended, and very good infrastructures have support recommendation of funding with reservation. These 3 ratings generally constitute %45 of all infrastructures. Good and uncompetitive infrastructures are not recommended for support and these infrastructures constitute %55 of all infrastructures (ARC, 2015c).

The infrastructure needs to prepare annual reports for the ARC covering both their financial situation and their research performance. Infrastructures are also obliged to report on a range of common and infrastructure specific Key Performance Indicators which are approved by the ARC. Each infrastructure undergoes a difficult and broad performance evaluation arranged by the ARC in its fourth year of operation. Continuance of funding for the remaining years depends on the outcome of the review (ARC, 2015a).

4.5. Czech Ministry of Education, Youth, and Sports (MSMT)

This section describes the evaluation model for research infrastructures designed by the Czech Ministry of Education, Youth and Sports (MSMT) with the project of individual national projects for the area of tertiary education, research and development.

The main objective of the evaluations performed by MSMT is stated as giving research infrastructures a tool that enabling them to explicitly define and assess themselves and deciding whether it matches to genuine and notable research infrastructure characters. The main aim of the evaluation methodology is specified as to combine and strategically structure the research infrastructure area and to expand the support for research infrastructures in the Czechia meeting the criteria of excellent quality, social need and utilization (MSMT, 2014a).

4.5.1. Overall Evaluation Process

The model described in this section was announced in 2014. The relevant authorities have studied the world examples in the model design process and adopted the informed peer-review evaluation method. The methodology used before was perceived as counting outputs, then scoring these outputs and funding according to these scores. The previous evaluation was also interpreted as a reduction of the complexity of impact to an excessively basic category of research outputs (MSMT, 2017a).

New evaluation system basically depends on the evaluation of experts and informing experts about the quantitative outputs of the infrastructure. For this reason, the evaluation system is also called as "informed-peer review" (Mahieu and Arnold, 2015). The main difference between the Czech model and the other countries is that the opinions of external users are obtained through questionnaire and participation in the evaluation.

The evaluation process consists of two interrelated evaluation stages. In the first stage, general concepts about infrastructure are explained and compliance with the evaluation indicators is assessed. Evaluation performed in this stage is carried out formally and contextually, including the strategic importance of the infrastructure (MSMT, 2014a).

The second stage of evaluation requires a more detailed presentation of the infrastructure and serves to make an open-ended assessment, such as infrastructure, quality and user potential. Foreign experts participate in the evaluation process from the relevant research areas (MSMT, 2014a). Administratively, MSMT is carrying out the evaluation process.

4.5.2. Evaluation Indicators

Evaluation indicators of the MSMT are outlined in Table 10. In Czechia case, evaluation indicators and sub-indicators are not matched one by one. Sub-indicators can be considered as horizontal and this provides a chance for applying them to all indicators. In this respect, Czechia evaluation indicators have quite unique characteristics. Information about indicators is collected by evaluation forms and evaluated by experts and panels with respect to these evaluation indicators.

Table 10 MSMT Evaluation Indicators

<table>
<tr><th>MSMT Evaluation Indicators</th><th>Sub-Indicators</th></tr>
<tr><td>A. Research and Innovation Potential, Quality</td><td rowspan="7">- Collaborations,
- Scientific Results (Publications, Patents, Contractual projects, other outputs)
- Impact on Czechia
- Use of infrastructure capacity: number and structure
- Contribution to the development of new technologies and innovations
- Technological equipment
- Infrastructure strategic management
- Education Activities</td></tr>
<tr><td>B. Utilization and Impact</td></tr>
<tr><td>C. Relevance for Research</td></tr>
<tr><td>D. Relevance for Industry</td></tr>
<tr><td>E. Relevance For Technology and Innovation</td></tr>
<tr><td>F. International Importance</td></tr>
<tr><td>G. Feasibility</td></tr>
</table>

Source: MSMT Evaluation Methodology for Research Infrastructures

The information requested from the infrastructures via the application forms in the first stage evaluation is given below (MSMT, 2014b):

- Introduction of research infrastructure
- Usage of infrastructure
- The level of implementation of the infrastructure and integration with national and international programs
- Relationship with the national research environment and research institutions
- SWOT analysis and managerial information
- Infrastructure expenses and activities

In the second stage evaluation, two separate forms are requested from the infrastructures. In the first form following detailed information about the infrastructure activities are requested (MSMT, 2014c):

- Detailed introduction to infrastructure
- Importance of infrastructure
- International relevance
- Utilization and outputs of infrastructure
- Benchmarking of infrastructure
- Research collaboration and collaboration with other stakeholders
- Management and feasibility

- Expenses and budget
- Portfolio of indicators

In the second form, the information of existing or potential users is requested from infrastructures to measure usage. After that, a questionnaire is sent to the users. In these surveys, the following information is obtained (MSMT, 2014c):

- Services provided by the infrastructure
- User evaluations of the services of the infrastructure
- Suggestions for improvement of existing services

Then the results of these surveys are analyzed and evaluations of the external use of the infrastructure are performed accordingly.

4.5.3. First Stage of Evaluation

In this stage, two or three foreign experts with an international reputation are appointed for each infrastructure evaluation by the scientific board constituted in each research area. Scientific boards composed of usually two international and one Czech expert in the relevant field. Both experts and scientific board have a chairman who leads and directs the evaluations. Appointed experts by the scientific board prepare reports on the evaluation of information gathered on first stage evaluation form specified above.

This review is mainly scientific review and elaborates the question whether the specific infrastructure really corresponds to research infrastructure norms in the relevant research area. At the end of the first stage, evaluation reports are submitted to the relevant scientific boards by the experts for further evaluation (MSMT, 2014a).

4.5.4. Second Stage of Evaluation

The second stage of evaluation requires a more detailed presentation of the infrastructure and lets infrastructure representatives to present their outputs to the scientific boards. Scientific board members conduct individual assessments for each infrastructure on the basis of first stage evaluation results and their own knowledge. Then, the relevant scientific board evaluates all infrastructures operated in their respective field, writes a summary report, and presents this report to the Evaluation Committee.

Evaluation Committee is consisting of chairmen of the experts and the chairmen of the scientific boards assigned by MSMT. The mission of the committee is to decide on the final recommendations to be made to the Ministry, to assign each infrastructure project to the relevant scientific board, to confirm the first stage evaluation from the experts, to accept the second stage evaluation from the scientific boards (MSMT, 2014a).

After all evaluations, for each evaluation indicators, the Evaluation Committee performs grading (for each infrastructure) with using a five-point scale from 1 to 5. For each evaluation indicator, the values of these grades should be described as well. The committee is not assigning a single total grade for the research infrastructures (Mahieu and Arnold, 2015).

The most up-to-date criticism of the Czech model is the use of too many international referees in evaluations. At this point, it is understood that international referees do not sufficiently know the dynamics of the research ecosystem in the country, which also leads to various problems as well (MSMT, 2017b).

5 DISCUSSION FOR MODEL PROPOSAL AND CONCLUSION

5.1. Brief Analysis of World Examples

Within the scope of this thesis, Australia, Czechia, France, Germany and Portugal evaluation systems are examined and it is seen that each country designs a system in the direction of its own dynamics and it has been found that every country has its own methodology. However, the existence of the most common aspects and approaches of the evaluation systems examined is also striking. Main methodologies from world examples are outlined in Table 11.

Table 11 Methodologies of World Examples

Country	Peer Review	Bibliometrics	Site Visit	Interview	Survey	Grading
Australia	✔			✔		Outstanding-Uncompetitive
Czechia	✔	✔			✔	1-5 Points
France	✔	✔	✔			No Grading
Germany	✔			✔		Outstanding-Sufficient
Portugal	✔	✔	✔	✔		Exceptional - Poor

When evaluation methodologies across the countries are further analyzed, it can be clearly stated that peer review is the main method and other methods are used for assisting peer review. It can also be seen that generally peer review and bibliometric methodologies are the methods used jointly. Beyond these methodologies, infrastructure representatives' opinions are gathered with site visits or interviews. Another common feature is about experts employed in the evaluation processes. International experts are employed in evaluations so that quality of the national infrastructures can be evaluated with international perspectives.

Quantitative evaluations are not used frequently. Instead of quantity, quality is the major concern in research infrastructure evaluations. In the world samples scrutinized, it is observed that pure bibliometric methods are avoided in the evaluation processes. Infrastructure performance and impact are not evaluated with some calculations made over numbers of research outputs such as projects, publications, and patents, etc. On the contrary, it is observed that the experts in the related field have been conducted evaluations and main focus in evaluations is quality rather than quantity. However, quantities and statistics of research outputs are also supplied to the experts.

Grading is another point that should be elaborated. As a result of the evaluation campaigns, infrastructures are not graded with a final numerical score such as 3 or 4. Instead, infrastructures are graded qualitatively such as "poor" to "exceptional" or "outstanding" to "uncompetitive" or "sufficient". It is known that in France, where the system has been developed for a long time, grading has been abandoned and experts prepare only reports on the performance and impact of the infrastructures.

Although evaluation methodologies show the difference, evaluation indicators are somewhat similar across the country examples examined. Scientific impact based indicators such as publications, scientific collaborations, and projects are main indicators which are used in evaluations frequently. Moreover, economic and technological impact indicators such as intellectual rights, spin-off firms, and collaborations with other parties are also used in almost all evaluations. In terms of social and educational indicators; national contribution, education activities supplied by the infrastructure and access measures can be emphasized.

In addition to these impacts, it can be observed that in almost all evaluation systems management quality is measured. In management quality evaluations; strategic plans, leadership capabilities, and human resource policy etc. are questioned. Qualified management structure is vital for the sustainability of research infrastructures. EIROforum is a charter based collaboration of powerful research infrastructures in Europe. In EIROforum discussion paper (2015), long-term sustainability of research infrastructures is evaluated with five main criteria which are scientific relevance, sustainable management and funding model, ability to attract scientific talent and socio-economic impact. Besides impact and scientific issues, it can be seen that management and human resources management are also crucial elements for long-term sustainability of research infrastructures.

5.2. Discussions for Suggested Indicator Set of Turkey

While suggesting indicators for the evaluations, Turkish dynamics with legal infrastructure and the global trend will be harmonized. In order to grasp Turkish dynamics, Law No. 6550 is examined. In Law No. 6550, duties and responsibilities of the research infrastructures are defined. Hence, these duties and responsibilities need to be satisfied by the research infrastructures by law and they should also be taken into consideration while determining evaluation indicators. Duties and responsibilities of research infrastructures are defined in Law No. 6550 Article 6 Clause 1 as follows:

a. Establishing necessary infrastructures related to the field of activity and operating these infrastructures
b. Performing education, basic and applied research, technology development, technology transfer and entrepreneurship activities
c. Providing continuous service to universities, public institutions, and private sector
ç. Preparing cooperation projects with universities, public institutions, and private sector
d. Carrying out research projects with their own income or with national and international resources
e. Supporting education activities conducted in universities
f. Filing and protecting patents and other intellectual and property rights
g. Providing consultancy services for intellectual and property rights and offering financial support for the acquisition and protection of rights
ğ. Establishing spin-off companies to ensure that the information produced in the research infrastructures and developed technologies are transformed into commercial values that contribute to the country's economy, industrial and social development
h. Educating users about infrastructure usage and security
ı. Acquiring necessary security measures related to the research infrastructure and employees in accordance with the quality assurance system and standards, accreditation, environmental and ethical legal arrangements
i. Cooperating with domestic and / or foreign bodies with contracts, protocols, and agreements

j. Organizing and participating in scientific meetings such as seminar, symposium, and congress and publishing in the field of activity

Also, Clause 2 of the Article 6 is specifying that research infrastructures need to prepare and send the vision, mission statements, strategic objectives, performance indicators and action plans to the Research Infrastructures Council (RIC) for approval.

By taking these duties and responsibilities as a base for suggested indicator set, country examples are elaborated accordingly. Moreover, impacts covered with specific indicators are also specified. Finally, examined country indicators are blended with these duties and responsibilities specified by the Law No. 6550 and suggested set of indicators are prepared and presented in Table 12 below.

Table 12 Suggested Evaluation Indicators for Research Infrastructures

Suggested Evaluation Indicator	Impacts Covered with Indicator	Countries Used This Indicator	Related Part Of Article 6
1. Publications and Citations	Scientific, Social	Australia, Czechia, France, Germany Portugal	Clause 1 Sub clause b, j
2. Scientific Projects	Scientific and Technological	Australia, Czechia, France, Germany Portugal	Clause 1 Sub clause ç, d
3. National Importance and Contribution	Social, Educational	Australia, Czechia, France, Germany Portugal	Clause 1
4. Intellectual Rights	Scientific, Technological and Economic	Australia, Czechia, France, Germany Portugal	Clause 1 Sub clause b, f, g
5. Spin-Off Firms	Scientific, Technological and Economic	Czechia, France, Portugal	Clause 1 Sub clause ğ
6. Scientific Events and Education	Scientific, Social, Educational	Australia, Czechia, France, Portugal	Clause 1 Sub clause e, h, j
7. Access and Usage	Scientific, Economic, Social, Educational	Czechia, France, Germany, Portugal	Clause 1 Sub clause c, h
8. Collaborations	Scientific, Economic and Social	Australia, Czechia, France, Germany Portugal	Clause 1 Sub clause ç, i
9. Management Quality and Research Strategy	-	Australia, Czechia, France, Germany Portugal	Clause 2

Scientific excellence is the primary concern in all evaluation systems and generally evaluated with publications, citations and scientific projects. In proposed indicator set, these are also included. Besides qualitative data of these indicators, qualitative sides such as the impact on the relevant scientific area should also be gathered. Moreover, non-academic publications' collection is also advised because it is a good indicator for measuring social impact of the research infrastructure.

National importance and contribution are also measured in all evaluation systems due to accountability concerns. This indicator can be used as a horizontal indicator that can be applied to all indicator types as well. Other than a horizontal feature of this indicator, research infrastructure's critical importance and impact on the visibility, attractiveness and international prestige of Turkish science can be questioned under this indicator.

Intellectual rights and spin-off firms have comprised scientific, technological and economic impacts. They are basically ways of commercialization of research. In Turkey, commercialization of research is one of the biggest problems in the research community (Özer, 2011). For this reason, these indicators are also valuable in terms of giving direction to research infrastructures as well. Under intellectual rights indicator data about patents, utility models, industrial designs and licenses could be collected. These outputs are all result of research activities and have technological and economic features as well. Spin-off firms are also established firms as a result of research activities. Spin-off firms can be evaluated with a contribution to employment, regional development, and high technology exports. Qualitative and quantitative information about these contributions can be gathered under this indicator.

Scientific events and education services or activities of research infrastructures are important in terms of dissemination of knowledge accumulated in these infrastructures. Scientific events such as conferences, symposiums, meetings, and courses could be evaluated under this indicator scope. Their number, diversity, and dissemination quality can be evaluated.

Law No. 6550 specifies that research infrastructures are open to the use of all users. For that reason, usage of and access to research infrastructures is needed to be evaluated as well. Under access indicator information about accessed user numbers, revenues gathered by accesses, access policies and strategies can be collected. Moreover, users of the research infrastructures can be collected from infrastructures and survey can be conducted to these users in order to grasp user's opinions and feedbacks.

Although main problems of the research infrastructures are solved with the Law No 6550, there are some critical issues to be solved with the help and guidance of evaluation model. First one is the lack of partnership and cooperation between research infrastructures and stakeholders, especially the private sector. In order to solve this problem, collaboration indicator is also added. Collaborations with the private sector, academic world, and other stakeholders are questioned and measured in all evaluation systems.

Research infrastructures' collaborations with the private sector and other stakeholders will be gathered by this indicator. Collaborations can be categorized as international, the private sector, and research collaborations and can be evaluated under these sections. Advisory and consultancy services provided by researchers or infrastructure, which is a good indicator of social impact, can also be collected under this indicator. Role and responsibility of research infrastructures in collaborations, quality, and importance of collaborations might be gathered as a qualitative data. Number of and fund volume of collaborations might be collected as a quantitative data.

The second problem specified is lack of institutional, effective and sustainable management. For that reason in addition to these impact indicators management quality of the infrastructures is also added to the suggested evaluation criteria. The second reason for adding this specific criterion is its effect on the sustainability of the research infrastructures as argued before. Moreover, according to Mahieu and Arnold (2015), in an escalating number of countries, evaluations at the national level contain evidence on the performance of research infrastructures and their management practices.

Under management indicator; strategic management, human resources management, access management and project, process, and quality management issues might be questioned. Also, the sufficiency of the research strategy and feasibility is also crucial for the sustainability of the research infrastructure. In country examples, this indicator is also questioned frequently and needs to be assessed in Turkey as well.

5.3. Discussions for Developing Methodology for Turkey

In order to sustain the acceptability and prestige of the process which will be carried out in Turkey, it is suggested as a priority to design evaluation model specific to Turkey by following the common standards and updates in country evaluation methodologies and taking into consideration Turkey's dynamics.

There is a universal tendency towards national research evaluation systems that rely on a combination of peer review and quantitative indicators (Colwell at al 2012). Also, Moed (2007) claimed that the future of research assessment is based on a smart mixture of bibliometrics and peer review. Moed further states that real difficulty is to unite the two methodologies in a way that the strength of the first pays off for the weaknesses of the second. In this respect, Dutch Committee on Quality Indicators in the Humanities (2011) specified that the most severe objections to peer review can be neutralized by means of practical assurances and by exploiting bibliometrics that provides an intersubjective basis to the evaluation of experts.

As in our world examples, a number of national research evaluation practices have independently evolved towards this methodology, frequently in response to extensive analyses of current evaluation methodologies supported by the government agencies handling these initiatives. The international practice has been combining the peer review with bibliometrics. This methodology exploits the ability of bibliometrics to represent data in a simplified overview while exploiting the ability of experts to make qualified judgments about impact, excellence, and other qualitative features that can not be reached through only with bibliometrics.

As a universal tendency, it can be concluded that there is an agreement reflected in both methodological recommendations made by research evaluation experts and evaluation practices accepted by government bodies that the robust method to research evaluations in the field level relies on a mixture of indicators and peer review.

In terms of Turkish ecosystem and dynamics of research infrastructures, supports on research infrastructures have just started in the 2000s in Turkey. For that reason, Turkish ecosystem of research infrastructure is relatively young. This situation leads time lag limitation for evaluations. Another handicap of research infrastructures being new is that they are not well-developed and sophisticated compared to analyzed countries infrastructures. Turkey has only several qualified research infrastructures which are eligible enough for competing with developed countries' research infrastructures in scientific, economic and social outputs. Moreover, as specified before these infrastructures have no clear institutional and managerial structure. This leads to problems in data management and data storage as well. The great majority of the infrastructures in Turkey have difficulty in providing data for evaluations. For these reasons, pure bibliometric evaluation is not recommended.

On the other hand, research infrastructures are among the most critical elements in the R&D ecosystem in Turkey. Advanced R&D activities carried out in universities are mainly performed in the research infrastructures and R&D ecosystem stakeholders are also aware of this situation. Main R&D activity, which is also having a connection with articles, patents, spin-off firms, and collaborations, is R&D projects.

In one of the leading research infrastructure located in one of the leading universities of Turkey, it has been seen that the number of TUBITAK projects carried out in infrastructure corresponds to more than half of the all TUBITAK projects in the university. Considering that TUBITAK is the main actor in Turkey in R&D funding area, it is clearly seen that the related infrastructure is undertaking a large part of R&D intensity of the university by itself. However, peer review is needed in order to evaluate these projects in terms of quality and contribution to the specific research area.

After elaborating national dynamics of the Turkish research infrastructure ecosystem, because of data and quality concerns, use of peer review and bibliometrics seems to be also the best methodology for Turkey. Lack of proper bibliometric data and quality needs on current data are important limitations for evaluations. These problems can be solved with help of qualitative review of experts supported with accessible bibliometric data.

After scrutinizing universal tendency and national dynamics, informed peer review is suggested as the main methodology for Turkey. Suggested evaluation methodology starts with an individual peer review performed by experts. All country examples are employed peer evaluations in their first stage of evaluations. As country examples are analyzed further, it can be seen that either site visits or interviews are carried out for interacting with infrastructure after peer reviews.

In this respect, in suggested evaluation methodology site visit is preferred because it allows more interaction with infrastructure and provides an opportunity to gather more and detailed information. Site visit is a regular tool used in institutional evaluations where detailed evaluations need closer contact with the evaluated bodies. This is also needed in Turkey case because lack of valuable research data necessitates the detailed and closer contact for grasping necessary information needed for the evaluations.

Access and usage measure is also critical and there is not proper and uniform data in research infrastructures. This limitation is obstructed the evaluation of usage and access. For that reason, as in Czechia example, where the access quality is measured by user opinions gathered with surveys, in suggested methodology user opinions will be collected by surveys.

In the last part of the evaluations examined, it is observed that performed evaluations in peer review and site-visit or interview stages are reviewed by thematic panels or committees before the final decision. The first reason for forming thematic panels is that it allows comparing the infrastructures in a related field. Also in terms of reliability of the evaluation process, it is needed to evaluate infrastructures on the same panel. Otherwise different experts might evaluate infrastructures differently and this might raise some concerns among evaluated infrastructures.

There are considerable differences among research areas. These differences can be stated in terms of output variety, publication patterns, and timelines, citation characteristics, publication language, collaboration features and needs, intensity of the use and need of financial, equipment and human resources and way researchers produce and transfer knowledge (Mahieu and Arnold, 2015). Thematic panels have ability to take field-specifics in concern during the evaluations. Nevertheless, there should be uniformity in evaluations among thematic areas and only partial area modifications should be permitted.

In terms of Turkish dynamics side, TUBITAK has long experience on panel evaluations. Currently, support programs are employing panel evaluations for decision making. Also, panel evaluations used by TUBITAK is favored by the researchers in Turkey. A survey conducted by TUBITAK in 2006 indicates that irrespective of publication records, geographical localities, and funding decisions, researchers in Turkish universities are persuaded of the efficiency of the panel evaluations conducted by TUBITAK (European Science Foundation, 2007).

For these reasons, final thematic panel evaluation is also recommended in suggested evaluation methodology. After forming the general framework, details of the suggested evaluation methodology is described in the next section.

5.4. Suggested Overall Evaluation Process

After considering all the world examples and Turkish research infrastructure dynamics, mainly informed peer review methodology is suggested. In detail, suggested evaluation system will contain individual peer review, site visit, and panel evaluation. In all these evaluation methodologies experts will be informed about the bibliometric data of the related infrastructure. The suggested overall evaluation process is outlined in Figure 2.

Figure 2 Suggested Evaluation Model for Turkey

5.4.1. First Stage of the Proposed Evaluation

As observed in all world examples, the first stage of evaluation contains evaluations on online forms by the experts. Similarly, in suggested methodology, online application forms which will be designed according to proposed evaluation criteria will be sent to the research infrastructures by the TUBITAK initially. Research infrastructures will fill the forms and send to the TUBITAK. Research infrastructures' bibliometric data and qualitative information about proposed evaluation indicators will be collected by these online forms. Once the online forms arrived in TUBITAK, selected experts will evaluate the information supplied by the infrastructures individually.

One industry oriented, one academia oriented expert from Turkey and one international expert from the relevant research area will be assigned for each evaluation. When world examples are scrutinized, generally international experts are appointed for peer review evaluations. Also, survey conducted by TUBITAK in 2006 among Turkish researchers further indicates that increasing the involvement of international reviewers was considered to be the key means of refining the current evaluation system used by TUBITAK (European Science Foundation, 2007). For these reasons, international participation in evaluations is strictly recommended.

A reason for suggesting experts amongst Turkish academia is they have knowledge about dynamics of the research ecosystem in Turkey. Industry oriented expert can contribute to evaluations of industry collaborations and management quality. One international expert will provide international vision to the evaluations. Reason for suggesting only one international expert is low maturity level of Turkish research infrastructure ecosystem compared to developed countries. As in Czechia example, some problems may arise because experts might not sufficiently capture the dynamics of the research ecosystem in Turkey.

Each expert will prepare individual evaluation reports for each infrastructure as a result of individual evaluations. Reports will include brief evaluations with critics, comments, and questions. These reports will form a basis for the second stage of evaluations. By these evaluations, experts will also prepare themselves for the second stage of evaluation by learning about infrastructure and preparing questions for the site visits. Moreover, these individual reports will also be shared with the infrastructures. This will increase the transparency of the evaluations and provide a chance to infrastructures for preparation of critics and comments in the first evaluation stage.

5.4.2. Second Stage of the Proposed Evaluation

The second stage of the evaluation is started with the site visits. Experts appointed for the first stage of the evaluation will also attain to the site visits. Moreover, in order to sustain a certain standard, observers from TUBITAK and Ministry of Development will also attend to the site visits. As in world examples of France and Portugal, site visits will start with brief introduction presentation of research infrastructure by the infrastructure managers. After that, infrastructure will be physically visited and research groups will introduce themselves and the infrastructure’s machine and equipment. Lastly, communication of experts with researchers and managers will take place. In third phase, questions and comments written in the first stage evaluations about the infrastructure will be debated with infrastructure researchers and managers.

After the site visits, 3 experts will prepare a site visit report jointly. This report will contain a short evaluation of the research infrastructure, strengths, and weaknesses of the infrastructure and brief advice for research infrastructure and for following thematic panel evaluation.

In site visit, grading is not suggested because of consistency and objectivity reasons. In order to provide consistent and objective evaluation, infrastructures in the same research area should be evaluated together. Different experts might evaluate with different perspectives, which can harm objectivity of the evaluations performed.

In the majority of the world samples examined, infrastructures are graded with "poor" to "exceptional" or 1 to 5 scales. This scaling is showing the level of achievement of research infrastructure. Infrastructure evaluation is not like a single project evaluation so the panel can only place the infrastructure in a certain range of success if there is a holistic evaluation (projects, patents, spin-off firms, etc.). For this reason, it would be more appropriate to give qualitative notes, such as "poor" and "exceptional", which will help to rank the infrastructures when necessary but will not give an idea of the magnitude of the difference. All in all, it is recommended to grade infrastructures with a 5 scale qualitative rating as “Poor”, “Fair”, “Good”, “Excellent” and “Exceptional”.

The suggested evaluation model should perform research area-focused assessments and should take infrastructure mission into account as well. Model should not expect an infrastructure to be successful at all evaluation indicators because they have different research areas and missions. For example, research infrastructure can collaborate with national defense companies, conduct applied research and provide significant projects for defense industry. Due to confidentiality issues, this specific infrastructure might not create desired number of publications, patents etc. Also in Germany, Czechia, and France's previous evaluation systems, grading is performed for evaluation indicators and whole infrastructure is not graded. For all these reasons, it is suggested that grading should be performed in evaluation indicators level.

In the world examples, generally, grading and final evaluation report writing are performed by thematic panels. These thematic panels are assessing all the evaluation reports of infrastructures in the related thematic area. These thematic panels include experts participating in the site visits and other specialists from the relevant thematic field. For that reason, at the end of the second evaluation stage, thematic panel constitution is also advised. Thematic panel will be constituted by site visit experts and observers, international experts and experts from private sector in Turkey as well. Moreover as in Portugal case, for interdisciplinary research infrastructures, more than one panel evaluation can be performed. In panel evaluation, site visit reports of all infrastructures will be argued and final recommendation with grades will be introduced to the Monitoring and Proficiency Evaluation Committee. This committee is preparing a recommendation on proficiency decision of research infrastructures for RIC by taking into consideration these evaluation results. Lastly, RIC is deciding on which research infrastructures will acquire proficiency status.

Almost in all evaluation systems, evaluation documents such as reports, recommendations and grades are announced to the public for transparency concerns. Also in our suggested methodology, all evaluation reports will become public when the evaluation process is finalized.

5.5. Conclusion

Evaluation institutes a key component of the research infrastructure policy cycle. In order to be effective and achieve the desired results, an evaluation system should be well planned and designed. For that reason, before suggesting an evaluation model for Turkey, literature is elaborated and prominent country examples are analyzed.

Following the introductory chapter consisting of research background, problem discussion, purpose and research questions, the second chapter begins with the presentation of research infrastructure concept. In this chapter, current legislation about research infrastructures and structure of research infrastructure system are also explained. The next chapter starts with research evaluation history and continues on challenges and purposes of the research evaluation. Evaluation focus and main evaluation tools are outlined and finally, impacts and impact indicators of research infrastructures are analyzed.

The fourth chapter in the study examines world examples of evaluation systems of research infrastructures in terms of methodologies used and indicators employed. In the last chapter, short analysis of examined evaluation systems is made and discussions are held on suggested indicator set and methodology for Turkey. In the end, details of the suggested evaluation model are explained.

Main impacts of research infrastructures in current literature are outlined and analyzed in scientific, economic, social, educational and technological impacts. Indicators employed under these impacts in the literature are also identified. In world examples, it is observed that every country has its own evaluation methodology. Nevertheless, the presence of the most shared practices in examined country systems is noteworthy. In general, peer review is the core tool and other tools are employed for supporting peer review. In addition to this, peer review and bibliometric evaluation tools are used together. Infrastructure's views are generally collected through site visits or interviews. In evaluations' final stage, completed evaluations up to that point are examined by thematic panels. Evaluation indicators seem to be parallel across the country examples.

Elaborated country indicators are combined with legal requirements specified by the Law No. 6550 and final indicator set is recommended. Also in methodology side, after analyzing all country examples and Turkish research infrastructure ecosystem, informed peer review methodology is suggested. In detail, consecutively; individual peer review, site visit, and panel evaluation tools are recommended for the evaluation system.

Main contribution of this thesis to the management and organization area is made by introducing and suggesting management evaluation practices with indicators to the suggested evaluation system. In evaluation indicators, under management quality indicator; vision, mission statements, strategic objectives, performance indicators, action plans of research infrastructures will be evaluated according to suggested evaluation system. In addition to this, human resources management, access management, project, process, and quality management issues are also considered in suggested indicator set.

This study provides several contributions to the scientific literature and research infrastructure evaluation campaign of Turkey. In regard to the scientific literature, the thesis is contributed to need of developing methodologies and indicators for evaluating research infrastructures. Moreover, literature and country examples are classified in terms of impacts and impact indicators of research infrastructures. In addition to scientific contributions, by proposed evaluation methodology and indicator set, the basis for a new research evaluation is specified and concrete model for evaluation implementation is suggested.

REFERENCES

Abramo, G., & D'Angelo, C. A. (2011). Evaluating research: From informed peer review to bibliometrics. Scientometrics, 87(3), 499-514.

Abramo, G., D’Angelo, C. A., Di Costa, F. (2011). National research assessment exercises: a comparison of peer review and bibliometrics rankings. Scientometrics, 89(3), 929.

Alagna, L. (2015). Socio-economic impact of research infrastructures. Organisation for Economic Co-operation and Development (OECD) Global Science Forum. Retrieved on November, 25, 2017, from: https://www.innovationpolicyplatform.org/system/files/ALAGNA.pdf

Allspach, K.V., Kancewicz H.J., Walter, N., White, O., Worms, J.C. (2015). European science foundation report on the foundation for science and technology evaluation of R&D Units 2013-2014. Strasbourg. Retrieved on November, 12, 2017 from:https://www.fct.pt/apoios/unidades/avaliacoes/2013/docs/ESF_FCT_report_Final.pdf

ARC (2015a). Funding rules for schemes under the Linkage Program for the year 2017-ARC Centres of Excellence. Canberra: Australia. Retrieved on January, 11, 2018 from:http://www.arc.gov.au/sites/default/files/filedepot/Public/NCGP/CE17/CE17_Funding_Rules_new.pdf

ARC (2015b). Presentation for ARC Centres of Excellence for funding commencing in 2017. May 2015, Canberra: Australia. Retrieved on January, 11, 2018 from: http://www.arc.gov.au/sites/default/files/filedepot/Public/Media%20&%20Resources%20Centre/Presentations/2015_Presentations/p201506_CE17_outreach_roadshow.pptx

ARC (2015c). How ARC proposals are assessed (and how to be a good assessor). May/June 2015. Canberra: Australia. Retrieved on January, 11, 2018 from: http://www.arc.gov.au/sites/default/files/filedepot/Public/Media%20&%20Resources%20Centre/Presentations/2015_Presentations/p20150506_How_to_assess_ARC_proposals.pdf

ARC (2017). Consultation on the ARC Centres of Excellence scheme for funding commencing in 2017. Canberra: Australia.

Baker, R. S. J. D. (2010). Data mining for education. International Encyclopedia of Education, 7(3), 112-118.

Barros, F. (2013). Centres of Excellence as a Tool for Capacity Building: Case Study Australia. Programme on Innovation, Higher Education and Research for Development Paris: OECD. Retrieved on January, 11, 2018 from: http://www.oecd.org/sti/Center%20of%20Excellence%20-%20Australia.pdf

Bilim ve Teknoloji Yüksek Kurulu, (2010) Araştırma altyapıları yol haritası ve araştırma merkezlerine ilişkin idari ve yasal düzenlemeler, Decree No: 2010/203, Retrieved on November, 27, 2017, from: https://www.tubitak.gov.tr/tubitak_content_files//BTYPD/btyk/22/BTYK22_Yeni_Kararlar.pdf

Boyack, K. W., & Jordan, P. (2011). Metrics associated with NIH funding: A high-level view. Journal of the American Medical Informatics Association, 18(4), 423-431.

Bozeman, B. (1993). Peer review and evaluation of R&D impacts. In evaluating R&D impacts: Methods and practice. Springer, Boston, MA.

Braben, D. W. (2004). Pioneering Research: A risk worth taking. New York: John Wiley and Sons Publication.

Brutscher, P. B., Wooding, S., Grant, J. (2008). Health research evaluation frameworks: An international comparison. Canadian Academy of Health Sciences, RAND Corporation Report.

Bryman, A. (2015). Social research methods. Oxford: Oxford University Press.

Buxton, M., & Hanney, S. (1996). How can payback from health services research be assessed? Journal of Health Services Research, 1(1), 35-43.

Canada Treasury Board (2007). Canadian Cost-Benefit Analysis Guide. Planning Branch, Treasury Board Secretariat. Retrieved on December, 3, 2016, from: https://www.tbs-sct.gc.ca/rtrap-parfa/analys/analys-eng.pdf

Cappy, A. (2015). The evaluation of research units at HCERES. Retrieved on 30.01.2018 from: https://dspacecris.eurocris.org/bitstream/11366/382/1/N8-HCERES_AMUE-final.ppt

CERN (2016). "Annual report 2016, Retrieved on December, 07, 2017, from: https://cds.cern.ch/record/2270805/files/486-1611-1-SM.pdf

Check J., Schutt R. K. (2012). Survey research. Research methods in education. Thousand Oaks, CA: Sage Publications

Chinese Academy of Science (2007). Research and advice on management policy of National Large Science and Technology Infrastructure. Policy Report.

Chubin, D. E. (1994). Grants peer review in theory and practice. Evaluation Review, 18(1), 20-30.

Clarke, S., Mawhinney, M., Swerdlow, R., Teichmann, D. (2013). Project preparation and CBA of RDI infrastructure projects. JASPERS Knowledge Economy and Energy Division. Retrieved on December, 12, 2017, from: http://www.jaspersnetwork.org/plugins/servlet/documentRepository/downloadDocument?documentId=184

Colwell, R., Blouw, M., Butler, L., Cozzens, S. E., Feller, I., Gingras, Y., et al. (2012). Informing research choices: Indicators and judgment. The Expert Panel on Science Performance and Research Funding, 142.

Committee on Quality Indicators in the Humanities (2011). Quality indicators for research in the humanities, Interim report, KNAW - Royal Netherlands Academy of Arts and Sciences, Netherlands. Retrieved on December, 13, 2017, https://www.knaw.nl/shared/resources/actueel/publicaties/pdf/quality-indicators-for-research-in-the-humanities

Deem, R. (2015). Recent research evaluations in the UK and Portugal: Methodologies, processes, controversies and consequences. Global Challenges, National Initiatives, and Institutional Responses – The Transformation of Higher Education Rotterdam, Sense Publishers

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., Torrance, G. W. (2015). Methods for the economic evaluation of health care programmes, Third Edition, Oxford University Press.

EIROforum (2015). EIROforum discussion paper: Long-term sustainability of research infrastructures. Retrieved on March, 13, 2017 from: https://www.eiroforum.org/wp-content/uploads/20150325_discussion-paper-research-infrastructures-sustainability.pdf

Eisner, E. W. (1998). The enlightened eye: Qualitative inquiry and the enhancement of educational practice. Columbus, OH.

Elias, P. (2009). Developing methodologies for the impact analysis of research infrastructures on innovation. European Commission Conference "Role of Research Infrastructures for a Competitive Knowledge Economy". Retrieved on October, 2, 2017, https://ec.europa.eu/research/infrastructures/pdf/elias_06_09.pdf

European Commission (2010). A vision for strengthening world-class research infrastructures in the ERA, Report of the expert group on research infrastructures. Retrieved on November, 15, 2017, https://ec.europa.eu/eurostat/cros/system/files/14_A%20vision%20for%20strengthening.pdf

European Science Foundation (2012). Working group report: The challenges of impact assessment. Working group 2: Impact assessment, Retrieved on June, 12, 2017, from: http://archives.esf.org/index.php?eID=tx_nawsecuredl&u=0&g=0&t=1519734481&hash=0fb560c41ecb18d38cbb6dded4d03bbac34ea350&file=/fileadmin/be_user/CEO_Unit/MO_FORA/MOFORUM_Eval_PFR__II_/Publications/WG2_new.pdf

European Science Foundation (2013). MERIL: Mapping of the European research infrastructure landscape. Retrieved on August, 23, 2017, from: http://www.esf.org/serving-science/ec-contracts-coordination/meril-mapping-of-the-european-research-infrastructure-landscape.html

European Science Foundation (2007). Peer review – Its present and future state. Conference report. Strasbourg. Retrieved on April, 12, 2018 from: http://www.allea.org/wp-content/uploads/2016/02/EU_ESF_Peer_review.pdf

Fayyad, U. (2004). Data mining grand challenges. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 1-2). Springer, Berlin, Heidelberg.

Fayyad, U., Wierse, A., Grinstein, G.G. (2002). Information visualization in data mining and knowledge discovery, San Francisco: Morgan Kaufmann.

FCT (2013). Evaluation Guide: FCT evaluation of R&D units 2013. Fundação para a Ciência e a Tecnologia: Lisbon. Retrieved on May, 9, 2017 from: https://www.fct.pt/apoios/unidades/avaliacoes/2013/docs/GuiaoAvaliacao_AvaliacaoUnidades2013.pdf

FCT (2017). Fundação para a Ciência e a Tecnologia Apresentação. Lisbon. Retrieved on May, 9, 2017 from: http://iceubi.ubi.pt/wp-content/uploads/2017/03/Apresentação_FCT_UBI_4.12.2017_v2.pdf

Flinders, D. J. (1997). Interviews: An introduction to qualitative research interviewing: Steinar Kvale. Thousand Oaks, CA: Sage Publications.

Frank, C., Battista, R., Butler, L. (2009). Making an impact: A preferred framework and indicators to measure returns on investment in health research. Ottawa, ON: Canadian Academy of Health.

Funtowicz, S., Ravetz, J. (2003). Post-normal science. International society for ecological economics. Online Encyclopedia of Ecological Economics. Retrieved on September, 30, 2017, from: https://pdfs.semanticscholar.org/ce91/a2cf9b7e05411fb5b1b9276b9aaf565bffb2.pdf

Georghiou, L. (2002). Impact and additionality of innovation policy. Innovation Science and Technology IWT Observatory, 40(1), 57-65.

German Federal Ministry of Education and Research (2011). Recommendations on scientific collections as research infrastructures. Retrieved on August, 27, 2017, from: http://www.wissenschaftsrat.de/download/archiv/10464-11-11_engl.pdf

Geuna, A., Martin, B. R. (2003). University research evaluation and funding: An international comparison. Minerva, 41(4), 277-304.

Gibbons, M., Georghiou, L. (1987). Evaluation of Research: A Selection of current practices. OECD Publications, Paris.

Godin, B., Doré, C. (2004). Measuring the impacts of science: Beyond the economic dimension. CSIIC Working Paper. Retrieved on August, 15, 2017, from: http://www.csiic.ca/PDF/Godin_Dore_Impacts.pdf

Griniece, E., Reid, A., Angelis, J. (2015). Evaluating and monitoring the socio-economic impact of investment in research infrastructures. Technopolis Group. Retrieved on December, 22, 2017, from: http://www.technopolis-group.com/wp-content/uploads/2015/04/2015_Technopolis_Group_guide_to_impact_assessment_of_research_infrastructures.pdf

Guthrie, S., Wamae, W., Diepeveen. S., Wooding, S., Grant, J. (2013). Measuring research a guide to research evaluation frameworks and tools, RAND Corporation Report. Retrieved on June, 12, 2017, from: https://www.rand.org/content/dam/rand/pubs/monographs/MG1200/MG1217/RAND_MG1217.pdf

Haddix, A. C., Teutsch, S. M., Corso, P. S. (2003). Prevention effectiveness: A guide to decision analysis and economic evaluation. Oxford University Press.

Hansson, F. (2010). Dialogue in or with the peer review? Evaluating research organizations in order to promote organizational learning. Science and Public Policy, 37(4), 239-251.

Hanney, S., Gonzalez-Block, M. A. (2011). Yes, research can inform health policy; but can we bridge the 'Do-Knowing It's Been Done'gap? Health Research Policy and Systems, 1(9), 23.

Haustein, S., Larivière, V. (2015). The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects. In Incentives and performance, Springer Publications.

HCERES (2010). Selection of the experts. Retrieved February 18, 2016, from: http://www.aeres-evaluation.com/Agency/Organisation/Experts-of-AERES2/Process-for-appointing-experts

HCERES (2014). Criteria for the evaluation of research institutions: The HCERES Standards. Retrieved January 16, 2017, from: http://www.hceres.fr/content/download/25453/395443/file/R%C3%A9f%C3%A9rentiel%20ANGLAIS%20HCERES-Entit%C3%A9s%20de%20Recherche-septembre%202015.pdf

HCERES (2015). Guidelines for research units. Retrieved January 16, 2017, from: http://www.hceres.fr/content/download/23524/364804/file/Carnet_de_visite_UR_VAG_B_2015_FR-24%20juil-2015.pdf

HCERES (2016). HCERES report on research unit: Cerebral Imaging and Neurological Handicaps, ICHN, U Toulouse 3, Inserm. Retrieved January 6, 2017, from: http://tonic.inserm.fr/medias/fichier/ichn_a2016ev0311384ls2pur160009679009469rt_1431964496345.pdf?ID_FICHE=146332&INLINE=FALSE

Health Communication Unit (1999). Conducting survey research: Version 2.0, Toronto: Centre for Health Promotion, University of Toronto.

Hicks, D. (2012). Performance-based university research funding systems. Research Policy, 41(2), 251-261.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., Rafols, I. (2015). The Leiden Manifesto for research metrics. Nature, 520(7548), 429.

Holm-Hansen, C. (2007). Conducting Interviews: Tips for conducting program evaluation. Wilder Research Publication

Hungary National Research Development and Innovation Office (2014). Research infrastructures in Hungary. Retrieved on February, 2, 2017 from: http://www.nih.gov.hu/strategy/research-infrastructures/research-infrastructures-in-hungary

Irvine, J., Ben, R. M. (1984). What direction for basic scientific research? In Science and Technology Policy in the 1990s and Beyond, Harlow: Longman.

Jonkers, K. & Zacharewicz, T. (2016). Research performance-based funding systems: A comparative assessment; JRC Science for Policy Report. Retrieved on June, 14, 2017, from:http://publications.jrc.ec.europa.eu/repository/bitstream/JRC101043/kj1a27837enn.pdf

Kalkınma Bakanlığı (2017). Üniversite ve kamu kurumları araştırma altyapıları, Retrieved on December, 17, 2017, from: http://www.kalkinma.gov.tr/Lists/Yaynlar/Attachments/758/Üniversite%20ve%20Kamu%20Kurumları%20Araştırma%20Altyapıları.pdf

King, J. (1987). A review of bibliometric and other science indicators and their role in research evaluation. Journal of Information Science, 13(5), 261-276.

Kronick, D. A. (1990). Peer review in 18th-century scientific journalism. Jama, 263(10), 1321-1322.

Laranja, M. (2009). The development of technology infrastructure in Portugal and the need to pull innovation using proactive intermediation policies. Technovation, 29(1), 23-34.

Lawrenz, F., Keiser, N., Lavoie, B. (2003). Evaluative site visits: A methodological review. The American Journal of Evaluation, 24(3), 341-352.

Lawrenz, F., Thao, M., Johnson, K. (2012). Expert panel reviews of research centers: The site visit process. Evaluation and Program Planning, 35(3), 390-397.

Lu, Y. C. (1981). Ex Ante evaluation of the separate effects of research and extension. In Evaluation of Agricultural Research, Proceedings of a Workshop, Minneapolis, MN, May 12-13, 1980, Miscellaneous Publication 8 (No. 49074). University of Minnesota, Agricultural Experiment Station.

Macilwain, C. (2009). The RAE: an assessment too far? Cell, 139(4), 643-646.

Mahieu, B., Arnold, E. (2015). R&D evaluation methodology and funding principles: Final report 1: The R&D Evaluation Methodology. Technopolis Group. Retrieved on January, 6, 2018 from: http://metodika.reformy-msmt.cz/cs/uvodni-strana-cz?file=vloz_soubor&id=524&task=download

Martin, B. R., Irvine, J. (1983). Assessing basic research: some partial indicators of scientific progress in radio astronomy. Research Policy, 12(2), 61-90.

Martin, B. R., (1997). Evaluating science and scientists: Factors affecting the acceptance of evaluation results. Evaluating Science and Scientists Budapest: Central European University Press.

Michalowski, S. (2014). The impacts of large research infrastructures on economic innovation and on society: case studies at CERN. OECD Global Science Forum. Paris. Retrieved on February, 11, 2018 from: France, https://www.oecd.org/sti/sci-tech/CERN-case-studies.pdf

Moed, H. F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. Science and Public Policy, 34(8), 575-583.

Moed, H. F., W. J. M., Burger, J. G., Frankfort, A. F. J. Van Raan. (1985). The use of bibliometric data for the measurement of university research. Research Policy, 14(3), 131–149.

Morris, Z. S., Wooding, S., Grant, J. (2011). The answer is 17 years, what is the question: understanding time lags in translational research. Journal of the Royal Society of Medicine, 104(12), 510-520.

MSMT (2014a). Evaluation methodology for research infrastructures. Prague: Czechia Retrieved on September, 01, 2017 from: https://www.msmt.cz/file/33834_1_1/

MSMT (2014b). Large research infrastructures “Evaluation form A” Prague, Czechia Retrieved on September, 01, 2017 from: http://www.msmt.cz/file/39360_1_1/

MSMT (2014c). Large research infrastructures “Evaluation form B” Prague, Czechia Retrieved on September, 01, 2017 from: http://www.msmt.cz/file/39402_1_1/

MSMT (2017a). Feedback on the draft summary report, Prague, Czechia Retrieved on September, 01, 2017 from:http://metodika.reformy-msmt.cz/cs/cck?file=vloz_soubor&id=385&task=download

MSMT (2017b). Feedback on the draft summary report, Prague, Czechia Retrieved on September, 01, 2017 from:http://metodika.reformy-msmt.cz/cs/cck?file=vloz_soubor&id=384&task=download

National Research Council (1998). Investing in research infrastructure in the behavioral and social sciences. Washington, DC: National Academies Press.

National Research Infrastructure Council (2011). Strategic framework for research infrastructure investment. Retrieved on October, 17, 2016, from: http://www.innovation.gov.au/Science/ResearchInfrastructure/Documents/StrategicFrameworkforResearchInfrastructureInvestment.pdf

Nelkin, D. (1979). The social responsibility of scientists. Annals of the New York Academy of Sciences, 334(1), 176-182.

OECD (2008). Organisation for Economic Co-operation and Development, Glossary of statistical terms. France, Paris.

OECD (2011a). Issue brief: Peer review, Retrieved on September, 22, 2017, from: http://www.oecd.org/innovation/policyplatform/48136766.pdf

OECD (2011b). Issue brief: research organization evaluation. Retrieved on December, 1, 2016, from: http://www.oecd.org/innovation/policyplatform/48136330.pdf

OECD (2016). OECD Global Science Forum workshop on methodologies and tools for assessing socio-economic impact of research infrastructures, summary report, Paris, Retrieved on October, 13, 2017, https://www.innovationpolicyplatform.org/system/files/Agenda07-OECD%20GLOBAL%20SCIENCE%20FORUM%20WORKSHOP%20ON%20METHODOLOGIES%20AND%20TOOLS%20FOR%20ASSESSING%20SOCIO-ECONOMIC%20IMPACT%20OF%20RESEARCH%20INFRASTRUCTURES%20_0.pdf

Okubo, Y. (1997). “Bibliometric indicators and analysis of research systems: Methods and examples”, OECD Science, Technology and Industry Working Papers, 1997/01, OECD Publishing.

Oppenheim, C. (1997). The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology. Journal of documentation, 53(5), 477-487.

Özer, Y. E. (2007). Girişimci üniversite modeli ve Türkiye. Uludağ Journal of Economy and Society. 2(85).

Pancotti, C., Pellegrin, J., Vignetti, S. (2014). Appraisal of research infrastructures: Approaches, methods and practical implications. Milan European Economy Workshops, Working Paper No. 2014-13. Retrieved on January, 21, 2018 from: http://wp.demm.unimi.it/tl_files/wp/2014/DEMM-2014_13wp.pdf

Patton, M.Q., (2015). Qualitative research & evaluation methods: Integrating theory and practice, 4. Edition, Thousand Oaks. CA Sage Publications.

Pereira, A. G., Funtowicz, S. (2005). Quality Assurance by Extended Peer Review: Tools to Inform Debates, Dialogues & Deliberations. Technikfolgenabschätzung Theorie und Praxis, 14(2), 74-79.

Ponto J. (2015). Understanding and evaluating survey research. Journal of the Advanced Practitioner in Oncology. 6(2), 168–171.

Pritchard, A. (1969). Statistical bibliography or bibliometrics. Journal of documentation, 25(4), 348-349.

Qiao, L., Mu, R., Chen, K. (2016). Scientific effects of large research infrastructures in China. Technological Forecasting & Social Change. 112(1), 102-112.

Research Councils of United Kingdom (2006). Efficiency and effectiveness of peer review. Polaris House Swindon. Retrieved on November, 3, 2017, from: http://www.rcuk.ac.uk/documents/documents/rcukprreport-pdf/

Rizzuto, C. (2012). Benefits of research infrastructures beyond science. Presentation at ERF Workshop "The Socio-Economic Relevance of Research Infrastructures". 31 May-1 June 2012, Hamburg. Retrieved on August, 27, 2017, from: http://www.cost.eu/download/Rizzuto

Rokach, L., Maimon, O. (2008). Decision trees: Theory and applications, series in machine perception and artificial intelligence, Singapore: World Scientific Publishing.

Rons, N., DeBruyn, A., Cornelis, J. (2011). Research evaluation per discipline: A peer review method and its outcomes. Research Evaluation, 17(1), 45–57.

Rose, H., Rose, S., Horrobin, D. F. (1971). Science and society. Harmondsworth: Penguin.

Ruegg, R., Feller, I. (2003). A toolkit for evaluating public R&D investment: Models, methods, and findings from ATP's first decade. US Department of Commerce, Technology Administration, National Institute of Standards and Technology.

Salasin, J., Hattery, L. H., Ramsaeus, T. (1980). The evaluation of federal research programs. MITRE Corporation, Metrek Division.

Scott, A. (2007). Peer review and the relevance of science. Futures, 39(7), 827-845.

Segers, P. (2015) Partnership for advanced computing in Europe key performance indicators, GSF-OECD Workshop November, 3, 2015, Paris. Retrieved on February, 25, 2018 from: https://www.innovationpolicyplatform.org/system/files/PRACE_Segers_2.pdf

Smith, L. C. (1981). Citation analysis, Library Trends, 3(1), 83-106.

Spaapen et al. (2011) Social Impact Assessment Methods through Productive Interactions Project Final Report. Retrieved on January, 07, 2018 from: http://www.siampi.eu/Content/SIAMPI_Final%20report.pdf

Stake, R. (1995). The art of case research, Thousand Oaks, Sage Publications.

Sumathipala, A. (2014). Research infrastructure. In S. Okpaku (Ed.), Essentials of Global Mental Health Cambridge: Cambridge University Press.

The Law on Supporting Research Infrastructures dated 10.07.2014 and numbered 6550. Retrieved on September, 01, 2017 from: http://www.resmigazete.gov.tr/eskiler/2014/07/20140710-3.htm

Turkstat (2017). 2016 Research and Development Activities Survey. Turkish Statistical Institute, Ankara.

Türkiye Büyük Millet Meclisi. (2014). Araştırma Altyapılarının Desteklenmesine Dair Kanun Tasarısı Ve Sanayi, Ticaret, Enerji, Tabii Kaynaklar, Bilgi Ve Teknoloji Komisyonu ile Plan ve Bütçe Komisyonu Raporları, 1/189

Tracy, S. J. (2013). Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact. UK: Wiley-Blackwell.

United States Agency for International Development (2013). Technical note: Evaluative case studies. Retrieved on September, 28, 2017 from: https://usaidlearninglab.org/sites/default/files/resource/files/case_study_tech_note_final_2013_1115.pdf

United States Department of Energy, Office of Energy Efficiency and Renewable Energy (2004). EERE Peer Review Guide, based on a survey of best practices for in-progress peer review. Retrieved on September, 30, 2017, from: https://www1.eere.energy.gov/analysis/pdfs/2004peerreviewguide.pdf

United States General Accountability Office, Program Evaluation and Methodology Division (1990). Case study evaluations. Retrieved on September, 28, 2017, from: https://www.gao.gov/assets/80/76069.pdf

University of Minnesota Research Infrastructure Task Force (2006). Transforming the university: recommendations of the Task Force on Research Infrastructure. Retrieved on August, 27, 2017, from: https://conservancy.umn.edu/bitstream/handle/11299/5620/1/acad_design_sci_final_exec.pdf

Van den Brink, M., De Haas, M., Van Del Heuvel, J., Spaapen, J., Elsen, M. Westenbrink, R. et al. (2010). Evaluating the societal relevance of academic research: A guide. Evaluating Research in Context Project, Rathenau Institute Publication. Retrieved on October, 21, 2017, from: https://www.rathenau.nl/en/file/1176/download?token=W88bEIIY

Ware, M. (2008). Peer review: benefits, perceptions and alternatives. Publishing Research Consortium, 4(1), 1-20.

Weimer, D. L., Vining, A. R. (1992). Policy analysis: Concepts and practice Englewood Cliffs, NJ: Prentice Hall.

Werskey, P. G., (1971). British scientists and ‘outsider’ politics, 1931-1945. Science Studies, 1(1), 67–83.

Westat, J.F., (2002) User-friendly handbook for project evaluation, NSF Directorate for Education and Human Resources. Retrieved on January, 11, 2018, from: https://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf

White, H. D. (1990). Author co-citation analysis: Overview and defense. Scholarly communication and bibliometrics, 84(1), 105-106.

Wouters, P. et al. (2015). The metric tide: literature review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management). HEFCE. Retrieved on December, 30, 2017, http://www.dcscience.net/2015_metrictideS1.pdf

WR (2013). Report on the science-driven evaluation of large research infrastructure projects for the national roadmap. Cologne, Germany. Retrieved on November, 18, 2017, from: https://www.wissenschaftsrat.de/download/archiv/2841-13_engl.pdf

WR (2017). Report on the science-driven evaluation of large research infrastructure projects for the national roadmap. Cologne, Germany. Retrieved on November, 18, 2017, from: https://www.wissenschaftsrat.de/download/archiv/6410-17_engl.pdf

Yin, R. (1993). Applications of case study research, Beverly Hills, Sage Publications.

Yin, R. (2009). Case study research. Design and methods. 4th Edition. Thousand Oaks, Sage Publications.

Zuijdam, F., Boekholt, P., Deuten, J. Meijer, I. Vermeulen, N. (2011). The role and added value of large-scale research facilities final report. Technopolis Group. Tallinn. Estonia Retrieved on October, 27, 2017, from: http://www.technopolis-group.com/report/role-added-value-large-scale-research-facilities/