Study of Large Scale Research Infrastructure Impact Assessment: Current Practices and Proposed Way Forward

Submitted to:

National Research Council of Canada

Prepared by:

Performance Management Network Inc.
Steve Montague, PMN
Gretchen B Jordan, 360 Innovation LLC

Date:

March 4, 2015

Table of Contents

List of Acronyms

CCGS	Canadian Coast Guard Ship
CFHT	Canada-France-Hawaii Telescope
CFI	Canada Foundation for Innovation
CLS	Canadian Light Source
DOE	Department of Energy
HAL	Hickling Arthurs Low
HQP	Highly Qualified People
IMOS	Integrated Marine Observing System
LSRI	Large Scale Research Infrastructure
M&E	Monitoring and Evaluation
NRC	National Research Council of Canada
ONC	Ocean Networks Canada
ROI	Return on Investment

Definitions

Impact Pathways	Impact pathways describe a chain of results, showing the linkages between the sequence of steps that result in impact. A theory of change adds to an impact pathway by describing the causal assumptions behind the links along these pathways.
Infratechnology	Infratechnologies are a varied set of technical tools, including scientific and engineering data, measurement and test methods, and practices and techniques, that are used in industry to support industrial research and development (R&D), marketing and production.
Theory of Change	A theory of change is an explanation that sets out the causal links between activities and outputs and the expected outcomes. The Treasury Board Secretariat suggests that a theory of change connects a program attributes to its goals and may include assumptions, risks and external factors. (See Supporting Effective Evaluations: A Guide to Developing Performance Measurement Strategies, Section 5.3 http://www.tbs-sct.gc.ca/cee/dpms-esmr/dpms-esmrtb-eng.asp). Assessing the theory of change will encompass the concept of incrementality – i.e. what happened differently because of the program.
Results Chain	The causal or logical relationship between activities and outputs and the outcomes of a given policy, program or initiative, that they are intended to produce. Usually displayed as a flow chart.
Value Propositions	Value propositions clearly identify the value expected to result from investment in the proposed policy, program or initiative. Value propositions lay the foundation for both performance measurement (monitoring) and evaluation.

Executive Summary

The objective of this project is to develop recommendations and identify best practices for measuring the performance, benefits and impacts of large scale research infrastructure (LSRI) based on a review of existing studies and literature and interviews with key LSRI facilities – with particular emphasis on the unique features of these facilities.

LSRI are research facilities with unique capabilities that serve users through merit-based access, and are usually of a scale or complexity that exceeds the capacity of a single organization, region or nation to fund, build and manage. Facilities may be aimed at fundamental discoveries to drive our understanding, focused on specific missions to address the requirements of their users, or some combination of both. They have both scientific impacts and social and economic impacts.

The literature review and analysis included nearly one hundred documents, including reports, articles, assessments, reviews and other literature. Consultations with 15 individuals representing 11 LSRI facilities and service groups in Canada and internationally complemented this review. Facility managers interviewed were interested in improving the assessment of impacts, considering their current approaches inadequate. In general LSRI report quantitative indicators of what can be counted such as expenditures in the local community, outputs, users, publications, and student and public engagement. Impact assessment is done with success stories and occasional economic cost benefit analysis of selected successes. The literature review of current practice found few suggestions for improvement, and those propose theoretical frameworks for evaluation of LSRI.

This report found that there are three major gaps in current practice and that if these are filled, wider impacts can be assessed earlier in the process and these can increase understanding of the contributions of LSRI as well as inform improvements.

1. The wide variance in the nature and context of LSRI conditions for assessment of performance is not described. The variety of purposes, target audiences, activities and how these activities will lead to desired results given specific existing conditions are seldom explicit. Logic models with underlying causal assumptions have not been done.

2. Related to the first, while LSRI typically consider a number of important quantitative indicators for management, not all of the pathways by which outcomes occur have been clearly explained and pursued. Current practice rarely considers public policy and effects of government-funded technical infrastructure. Also detailed descriptions of common sequences of outcomes currently used in assessment of more ordinary S&T organizations are not used for LSRI. Examples are stages of building trusted relationships with partners or stages of technology development and commercialization.

3. Monitoring and evaluation approaches were found to be somewhat limited and inconsistent when it came to outcomes. Opportunities exist to monitor and track more impact pathways, document sequences of outcomes, and improve on both success stories and periodic assessment with careful case studies building on monitoring data and assessment of progress along impact pathways. In particular, the relationships among the metrics should be assessed to paint a broader picture of impacts.

This report builds on accepted best practice in assessment of organizational impacts and adapts them for LSRI. The essence of this is that assessment answers questions about program/organization logic: why is the program important to science and society, what sequence of results occur from which target audiences, and how did the program influence that through its activities, taking into account differing conditions.

An analysis of the literature, supported by consultations, suggests that there are some common sequences of results from various groups which can serve to outline performance expectations. These can be summarized as impact pathways.

1. Create a Research Structure that Supports Discovery and Innovation.
2. Build Research Capacity – Knowledge Base, Highly Qualified People, Research Tools.
3. Contribute to New Technologies, Competitive Companies, Markets and Clusters
4. Inform Government Policies and Decisions.
5. Inspire Students and Public Appreciation of Science and Technology.
6. Contribute Directly to Local and High Tech Economic Activity.

Once contextual information has been accumulated and assessed, analysts can then establish the selection of impact pathway(s) and monitoring and evaluation approaches appropriate for the type of analysis required.[1] Combinations of impact pathway results logic(s) should be defined, and – in concert with key stakeholders – a logic model (or models) should be constructed along with a performance measurement strategy**.**

In order to appropriately assess performance it is recommended that each LSRI implement a multi-year assessment plan that has three levels of analysis integrated over a period of time from four to ten years depending on funding cycles and LSRI context.

The three levels of analysis are:

1. Routinely collect data on inputs, activities, outputs, and engagement (users, partners, others involved), This includes looking at the quality of facility operations, relevance of research options provided, people trained, outreach events, and research projects enabled. It also includes documenting major external influences (e.g., funding for the whole field, breakthroughs elsewhere). What is collected depends on the LSRI logic model and context.

2. As required, mid-term reviews will assess progress and early effects. This includes looking at the continued relevance of research options provided, research results as measured by publications, collaborations, significance of that research, and user satisfaction.

3. Periodically, larger assessment efforts can determine longer-term scientific and socio-economic impacts, both qualitative and quantitative, by using, updating, and expanding on the data and analysis completed in levels 1 and 2.

In summary, the findings suggest that LSRI operate in complex, dynamic environments. They often serve multiple missions, have a diverse set of stakeholders and several impact pathways. The approach to developing questions, indicators and methods of assessment must be duly considered in this context. This report recommends the use of a basic set of templates relating to performance levels and time frames, key questions for program description and context leading to a selection of impact pathways (and defined theories of change). These pathways should be analyzed by distinct monitoring, periodic review and in-depth assessment strategies. The models and templates offered in this report should be considered to be preliminary, indicative of key principles and hopefully inspirational to further refinement, trial and adjustment.

The recommended approach for LSRI performance monitoring and evaluation is made up of three key elements:

[1] Analysis of any kind should always start with consideration of the decisions to be supported and the issues / questions to be addressed. This is true for officially sanctioned Government of Canada evaluations as well as for other types of review and assessment.

i. Consider key characteristics of LSRI relating to why they exist, what outcomes they produce, who they work with and how they operate.

ii. Outline impact pathways so that performance can be assessed by referring to a sequence of results and appropriate indicators linked to those pathways.

iii. Establish monitoring and evaluation strategies which operate at the level of on-going operations, a mid-term review related to early and intermediate results and periodic impact evaluations.

Canadian funding authorities should consider a pilot trial or trials for this approach for upcoming LSRI assessments. Through such pilots and early trials, the detailed elements of context, impact pathways, indicators and approaches can be refined, improved and shared over time to allow for generative learning in this emerging area.

1.0 Objective

Large scale research infrastructure (LSRI) facilities are among the most sizeable science expenditures and require not only significant investments but also long term planning and commitment – sometimes more than 30 years. In addition to the large initial capital expenditure, annual operating costs are usually high. The size of the expenditure means that these facilities are highly visible to the research community, to the public, and in the political realm where decision making occurs. Researchers present cases for building and supporting facilities in their area of research so that they can undertake cutting edge research, whether it be fundamental, mission-oriented or both. The scientific and technical benefits from these facilities can be large. Considering the investment of resources at a national level, evidence must be provided to demonstrate that these investments have produced significant returns. For this reason, it is important to be able to reliably assess “big science” investments and their impacts.

This assessment seeks to enhance the ability of Canadian agencies with an interest in funding LSRI, and the LSRI facilities themselves, to evaluate the nature and extent of the impacts of LSRI on science, the economy and society and the ways in which those effects are generated.

The objective of this study is three-fold:

1. To review existing studies and literature, and conduct interviews with individuals from domestic and international LSRI, to identify methodologies that measure the value of LSRI, from both the scientific and socio-economic points of view.

2. To identify best practices and systems of measurement that can be applied to LSRI, in order to measure their scientific, technical, and socio-economic impacts.

3. To develop recommendations on performance metrics, indicators and methodologies for the measurement of benefits and impacts of Canadian investments in LSRI, with a focus on:

    a. Metrics to measure the scientific and technical impacts of LSRI, with a particular emphasis on the unique features of these facilities.
    b. Metrics to assess the socio-economic impacts of LSRI.

This report provides observations regarding the nature of LSRI and suggests an approach to classifying the context and value propositions for LSRI in order to appropriately frame performance.

LSRI are research facilities with unique capabilities that serve users through merit-based access, and are usually of a scale or complexity that exceeds the capacity of a single organization, region or nation to fund, build and manage. Facilities may be aimed at fundamental discoveries to drive our understanding, focused on specific missions to address the requirements of their users, or some combination of both. They have world-class capabilities that are globally competitive, and usually draw users and funders from other countries.

The analysis and synthesis for this study has first sought to understand the entities being considered (LSRI in the Canadian context), second to distill the logic behind the key results for these entities, and third, to propose appropriate means to monitor and evaluate performance.

The literature review and analysis included nearly one hundred documents, including reports, articles, assessments, reviews and other literature. . In some cases this approach required going outside the ‘Big Science’ and LSRI literature to examine work related to the monitoring and evaluation of research science and innovation more generally. To complement this review, consultations were conducted with 15 individuals, representing 11 LSRI facilities and service groups in Canada and internationally, in order to assess current practices, develop an appropriate assessment framework for LSRI and establish appropriate monitoring and evaluation approaches. **Appendix A** provides a list of document references. **Appendix B** lists the individuals and institutions consulted.

2.0 Methodology

A literature review was undertaken covering literature and practice in impact assessment of LSRI, roadmapping exercises, specific impact assessment reports, and the annual reports of those Canadian and international LSRI which were interviewed for the study. The bibliography (**Appendix A)** lists the reports and documents considered in this review. The study team focused on identifying broad similarities and differences within the literature and found, as noted in Autio's (2014) review of LSRI and innovation, that the LSRI literature is "sparse, non-cumulative and fragmented..."

In addition to the literature review, consultations (typically an interview with follow-up correspondence) were conducted with representatives of 11 LSRI:

1. TRIUMF, Canada
2. Canada-France-Hawaii Telescope (CFHT)
3. SNOLAB, Canada
4. CCGS *Amundsen*, Canada
5. Compute Canada , Canada
6. Ocean Networks Canada (ONC), Canada
7. Canadian Light Source (CLS), Canada
8. Integrated Marine Observing System (IMOS), Australia
9. Helmholtz Association, Germany
10. Institut Laue-Langevin, France
11. Fermilab, USA

Prior to the interviews, the study team reviewed strategic plans, annual reports, and any larger evaluative studies that had been undertaken related to the facilities. Three broad questions were asked:

1. How do you define successful performance for your facility?
2. What are your current practices – methods and metrics – for results planning, monitoring, measurement and evaluation? and
3. What practices would you see as appropriate in the future to undertake results planning, monitoring, measurement and evaluation for facilities such as yours?

The consulted groups were all very cooperative – often with more than one representative from the LSRI participating in the interview. Overall, Canadian and international LSRI representatives provided similar areas of emphasis and expressed similar concerns regarding the need to better describe performance.[2] These consultations, with both domestic and international LSRI, essentially established the key insights for this study in terms of context, impact pathways and the appropriate selection, collection and use of indicators. Highlight observations are contained in the following sections.

[2] Some differences may be present in terms of international LSRI showing, on margin, a higher emphasis on international panels of experts and for two of the four international LSRI – a stronger emphasis on top-down target setting.

3.0 Scope

3.1. Scope of Impacts Covered

This study focuses on the assessment of impacts, with particular attention to metrics related to the unique features of LSRI. Impacts are considered in two broad areas:

a. Metrics to measure the scientific and technical impacts of LSRI
b. Metrics to assess the socio-economic impacts of LSRI.

Some contextual factors have been included that are known to influence the success of LSRI, since these are often monitored by facilities for other purposes (e.g., operational planning). Assessing these contextual and process aspects enhances the learning that can arise from an impact assessment. When the assessment takes place at an early phase, such learning can allow for mid-course corrections.

3.2 The Nature of LSRI

In the literature, LSRI are typically differentiated based on such criteria as: single discipline vs. multi-disciplinary; and single location vs. mobile, multiple location (distributed) or virtual. Consultations, combined with the study team's experience in research evaluation, suggested that these designations are insufficient for the purpose of understanding LSRI and their impacts (see section 5.1). Rather than categorizing LSRI by facilities type, LSRI may be more effectively considered in terms of four simple questions, the answers to which define the logic of the LSRI effort:

- Why? (why was it created and why does it continue to exist?);
- What? (what role does it play for its users and others?);
- Who? (who are the stakeholders that engage with and benefit from this facility?); and,
- How? (how does it operate, train, connect, etc.?).

This approach was found to be important in order to appropriately categorize and assess LSRI and their impacts. Of particular importance are the context of the program and its design, and the external influences that are likely to help or hinder the successful achievement of the desired outcomes. Major influences surrounding these facilities are often not included in assumptions about how programs will perform. These influences are revealed only during the implementation of the program (LSRI), if progress is assessed regularly, and may change over time. As such, they warrant monitoring. Also, the tracking of major external events, scientific and otherwise, will be useful over time in preparation for a future assessment which needs to differentiate the benefits that can be attributed to the contribution of the LSRI (the value from the LSRI's impact pathway) from other plausible explanations based on external influences. Some of these contextual dimensions are considered in **Table 1**.

Table 1 is intended as a tool that can be used as part of an initial LSRI assessment to assist in selecting the appropriate impact pathways, indicators, relevant benchmarks or targets for indicators of the scientific and socioeconomic performance of LSRI and the timing and type of review processes to be used. It also identifies potential external influences to consider and monitor to assist in attempts to attribute outcomes to the LSRI. Assessing contextual and process aspects enhances the learning that can arise from an impact assessment. When the assessment takes place at an early phase, such learning can allow for mid-course corrections.

Table 1: Key LSRI Performance Dimensions		
Performance level	**The Nature of Impact Pathway(s) – Value Propositions**	**External Influences** (Examples)
Ultimate Outcomes / Impacts (Why?)	• Strategic science and technology (S&T) leadership, platform for discovery, innovation; • S&T capacity (knowledge, tools, technology, highly qualified people (HQP)); • Contributions to the economy (local, direct industry, infratechnology, indirect industry or policy); • Contributions to societal well- being (direct through policies influenced; indirect through industry or policy action).	• Global S&T spending in field; • Serendipity; • General economic conditions; • Shifting societal challenges.
Immediate + Intermediate Outcomes (What?)	Outputs/outcomes dependent on: • Role and place on spectrum of scientific discovery; • Technological innovativeness and applicability elsewhere; • Users' needs, range of applications; • Place within larger research community; • Hub effects; and • Extent to which LSRI engages in educational activities.	• Speed of evolution in the S&T field; • Progress elsewhere; • Absorptive capacity in application areas, and communities targeted via outreach.
Engagement (Who?)	• Extent of sector and/or discipline specificity; • Degree of cohesiveness in research decision-making; • Extent that LSRI links into international science networks; • Extent to which industry is directly involved; • Extent to which governing bodies are directly involved.	• Global trends in this area; • Political influences on collaboration; • Sector capacity and readiness to engage.
Inputs and Activities (How?)	• Scientific field(s) involved; • Degree of centralization, capital intensity, economies of scale and scope; • Degree of LSRI participation in research; • Existence of natural by-products of the research or location. • The operating time and life cycle of the LSRI.	• Existing national strengths, strategy; • Speed of evolution in the S&T field; • Authorities, governance and overall management norms.

These dimensions can be used as a list of features to characterize the nature of an LSRI and help to frame the appropriate choice and weighting of performance indicators, and the approaches used to monitor and evaluate performance.

For instance, an astronomical observatory will be geared to fundamental science discoveries with the global astronomy community as its user base. Technology spillovers occur from industry involvement in constructions and development.

Like most LSRI, the observatory needs to consider operational efficiencies and economies. It also focuses on the careful validation of the quality of data recorded for the global astronomy community. One of its key challenges is the need to balance consistent levels of service with offerings of new features such as UV panoramic imaging. The LSRI knows that it is adding value and serving a niche by noting its international client base (who have options to go elsewhere) and through its contribution to published work. (Which is tracked systematically – see Crabtree 2001 updated in 2014)

By contrast, a distributed or mobile LSRI dedicated to environmental monitoring may have a user community of governments, policy makers and policy implementers – as well as private groups in resource extraction or transportation dealing with the natural environment. The key focus will be on providing a reliable, high quality platform.

Each of these LSRI can be seen to have a distinct performance framework, characterized by their contextual elements (including external factors) relating to inputs, activities and outputs (How?), engagement (Who?), direct outcomes relating to user and stakeholder behaviors (What?) and impacts or ultimate outcomes (Why?).

4.0 Summary Findings from Review of Literature and Interviews

4.1 LSRI Context

The literature review suggested that the essential LSRI context categories related to the type of facilities and their discipline included: 1) Single Discipline Facilities; 2) Multidisciplinary Facilities; and 3) Distributed Facilities.

Interviews with individual LSRI suggested that the context of LSRI is much more nuanced than is typically acknowledged in the literature. LSRI were found to be single location, distributed and virtual, and single or multi disciplinary (as per Technopolis (2013) definitions) but there were also significant differences in vision, mission and strategy, and differences in outputs and early outcomes, depending on who is engaged and the types of activities undertaken. All of the LSRI examined aim for scientific and technical leadership and contributions to the knowledge base and the training of HQP. Likewise, all expressed that they expected the research they enable to lead to innovation, more competitive industry and local economic benefits.

All of the LSRI interviewed work directly with industry in the development and building of new facilities and equipment. Some, however, are directly involved with industry in their operational activities, either because they supply a product, such as medical isotopes (e.g., TRIUMF and CLS) or, like SNOLAB, they cooperate on research that is of interest to a particular industry where the application of data management techniques to collect, filter and analyze mining and exploration – related data and serve as a data processing facility then can assist mining and exploration companies.[3] Some, like ONC, work directly with government decision-makers and indigenous populations on collecting and making available data of interest. Others, like Compute Canada, cooperate with industry to enable simulation and big data analysis across many scientific fields. The CCGS Amundsen goes beyond multidisciplinary and is trans-sectoral, working across academia, industry and levels of governments.

4.2 Theories of Change

In the words of one recent review of LSRI (Autio 2014) there is a "lack of guiding theory" for LSRI and innovation impact assessment. The science and innovation literature, in general, suggests theories connecting scientific research to impacts on the science itself, on marketplaces and on policies – but these outcomes are not unique to LSRI.

As part of consultations, it was revealed that at least one international LSRI, Australia's Integrated Marine Observing System, has focussed on the take-up and use of the information they generate and has began to consider impact pathways.[4]

Interviews with other LSRI representatives confirmed that impact pathways (other than the 'Research Structure' pathway they all have in common), such as building unique research tools, knowledge, and qualified people are present in most LSRI, although they are most often not stated explicitly or well-defined. Rather, they tend to be implicit (to be inferred from indicators) rather than lay out explicitly, either verbally or graphically. All of those interviewed noted the contribution of the infrastructure (facility, data,

---

[3] For a full explanation see http://www.thesudburystar.com/2013/10/12/project-to-mine-data-good-for-snolab; http://www.snolab.ca/news/2013-10-11-new-mining-exploration-data-centre-snolab

[4] The Impact Pathways for IMOS relate to research, evaluation and training, research projects and programs, multi-decade analyses, remote sensing products, research modelling systems and operating forecasting systems – all 'fed' by IMOS data and products. See presentation by Tim Moltmann, *Performance measurement for an ocean observing system – perspectives from Australia's IMOS* (University of Tasmania), November 2014

equipment) to S&T, and how these contributions are unique and serve a particular niche. All respondents highlighted the importance of LSRI to enabling and facilitating both international collaboration and access to other LSRI around the world. Respondents were aware that the training of researchers, particularly graduate students performing research at the LSRI, led to an increase of HQP both within and outside of, academia. Some facilities have taken a more active approach to outreach with students and the public, or with specific populations. Those who work directly with policy makers are aware of the impacts of their work on policy – however policy impacts did not figure prominently in either interviews, discussions nor in related reports.

The impact pathway with industry has multiple strands, both direct and indirect. The direct standard relates to the local economic impacts associated with infrastructure construction, servicing and maintenance. The indirect strand relates to industrial and marketplace development and use of products and services resulting from LSRI work. Those interviewed did not appear to distinguish between these two strands, although it would seem safe to assume that it would be much easier to follow progress towards a given impact when there is a direct relationship with industry (either from facility upgrades or industry use of the facility). Overall, impacts from new products or technologies were recognized but few LSRI were looking for changes in processes and only one mentioned that a change in practice (management of large facilities) had resulted as an outcome of interactions with industry. None of the LSRI were looking for the particular impacts of infratechnologies, including standards, on industry.

4.3 Impacts Investigated Differ Across Studies

In the literature there is considerable diversity in the specific impacts measured in LSRI impact assessments, reflecting the lack of an agreed upon program theory. The LSRI Roadmaps reviewed are projections rather than assessments but the criteria for selecting new infrastructure shed light on expected performance. Criteria include scientific significance, fit with and added value to national S&T goals and leadership, use by a large portion of one or more S&T communities, training of researchers, and added value to industry or the common good either in the construction phase or the longer term. The CFI/MSI Outcome measurement study has three Core Categories: Capacity Building, Leadership and Research Enabled, and Extrinsic Benefits. Two different schemes for assessing return on investment are suggested in a 2011 framework where Astronomy is the example and 2013 where TRIUMF is featured. In 2011 five types of socio-economic impact were assessed: Knowledge generation, Knowledge Use, Development of HQP, Contribution to Partnerships, and Contribution to Innovation. The 2013 framework was organized around three levels of expected returns: Disciplinary (Excellence, Community Support, Impact, Partnerships); Cross-Disciplinary (Relative Science Impact, Synergy, Urgency); and Strategic (Scientific competitiveness, Economic, social benefits, International leadership, National priority).

The two recent literature reviews pay attention to the unique nature of LSRI. Technopolis (2013) suggests four impact pathways: Purchase/development of advanced equipment; Allow research on new questions achieve breakthroughs; Provide access to unique, scarce equipment, data for diverse user base; and Provide focal point for clusters of scientists, high tech companies. Autio (2014) suggests that LSRI contribute to innovation by: Scientific research, Human resource development, Industry collaboration, Spin-off and spin-out companies, Infrastructure building and maintenance, International collaboration, and Service provision.

4.4 Key Indicators

The findings from the literature review and LSRI consultations suggest that there are three types of indicators for LSRI:

i. Indicators used for operational management;
ii. Quantitative and qualitative indicators of reach and selective outputs and outcomes; and,
iii. Emerging indicators, which could more rightly be called evaluative reviews, relating to socio-economic and possibly other impacts.

i. **Indicators used for operational management** tended to focus on the following:

- Facility efficiency
    - Examples include: down time (sometimes adjusted by external or internal factors); 'booking' (subscription and oversubscription) rates; delivery according to agreed-upon timelines; overhead costs and other costs related to facility or equipment use; and wait times for set up between tests.
    - Project management metrics, in general, related to the above, appear to be emerging as more prominent among LSRI operational indicators.
- Quality indicators
    - These may be related to technical/equipment faults, re-test requirements, accuracy measures, reliability measures, etc.
    - How does the equipment compare to the state of the art (quantitative and measures and expert opinion); number of international researchers attracted.
- Human resources-related measures
    - Examples include: tracking safety in terms of time lost due to injury; staff morale; absenteeism; and the ability to attract new hires, etc.

ii. **Quantitative and qualitative indicators of reach** and selective outcomes were found in both the literature review and in interviews and tended to focus on the following:

- Reach (Users):
    - The number (and percentage) of domestic and international (world class) users was very important for many groups.
    - The reach of the LSRI within various types of communities was tracked with varying degrees of consistency.
- Selective Outcomes:
    - Formal user satisfaction surveys were conducted by some LSRI, but this was not necessarily the norm.
    - Quantitative science metrics such as publications and citations related to LSRI use was systematically tracked for some groups – but not others.
    - Other quantitative metrics are used in some cases related to the number of highly qualified people (including students, academics, researchers) using the LSRI.

iii. **Emerging indicators related to socio-economic impact:**

- Areas mentioned in consultations related to the idea of tracking economic impacts (economic impact studies had been conducted by some LSRI – but the credibility of some of the impact numbers used was called into question by facility proponents themselves in some cases) and, more generally, several respondents noted that case studies and qualitative results stories were also important areas to be developed.[5]
- As noted above, in a few cases, LSRI are seeking links between outputs and further outcomes and impact. Some track user satisfaction with services. Where researchers apply for continued funding, they are asked to report on use that has been made of previously funded research. Some have been tracking links between use of LSRI equipment or services and use of that information in publications. A few have tried to follow the career paths of people trained at the LSRI.

[5] No 'best practices' of case studies were offered in consultations – nor found in the LSRI literature *per se*. Good practices for research related case studies may be considered to generally include a strong discussion of impact pathways or a theory of change, contextual factors which affect results and a systematic approach to assessing the logic sequence of events and the plausible contribution of research to outcomes. See Mayne and Stern (2013).

Metrics mentioned as collected by most LSRI representatives interviewed included the following:

Operational Management

- Facility operations and quality of operations, such as hours of available beam time, reliability, and unique or special characteristics of the operation, or amount and qualities of data available.
- Amount and quality of technical support services, such as set up and trouble-shooting equipment and *ad hoc* consulting.

Reach and Selective Outputs and Outcomes

- Number of users of the LSRI and their characteristics, compared to requests for use.
- Number of collaborations, location, organization, funds or in kind contributed, fees paid.
- Access to, and participation in, international projects or facilities
- Outputs such as professional conferences and events hosted, education and outreach events, level of participation
- HQP trained, by level of education and/or research position.
- Publications, co-authorship, citations, compared to expected rate or those of peer organizations.
- Professional recognition, awards.
- Media mentions.

Indicators of social or economic impact

- LSRI employment and spending in the local/regional area, and with high technology firms on equipment upgrades.
- Success stories, as these become visible, including patents, spin-off companies, technologies developed, policies influenced, or students inspired to pursue a career in research.
- Economic benefit- cost analysis of selected cases.

For most of the Canadian LSRI consulted, the Canada Foundation for Innovation (CFI) requirements featured prominently in their indicator selection. One such schema is shown below:

	Example: CFI Indicator Requirements
Business Development	Number Of Users
	Number Of Highly Qualified Personnel
	Number of Graduated Students
	Total LSRI Employees
	Number of Scientific Contributions
Technical Contributions	Patents Filed
	Spin Offs
	Networking Meetings On Economic/Commercialisation
Level of Use	Number of Projects
	Programme value
	Visits – Staff
	Visits – Users
	Visits – Visitors
Reliability and Service	Percentage of Experiment-Related Deliverables Met Within Agreed Timescales
	Number of Corrective Actions
	Percentage Corrective Actions Closed Within Prescribed Timelines

	Example: CFI Indicator Requirements
	Number Of Shifts Lost Due To Events Within LSRI Control
	Number Of Shifts Lost Due To Events Outwith LSRI Control
Physical Safety	Number of Lost Time Injuries (LTI)
	Lost Time Injury Frequency (LTIF)
	Number Injuries and Illnesses Reported
	Number of Incidents Reports, including Near Miss
Programme and Staff Development	Number Workshops/Meetings Held At LSRI or Sponsored
	Number Of Professional Presentations from LSRI Staff
	Number of High Esteem Engagements by LSRI Staff
	Number Of New Partnerships/ MOUs
	Number of New Projects
	Expressions of Interest
	Person Days Of Training

4.5 Approaches and Methods of Assessment

Consultations suggested that the most common approach to assessment found in LSRI was the use of quantitative operational management and selective impact metrics reported periodically – usually at the request of funders. As the summary below, shows the types of approaches which were most common include operational assessments of quality, efficiency and related project management combined with selective use of approaches to assess science impact through bibliometrics or more anecdotally through S&T leadership indicators such as awards and other use of expert judgment. User satisfaction surveys were found to be used – but not consistently. A key approach gap is the fact that LSRI often capture success stories anecdotally and conduct follow-ups sporadically. This is also the case for return on investment and cost-benefit studies, which were not a matter of routine and, in some cases, vary in structure and approach discipline. Essentially, LSRI do not always follow a systematic process to show a plausible cause-effect connection between what the LSRI does, who it reaches and what difference it makes. In many cases, LSRI appear to capture and report indicators with putting them into the context of an overall performance story.

Common assessment approaches across LSRI:

- Self Assessment. LSRI carefully monitor quality of operations, reach and selected outputs and outcomes. . They routinely make assessments to understand their national and global competitiveness. S&T leadership is demonstrated through multiple indicators (awards, professional service, attracting and retaining researchers, and S&T community support.
- Bibliometric analysis. Bibliometric analysis is always conducted, although it can be a challenge to locate all of the publications associated with research done using the LSRI in the case of merit-based access to the facility (a few use data mining techniques).Counts are undertaken of publications. Quality is often inferred from journal impact factor and co-authorship is examined to investigate collaboration patterns in terms of organization, location, and discipline. Some LSRI benchmark against similar facilities and a few calculate an h-index ranking (which experts do not consider credible for individual researchers).
- Success and impact stories. All LSRI collect anecdotes about success and impacts in scientific and socio-economic areas, but few complete the verification or expansion steps on how these successes occurred that would be done in a case study approach to validate attribution. A few LSRI conduct follow up studies to see what happened as a result of their research, training or outreach activities. Most LSRI have calculated the return on LSRI investment within local and regional economies. Several have done formal cost-benefit studies using economic modeling of

selected cases of success. Like all cost-benefit studies, these require many assumptions and documenting the attribution of a portion of the benefits to the action of the LSRI was not done in the studies we reviewed.

A number of respondents expressed a wish to go farther to connect their work to longer term results. For the most part interviewees were not clear on exactly how one would attribute longer term results to LSRI work.[6] In some cases it would seem that more systematic, overall qualitative judgment may be advisable over indicator table reporting.

In one international case the LSRI respondent (from a 'discovery' oriented LSRI) laid out the highly summarized and qualitative questions they ask of an esteemed international panel as follows:

i. Questions concerning scientific quality and strategic relevance
   a) Does the facility make it possible to carry out excellent scientific work? How do you rate the facility on a national, European or international level?
   b) How do you rate the scientific results produced by the facility?
   c) How would you judge the technical design and implementation?
   d) What role does the facility play in national or international roadmaps in the respective research field?

ii. Questions concerning users
   a) How would you judge the technical and scientific support provided for external users?
   b) To what extent does the facility explore the opportunities for interdisciplinary use?
   c) Do the procedures ensure equal access to the facility for scientists from the XX Association and for external users?
   d) Are the present staff qualifications and management structures appropriate to a user-oriented facility aiming to meet international standards?

iii. Questions concerning the appropriateness of resources used and future costs
   a) Is the cost planning for the five-year period realistic?
   b) Taking into account the cost, the planned availability and the scientific demand, is the facility operated on a sufficient level to meet requirements (running time and average use)?
   c) Do you consider the estimates of the remaining life span of the facility to be realistic? What is your assessment of the plans concerning a replacement of the infrastructure and its possible costs? Do you see potential financial risks?

While the above questions are no doubt raised in Canadian LSRI reviews, these kinds of overall questions rarely appear in the documentation reviewed surrounding assessments made available for this study which included annual reports and some review materials.

High level panel assessments and other approaches that were not found in LSRI assessments but that would appear to have merit for LSRI are summarized below. These include the use of case studies that trace the contribution of science investments along an impact pathway or results chain and network analysis may be used alone or as part of case studies to show the connections over time which are being made (connections can be tracked via research agreements, contracts, joint studies, co-publications or other less formal interactions, etc.). As one international LSRI respondent said, *"We are a key node in the science network".* If this is true, then analytical approaches examining the role of LSRI as nodes in international science networks would seem to represent an obvious and key area of exploration – likely to become more important in the very near future.[7]

---

[6] The notion of attribution or more correctly 'contribution' in science related activities has been explored in certain Canadian evaluation studies and is discussed by Mayne and Stern (2013) as best done using a theory based approach such as contribution analysis – applied on a case by case basis.
[7] Note the recent rise of the area of Social Physics – developing via analysts like Pentland from MIT who use rich interaction databases to predict innovation success, productivity and other desirable social outcomes.

Other approaches that may be applicable include:

- Analysis of monitoring data within context considering trends or measuring against expectations. A few research organizations assess progress on S&T goals against expectations that have been proposed by an expert panel as performance that would represent significant achievements.
- Assessing international leadership in the field (link to relevance). A methodology for assessing international leadership in a field was developed by a U.S. National Academy panel. Reviewing considerable data, panels of experts determine whether a country is the leader, a close follower, or has the capabilities to absorb breakthroughs made elsewhere.
- Agglomeration effects. There have been attempts to measure the agglomeration effect of a facility, that is, its influence on the location of suppliers and other research organizations in a cluster near the LSRI, and the effects of the cluster on the community and region.
- Contribution analysis. Contribution analysis has been applied to research programs with success (Mayne 2012). This is a form of attribution analysis that demonstrates the program's contribution to outcomes by showing events and accomplishments along a program logic pathway as well as by investigating plausible alternative influences on the outcomes.
- Historical tracing between science and technology. Historical tracing between science and technology has been applied in the past few years by the U.S. Department of Energy (DOE) and others. In this technique, mixed methods including interviews, document review, publication and patent analysis, and market analysis are used to trace connections either forward from a scientific breakthrough to technology development utilizing that, or backward from a technology to the science underlying it. This technique can be used to supplement a cost-benefit analysis.
- Network analysis. Network analysis has become a popular way to demonstrate changes in participation and collaboration over time. Data can be partnerships, joint projects, or co-authored papers. Software is used to draw relationship diagrams, and if an entity is a hub or "node" in the network that is noticeable. The method is in its infancy for there is not yet theory on what a network diagram should look like to demonstrate specific objectives.

4.6 Gaps in Current LSRI Impact Assessment Practices

Three significant areas of limitations or gaps were found in the LSRI performance assessments and reports reviewed:

(1) The wide variance in the nature and context of LSRI conditions for assessment of performance is not described. The variety of purposes, target audiences, activities and how these activities will lead to desired results given specific existing conditions are seldom explicit. Logic models with underlying causal assumptions have not been done.

(2) Related to the first, while LSRI typically consider a number of important quantitative indicators for management, not all of the pathways by which outcomes occur have been clearly explained and pursued. Current practice rarely considers public policy and effects of government-funded technical infrastructure. Also detailed descriptions of common sequences of outcomes currently used in assessment of more ordinary S&T organizations are not used for LSRI. Examples are stages of building trusted relationships with partners or stages of technology development and commercialization.

(3) Monitoring and evaluation approaches were found to be somewhat limited and inconsistent when it came to outcomes (see section 4.4). Opportunities exist to monitor and track more impact pathways, document sequences of outcomes, and improve on both success stories and periodic assessment with careful case studies building on monitoring data and assessment of progress along impact pathways. In particular, the relationships among the metrics should be assessed to paint a broader picture of impacts.

In summary, LSRI have suffered from a lack of systematic analysis in terms of frameworks which will address the question of how these facilities contribute to sequences of results for various groups in differing conditions. The result has been isolated quantitative indicator reporting and analysis which requires a great deal of uninformed inference in order to judge success and practically eliminates the ability to learn what works for whom in what conditions and why.

5.0 Identification of Best Practices and Systems of Measurement

Sections 3 and 4 identified the highly variable context and the variance in practices found in LSRI impact assessment.

This section turns to addressing two of the three gaps found in this section.[8] Section 5.1 addresses the relative absence of clarified impact pathways. Section 5.2 addresses related inconsistencies in performance questions, indicators and methodologies.

5.1 Proposed Best Practices – LSRI Logic Model with Impact Pathways

Notwithstanding wide variances in the context of LSRI, consultations and literature review show that there are some common impact pathways which can still serve to outline performance expectations. The following proposed model represents a distillation of the pathways observed during the conduct of this assignment.

The elements of the proposed LSRI logic model include activities, outputs, reach, and a sequence of outcomes for six impact pathways. The model also suggests important aspects of context to be considered, given the diversity of LSRI, their research areas and potential applications.

Pathways have been elaborated to show areas of impact from direct engagement with target groups. These impact pathways are simplifications to demonstrate the bigger picture. Not all LSRI engage along every pathway. Study team research and consultations have suggested that all LSRI engage with the research community and researchers on projects and do outreach to students and the public. All work with industry (including industries who are also users of facilities) on facility upgrades and some on research or products of the facility, such as data or isotopes. A few will have impact through their engagement with government agencies and policy makers. These six pathways are in reality interconnected, often ‘recursive’ and complex. They are not separate and linear – though they are depicted as such in **Figure 1** for illustrative and practical use. The pathways can be considered as:

1. Create a Research Structure that Supports Discovery and Innovation.
2. Build Research Capacity – Knowledge Base, HQP, Research Tools.
3. Contribute to New Technologies, Competitive Companies, Markets and Clusters.
4. Inform Government Policies and Decisions.
5. Inspire Students and Public Appreciation of Science and Technology.
6. Contribute Directly to Local and High Tech Economic Activity.

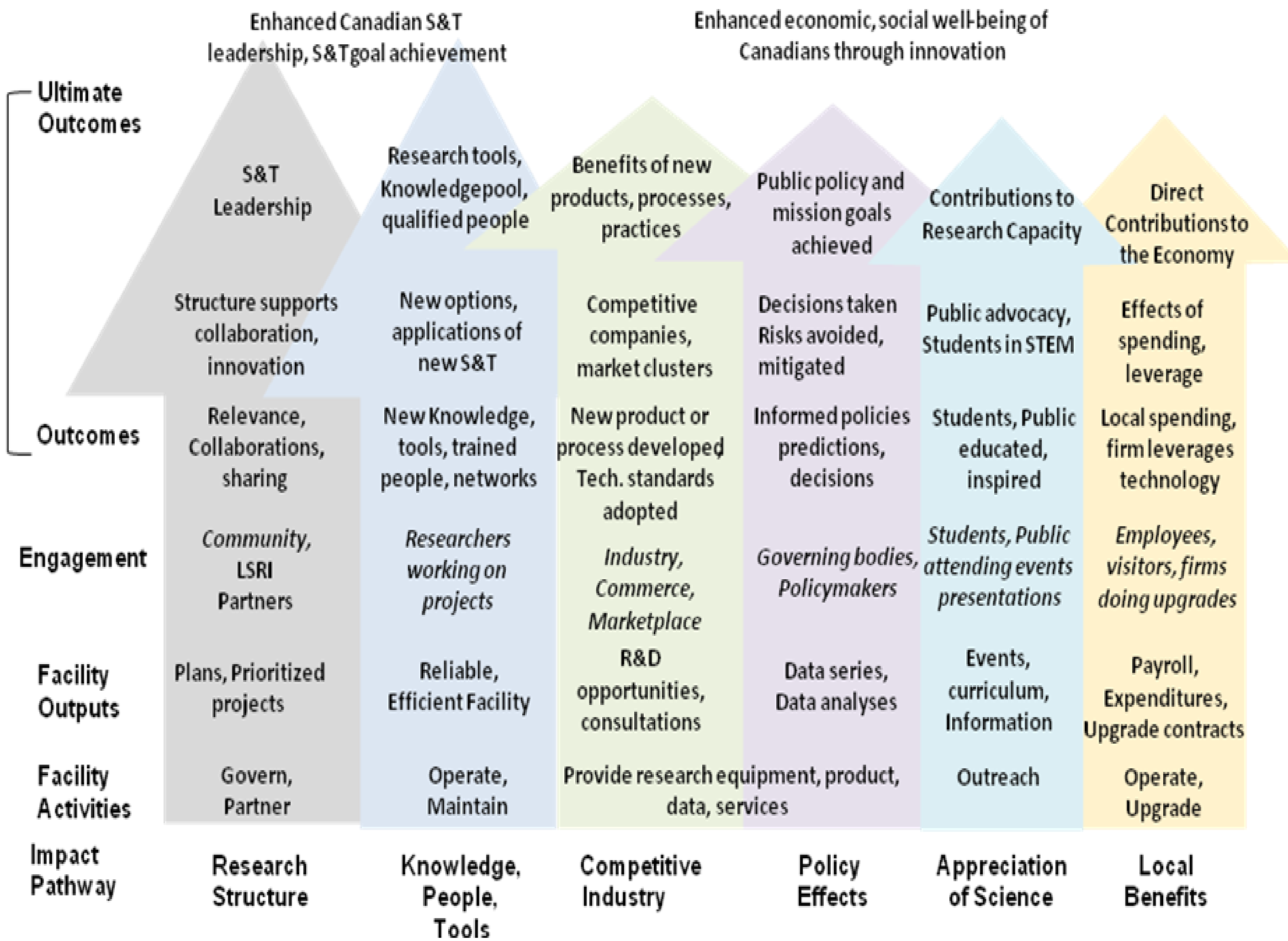
The following diagram outlines the six impact pathways together as basic result chains proceeding from facility activities through outputs, the reach and engagement of key stakeholders[9] and then through a

[8] The first context gap is actually illustrated in Table 1 and is addressed in section 6.

[9] The concept of reach and engagement has been explored in research, technology and development and technology for years. See Montague and Teather, *Performance Measurement, Management, and Reporting for S&T Organizations – An Overview*, Journal of Technology Transfer, Vol. 22, Summer 1997, pp 5-12., Canadian Academy of Health Sciences (2009). *Return on Investment in Health Research 2009. A Preferred Framework and Indicators to Measure Returns on Investment in Health Research* and more recently by Mayne and Johnson *(ibid)* 2014. Emerging work suggests that the building of reach into theories of change can be vital to science-based performance.

sequence of behavioral outcomes leading to ultimate outcomes that relate to the LSRI S&T leadership, goal achievement and socio-economic outcomes.

Figure 1: A Generic Logic Model with Six Impact Pathways



The following sections address each pathway in some detail.

5.1.1 Pathway: Create a Research Structure That Supports Discovery and Innovation

In an era when science and technology are seen as key to solving national and global challenges –the knowledge economy – LSRI play an important role. LSRI exist because they open possibilities for research not otherwise possible because of scale and scope or extreme environments. These characteristics of LSRI require governance, partnership and network building, policies of access, and operations and maintenance, all of which bring further benefits. Access to partners and facilities in other countries through an LSRI is an important benefit of Canadian facilities. Shared equipment and data takes advantage of economies of scale, saving user time and money. Plans and community involvement are necessary to perform the most relevant research. Operation of an efficient, reliable and supportive research platform is both technically difficult and essential. The scale of the effort also provides a focal point for researchers which can create a critical mass in a research area or stimulate a cluster of high technology businesses that work with or support the facility. Sharing of resources puts participants in proximity to other disciplines, thus increasing interdisciplinarity and transdisciplinarity (blending of disciplines).

The engineering and organizational models required for LSRI also bring together cross-functional teams, to design, build, and operate equipment. The process of bridging a diversity of disciplines and functions through teams has been shown to foster discovery and innovation. Technical infrastructure enables the development of a myriad of new products, processes and practices in both private and public sectors. Discovery occurs on the margins of fields. Technology development, commercialization and adoption are faster when each step anticipates the requirements and constraints of the next step.[10]

Some of the key characteristics of research and technology infrastructure have been articulated by Tassey (2008) to include a wide variety of infratechnologies and associated standards which are essential to conduct R&D. Examples of infratechnologies are measurement and test methods, process and quality control techniques, evaluated scientific and engineering data, the technical basis for product interfaces and the ability to communicate research results in an unambiguous manner. Tassey notes that technology infrastructure tends to be ubiquitous, somewhat invisible and displays the characteristics of both a public and a private good (because its use can help communicate across groups on a common basis so each user can benefit from the use, as well as providing a competitive advantage to those who link to it more quickly than others.) Calculations would include the science productivity impacts of infratechnologies like databases, standards, benchmarks and means to interface results (Link 1996 and Tassey 2008).

5.1.2 Pathway: Build Research Capacity – Knowledge Base, HQP, Research Tools

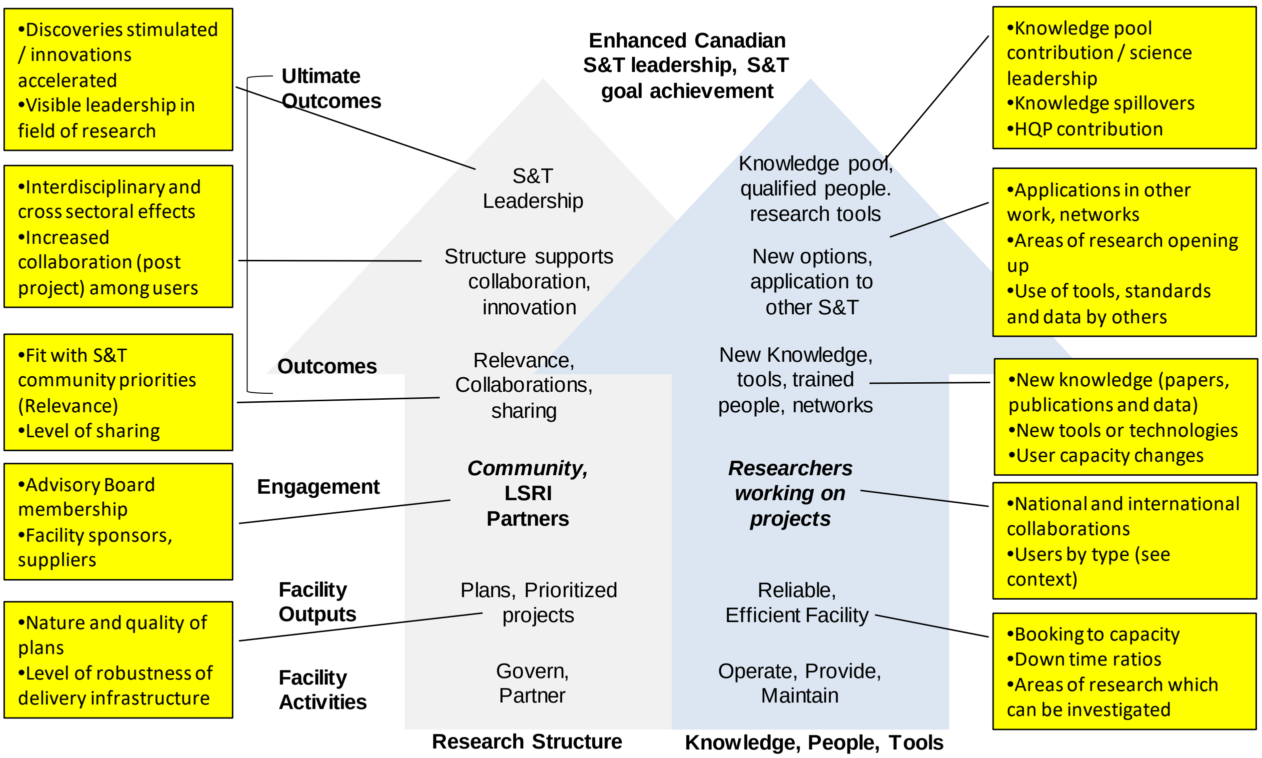
This pathway relates to science, knowledge and know-how production. Much has been written about how to assess both the process and the science outcomes. While advances in knowledge and new or improved research tools are usually discussed together, additions to knowledge in the heads and skills of (highly qualified) people is sometimes separated from the larger term “research capacity”. The pathway is a combination of building and improving research tools (equipment, databases), and the selection, implementation and findings of research projects using those tools. Support staff train people on the equipment and research processes and adds to their knowledge of the field and problem area. As part of research processes involving the use of LSRI facilities, students also are presented with an opportunity to meet and learn from leading scientists and each other. They carry this tacit knowledge, and the resulting social connections, with them whether to careers in academia or elsewhere. The pool of research capacity is available to the research community, industry, government and the public. Capacity not only drives Canadian scientific leadership but also allows Canadian researchers to absorb important research

[10] Key theories of change which inform this pathway include big science theories (Weinberg 1967, Simmons *et al* 2013, Autio 2014) innovation diffusion theory (Rogers 2003), spillover calculations (Jaffe 1986), direct and indirect pathways (Ruegg 2000), and direct attempts to apply economic social and private return theories (Mansfield *et al* 1977).

results from elsewhere (to be 'fast followers'). Capacity has been a focus of performance frameworks for many science programs, including the National Science Foundation in the U.S. and the Natural Sciences and Engineering Research Council (NSERC), National Research Council Canada (NRC), Canada Foundation for Innovation (CFI) and various others in Canada.[11]

Research and consultations suggested that the governance and operations of LSRI are inextricably linked. For this reason **Figure 2** shows these two 'core' pathways together – along with some proposed indicators.

Figure 2: Create a Research Structure That Supports Discovery and Innovation and Build Research Capacity – Some Select Metrics



These pathways follow a science support and impact logic – enabling better science, HQP networks and knowledge pool improvements. Qualitative contextualized approaches (e.g. case studies) will be important to complement quantitative indicators.

5.1.3 Pathway: Contribute to New Technologies, Competitive Companies, Markets and Clusters

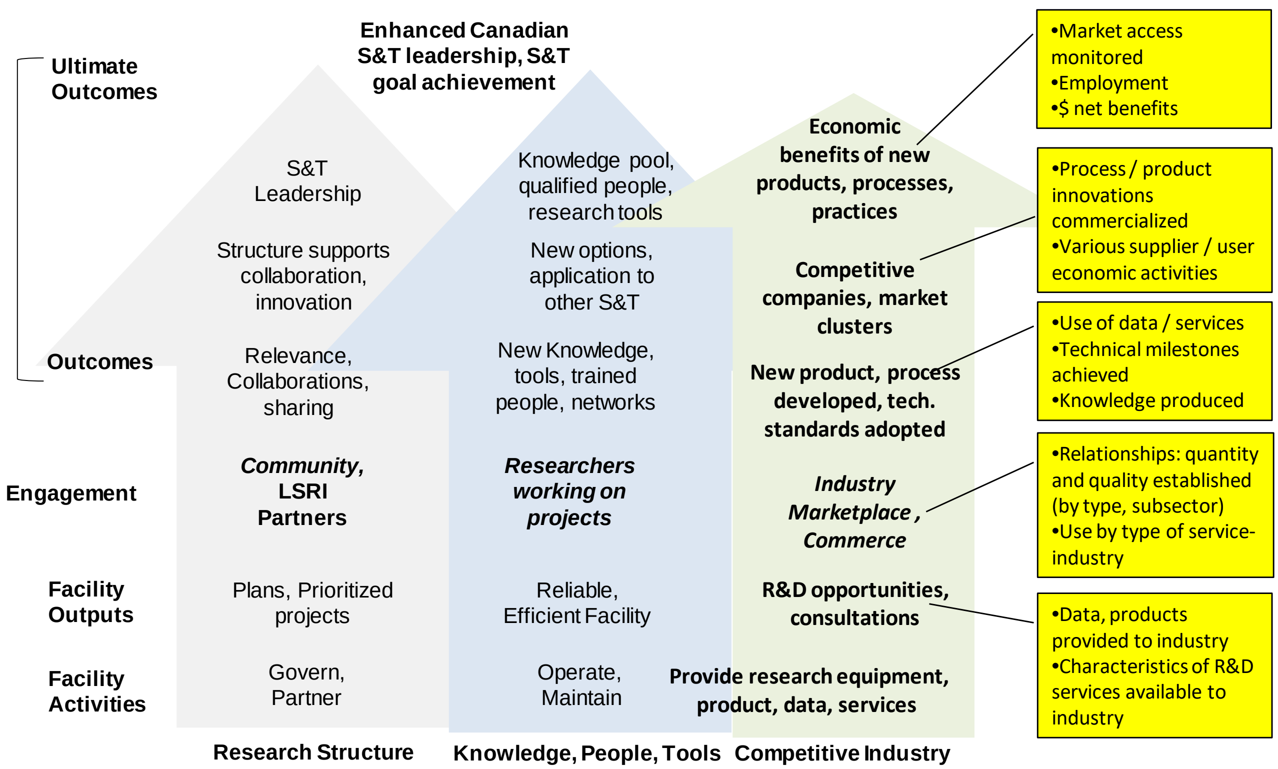
This pathway involves engaging and supporting companies, directly and indirectly, to innovate and develop products and processes which generate private returns for (Canadian) companies and for those employed by them as well as potential benefits to consumers through higher quality, more accessible and

[11] Key theories of change which inform this pathway include innovation theory such as Kline and Rosenberg (1986), some of the innovation diffusion and pathways theories noted above, as well as the notion of social capital influences (Autio *et al* 2004, 2014). Elements of socioecological theory could also be considered when examining this pathway (Bronfenbrenner 1979).

lower cost goods and services. More broadly, such development impacts may contribute to clusters of economic development.[12]

Figure 3 shows this pathway with some possible indicators.

Figure 3: Contribute to New Technologies, Competitive Companies, Markets Clusters – Some Select Metrics



This pathway follows a product-process commercialization (marketplace innovation) logic. Key impacts relate to the benefits of commercialization and innovation such as employment and net financial benefits to industry. The potential for this pathway varies extensively by LSRI.

5.1.4 Pathway: Inform Government Policies and Decisions

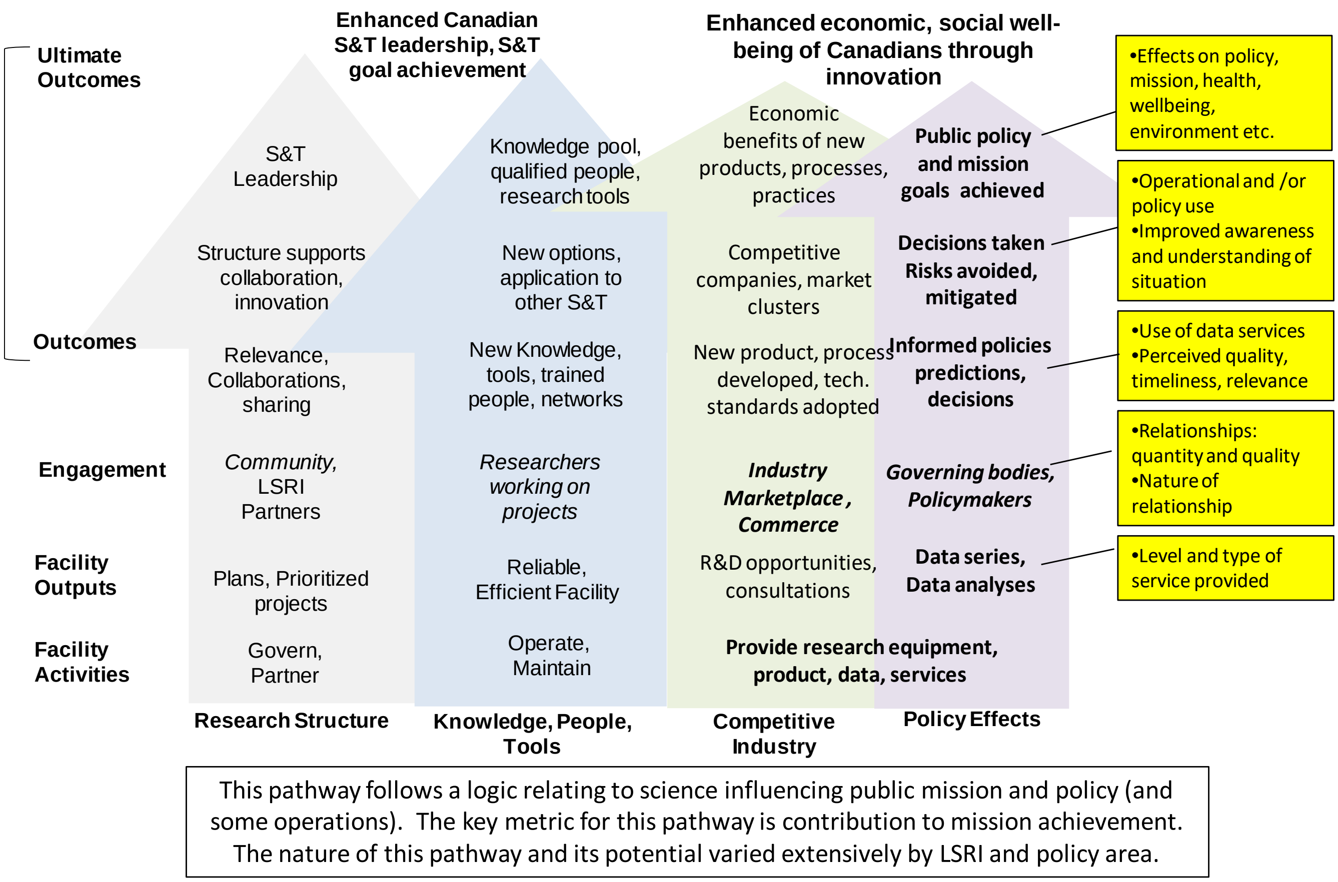
This pathway involves the support of government policies and decision making. Research can have impacts on policymakers' understanding of problems and possible responses to them. Research can also help with agenda setting in establishing priorities, making substantive changes in policies, protocols, regulations and investments and the way in which policies are delivered. In the case of LSRI, there are cases where the information generated can help with operational decision making as well as with policy.

[12] Key theories of change which inform this pathway include those included above for discovery and innovation – with special emphasis on private impacts (Mansfield *et al* 1977) and other economic impact theorists such as Lipsey and Beckar (1995) who discuss the effects of context and 'structure' on innovation. The work of Tassey (2008) and Link (1996) has noted the impact of technology infrastructure on firm competitiveness. Note that NRC has conducted significant past research on the impact of technology clusters. See Portfolio Evaluation of the NRC Technology Cluster Initiatives (2012).

(e.g. ocean monitoring helps with coast guard navigational decisions. It also helps with long-term climate change policy.)[13]

Figure 4 shows this pathway with some possible indicators.

Figure 4: Inform Government Policies and Decisions – Some Select Metrics



5.1.5 Pathway: Inspire Students and Public Appreciation of Science and Technology

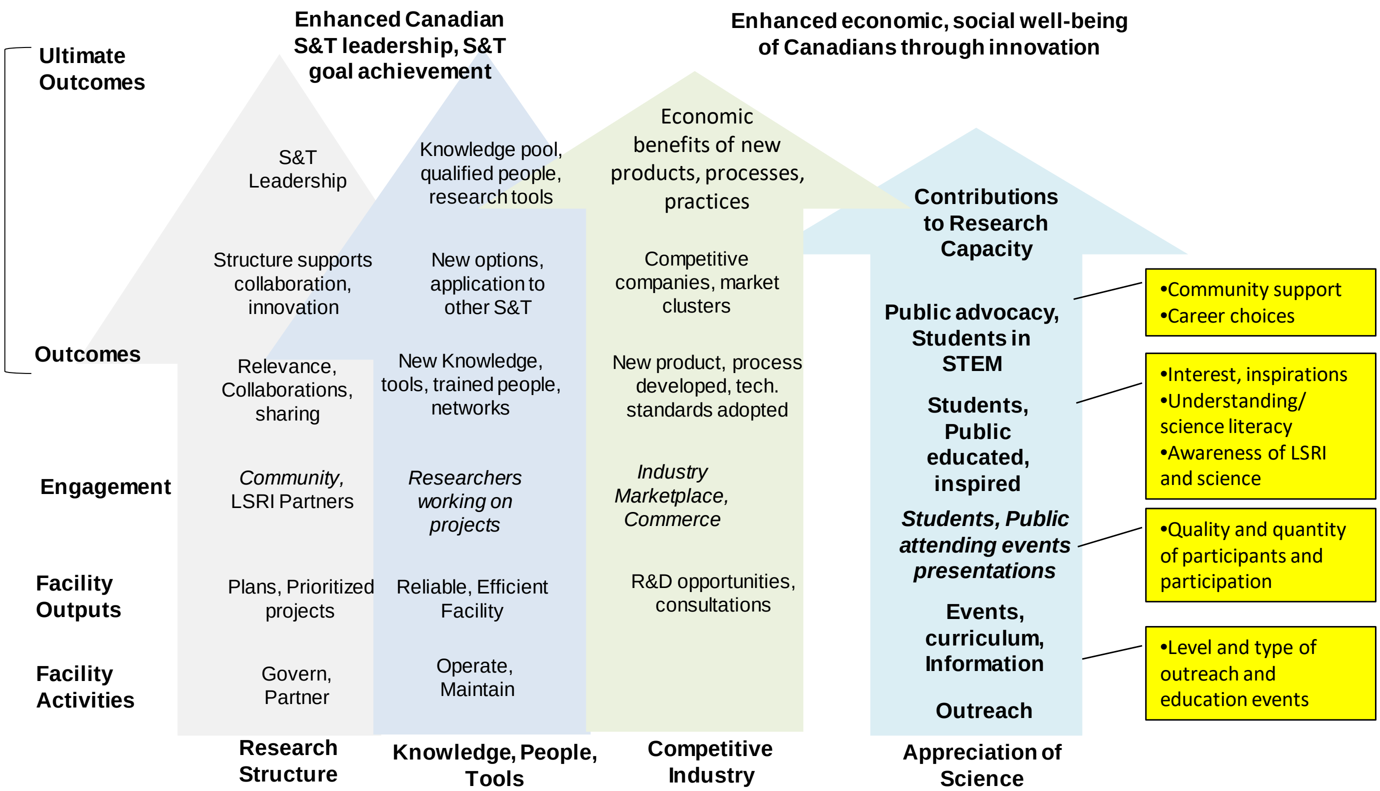
This pathway involves engaging, informing and inspiring the public, including students and the broad community, to create an awareness and appreciation of science. The combination of novel science and novel technology on an unusually large scale, or the unique physical location of facilities in some cases, stimulates the imagination and draws attention to scientific pursuits. Particularly when accompanied by related curriculum or research on site, this awareness may lead to the inspiration to pursue further

[13] The key theories informing this pathway relate to those for influence more generally (Kirkpatrick, Prochaska and others) as well as to the recent work of analysts looking at the influence of research and development on public policy (Sumner 2009, Steven 2007). In the U.S., considerable discourse on the effects of research on policy can be found in studies done by the Environmental Protection Agency to fulfill requirements of the Clean Air Act to evaluate and, if appropriate, revise existing criteria for pollutants every five years to reflect advances in scientific knowledge on the effects of the pollutant on public health and welfare and to recommend air quality standards (U.S. EPA 2009, Pahl *et al* 2008)

education in science, technology, engineering or mathematics (STEM), which can then lead to a related career. In any case, outreach is expected to garner public support for scientific endeavours.[14]

Figure 5 shows this pathway with some possible indicators.

Figure 5: Student and General Public Appreciation of Science – Some Select Metrics



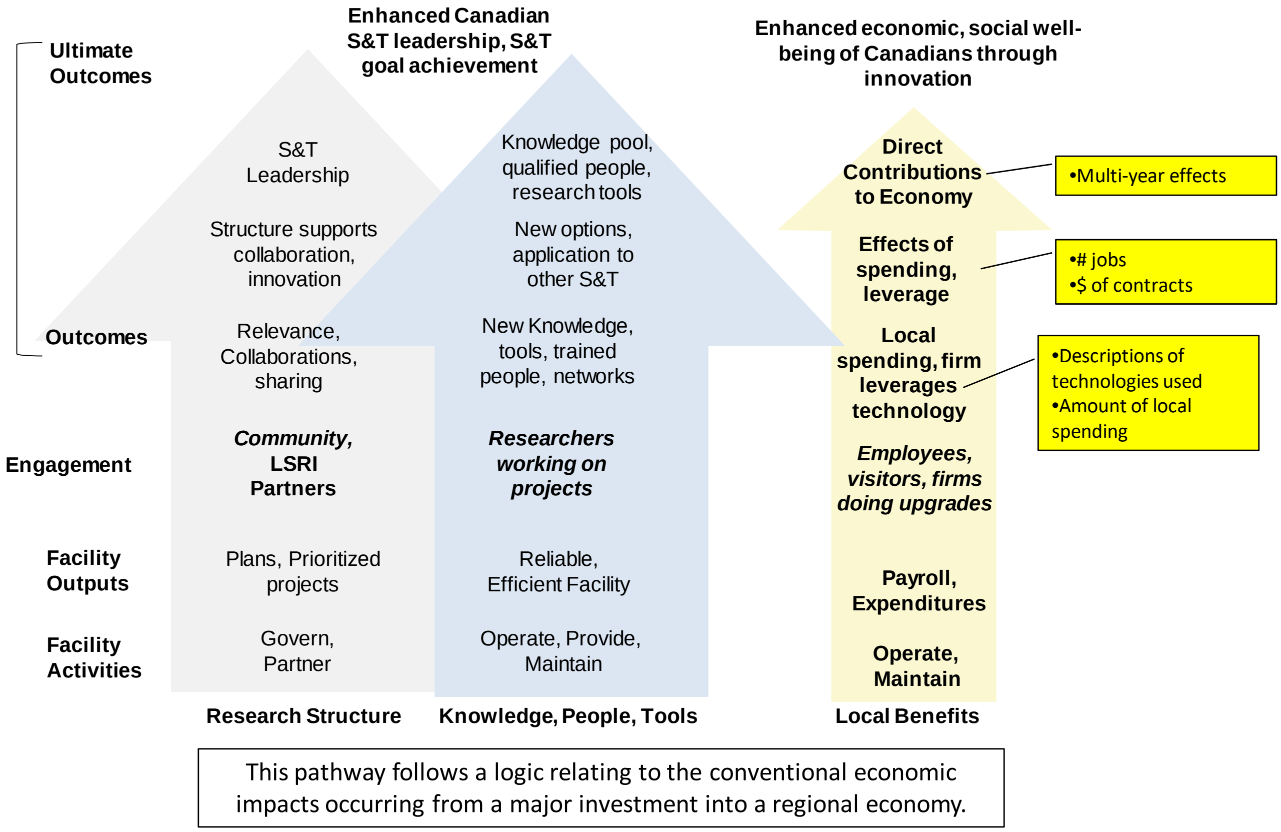
This pathway follows an education and social marketing logic. The key metrics involve tracking engagement, reactions and then knowledge and attitudinal changes. Difficulties attributing longer term results means that only limited effort should be taken to track this path in most cases.

5.1.6 Pathway: Contribute Directly to Local and High Tech Economic Activity

A unique feature of LSRI is that they contribute to the economy by being a large employer, investing in high tech equipment as well as other services such as construction and maintenance and various service supports, and attracting visitors who require lodging and other local services. The impacts from this kind of activity relate to economic stimulus effects.[15] This impact pathway is shown in **Figure 6.**

[14] Key theories of change which inform this pathway relate to learning (Kirkpatrick 1959) social (or even private) marketing theories (Prochaska *et al* 1992).
15 Typically local economic stimulus effects are measured by economic impact models such as those developed by Statistics Canada (see Assessment of the Sudbury Neutrino Observatory and SNOLAB 2012, KPMG). This type of modeling was used in at least one of the cost-benefit studies completed by the LSRI consulted.

Figure 6: Direct Economic Impacts of Facility Spending and Upgrades – Some Select Metrics



5.2 Best Practice in Performance Questions, Indicators and Methodologies

The pathways elaborated in section 5 have illustrative results statements and indicators. This section complements these statements with a set of key questions, potential indicator approaches and sources which may be considered in the monitoring, review and assessment of LSRI pathways.

The following tables represent the best practices found and suggested by the research team and suggest a possible starting menu of selections rather than a prescriptive form to complete.[16]

[16] Note the use of context assessment with the impact pathways and questions-indicators-methodologies chart will be shown in section 6.

5.2.1 Create a Research Structure That Supports Discovery and Innovation

Table 2: Performance Measurement and Evaluation Matrix for Create a Research Structure That Supports Discovery and Innovation

Topic	Questions	Indicators*	*Assessment Approach* Data Sources
Outputs	Plans (efficiency, effectiveness, prioritization)? Relevance? Robustness of delivery infrastructure Progress in required areas (e.g., safety) Deferred maintenance? Ratio capital/ operation? Equipment based service support provided?	Plans, budgets (actual vs. planned) Alignment of plans with community priorities; Partners recruited # staff employed $ of supplies purchased Functions/abilities of equipment, data Reliability of equipment, data Subscription rate	*Data analysis of planned vs. actual considering context* *Benchmarking with similar facilities* Management data Documents Self-assessment Interviews
Engagement	Partnerships? Community support? Advisory boards: number, type of advice sought? S&T community events sponsored, attended that inform plans, priorities, governance? Service users Grad students involved?	Board membership rosters, biographies Event name, purpose, attendance # international conferences hosted Exchange of staff	*Data analysis considering context* *Network analysis* Management data Documents Interviews Correspondence / communications
Short -term Outcomes (Outputs of users or Change in user behavior)	Fit with its S&T community priorities, national strategy and priorities, strengths? 'Value' of services provided Uniqueness of capabilities it provides? Collaborations formed? Synergies? Focal point/Hub? Sharing of resources?	# facility sponsors, amount of co-funding # of collaborations, national and international Research data generated Level of user / client satisfaction, Responsiveness to user needs How better science is enabled.	*Data analysis considering context* *Network analysis* Management data Document review Interviews Self assessment
Intermediate-term Outcomes	Reported success (e.g. in press releases)? Increased collaboration? Increased interdisciplinary, cross-sectoral or functional research teams? Focal point for clustering, critical mass development?	Trend in collaborations Team expertise Interdisciplinarity of publications # of researchers working in area supported by facility Money saved through economies of scale in infrastructure	*Case studies to analyze change over time (or anecdotes).* *Bibliometrics* Network analysis Document review Interviews Publications Self assessment
Contributing to Ultimate Outcomes	Stimulated discoveries? Accelerated innovations? Kept researchers on cutting edge, opening options for further discoveries? Science Leadership?	Media reporting of discoveries Demand for facility use Recognition, awards Ranking in field Leadership in international groups	*Contextualized case analyses; Expert Panel, possibly International Benchmarking Panel for Field* Data from above assessments

*All indicators will be depicted as level, %, # and / or will state actual vs. planed or vs. benchmark expectations.

The questions addressed through indicators and assessment for this pathway relate to identifying elements that are specific to large scale research infrastructure. The emphasis here is on the work of the facility, its quality and significance to the research it supports, national and international collaborations, and S&T leadership (pathway 1). Pathway 1 is inextricably linked to research capacity in terms of knowledge, HQP and research tools impacts (pathway 2).

5.2.2 Build Research Capacity – Knowledge Base, HQP, Research Tools

Table 3: Performance Measurement and Evaluation Matrix for Build Research Capacity - Knowledge Base, HQP, Research Tools

Topic	Questions	Indicators*	Assessment *Approach* Data Sources
Outputs	Areas of research questions that can be investigated? Supportive services provided? Relevance of these?	# projects complete Back log of experiments Reliability of equipment Response time to data requests Subscription rate	*Data analysis of planned vs. actual considering context* Management data Document review Self-assessment Interviews
Engagement	Make up of users (multiple groups?) Relationships developed with users, among users	Number of users by type # of national and international collaborations # international conferences hosted # high esteem engagements	*Data analysis* *Network analysis* Management data Document review User survey Communication tracking
Short -term Outcomes	New knowledge from research performed; New tools or techniques developing? Training users receive? Progress made, as reported in request for continued funding	Quality, quantity of peer-reviewed publications Journal impact factor # of papers with international co-authors # invitations to access international projects, be on committees # HQP trained in Canada	*Data analysis considering context* *Bibliometrics* Management data Document review User survey Interviews Publications Self assessment
Intermediate-term Outcomes	Have research findings, new tools, been used in further research? Answered important questions? Have new tools, techniques, software been produced for use by others? Have technical standards been developed and adopted? Are new areas of research being opened up? Have researchers collaborated on research?	Researchers received follow on funding (#, %, $) Awards Citations Speed of citations New scientific paradigm Used in applied research Use by other researchers IT Index Examples of use to solve problems beyond the field	*Data analysis including expert judgment or using expert pane.* *Bibliometrics* Document review Self assessment Interviews Bibliometric data Expert opinion Follow up with past student researchers
Contributing to Ultimate Outcomes	What is the contribution to HQP? Has this knowledge spilled over to other fields? To know-how in industry? S&T Excellence, Leadership?	# HQP attracted and retained #joint or adjunct positions in outside organizations Ranking in field Leadership in international groups	*Data analysis including expert judgment or using expert panel* *Specialized case studies* Bibliometrics Document review Interviews Publications

*All indicators will be depicted as level, %, # and / or will state actual vs. planed or vs. benchmark expectations.

The questions addressed through the indicators and assessment for this pathway are identical to those normally used to assess research more generally. These measure the increase in knowledge, HQP and research tools (research capacity) which can be largely tracked at outcome levels through publications, citations and other quantitative metrics as well as by case studies and area analysis of a sequence or 'chain' of results contributed to by the LSRI.

5.2.3 Contribute to New Technologies, Competitive Companies, Markets and Clusters

Table 4: Performance Measurement and Evaluation Matrix for Contribute to New Technologies, Competitive Companies, Markets and Clusters

Topic	Questions	Indicators*	Assessment *Approach* Data Sources
Outputs	Relevance, value added for industry? Consultations by facility staff Characteristics of R&D available to industry? Data, Product provided directly to industry	# consultations, technical services provided Amount, quality of data or product available #demonstrations hosted	*Data analysis of planned vs. actual considering context* Document review Self-assessment Interviews
Engagement	What has been the relationship with industrial companies: duration, alignment of interests? What types of sectors and companies? Separate advisory group?	# joint projects # and types of organizations Cost share; in kind contributions Amount of product purchased, facility used Repeat use of LSRI	*Data analysis considering context* *Network analysis* Document review Interviews
Short -term Outcomes	New knowledge gained from R&D projects? What was product or data used for, for whom?	# HQP exchanged with industry; going to industry Technical milestones achieved Movement through development stages R&D tests run Patents applied for	*Data analysis considering context* *Patent analysis* Document review User survey Interviews with contractors, former users Data requests to firms using products or data Trade journals
Intermediate-term Outcomes	Have companies formed/ moved to provide support services? Have new products been commercialized? Have companies a competitive advantage because of new skills, new product, reputation or new markets?	# of technologies developed by user or transferred to industry # technologies developed for facility adapted for other applications; examples of innovations # Companies spun off Examples of access to world markets	*Success stories/ contextualized case analyses* Document review Interviews Follow up with former participants Patent analysis
Contributing to Ultimate Outcomes	Local spending and jobs during upgrade; When new products or services are used, what are the benefits of use (cost savings, health improves, etc.)	# jobs created Construction contracts ($) Visitor spending estimate Sales of commercialized products; effects of use	*Success stories/ contextualized case analyses* *Evaluation study* (could be cost-benefit) Document review Secondary statistics Economic analysis

*All indicators will be depicted as level, %, # and / or will state actual vs. planed or vs. benchmark expectations.

The questions addressed through the indicators for this pathway relate to the level and nature of engagement with industry to maintain and support the LSRI itself or to technologies developed for private sector use. In some cases, these questions may reference actual services and data used. Client tracking in terms of use of the facility and milestones related to technology development and applications may be of high importance. Specialized case and cost-benefit analyses would seem directly applicable in many instances.

5.2.4 Inform Government Policies and Decisions

Table 5: Performance Measurement and Evaluation Matrix for Inform Government Policies and Decisions

Topic	Questions	Indicators*	Assessment *Approach* Data Sources
Outputs	Relevance of facility for governments, policy makers? What was provided? How much was the demand for it?	Data series available Data analyses, forecasting performed, available Validations completed Data requests received Consultations, visits	*Data analysis of planned vs. actual considering context* Document review Interviews
Engagement	Who are partners, users and other stakeholders? Trans-sectoral projects? Work with indigenous communities? How are they engaged?	Meetings with government officials, staff, stakeholders Collaborative agreements	*Data analysis considering context* *Network analysis* Document review Interviews
Short -term Outcomes	What has been done with the data or service provided? How has it changed attitudes and behavior?	Perceived quality of data New demonstrated awareness, understanding Questions informed by data, analysis provided Policy discussions informed	*Data analysis considering context* Document review User survey Interviews with agencies involved
Intermediate-term Outcomes	How and where has the knowledge been applied?	Application or change in policy, regulations, changes in behavior, in decisions or policies Risks avoided or mitigated Disputes settled Information provided for international agreements	*Success stories/ contextualized case analyses* Document review Interviews Data mining policy documents Analysis prepared for regulatory hearings, decision discussions Media reports
Contributing to Ultimate Outcomes	Benefits from these changes in behavior, decisions or policies (reduced storm damage, lower costs, etc.)	Effects of policy change or decision (cost savings, emissions reduced, tanker safety, etc.) Effects of that on health, etc.	*Success stories/ contextualized case analyses* *Evaluation study (could be cost-benefit)* Document review Secondary statistics Economic analysis

*All indicators will be depicted as level, %, # and / or will state actual vs. planed or vs. benchmark expectations.

The questions addressed through the indicators and assessment for the government policies and decision support impact pathway relate to the engagement of partners, users and other stakeholders leading to direct outcomes resulting in improved awareness, and understanding as well as the use of research results for decision-making.[17] The nature of this type of impact pathway lends itself to case study approaches.

[17] As demonstrated through use tracking and feedback (formal through surveys or informal via correspondence and / or the content of meeting minutes.

5.2.5 Inspire Students and General Public Appreciation of Science and Technology

Table 6: Performance Measurement and Evaluation Matrix for Inspire Students and General Public Appreciation of Science and Technology

Topic	Questions	Indicators*	Assessment *Approach* Data Sources
Outputs	Tours, talks, events? Educational products produced? Local input sought?	Number, attendance/type at tours, talks Web pages, social media posts # of programs, events administered Educational products produced Media mentions, citations	*Data analysis of planned vs. actual considering context* Document review Self-assessment Interviews
Engagement	What has been the type and level of engagement with students, teachers, parents, local community leaders, general public	Characteristics of participants Partnerships, formal and informal; cost share Visits by media, political leaders Initiatives to solve local problems	*Data analysis considering context* *Network analysis* Document review Interviews
Short -term Outcomes	What learning and research opportunities have been provided, to whom? Have students, public continued to be engaged? Media mentions, media reach?	Trends in attendance at scientific lectures; in use of curriculum Media citations; web downloads, social media shares # students doing research projects Awareness of facility Local community leaders participating	*Data analysis considering context* Document review Participant survey Follow up participant survey Interviews Self assessment
Intermediate-term Outcomes	Are students taking STEM courses, degrees, pursuing STEM careers? Is there more advocacy by the public for science?	# students in STEM classes, grade level Parental support for STEM # students graduating in field(s) of facility Examples of public advocacy for facility, for science generally	*Case study to analyze change over time.* Document review Interviews Secondary statistics
Contributing to Ultimate Outcomes	Contributes to HQP; Strengthens a culture of science and S&T excellence.	Career paths of former users, students Percent of public in favor of public $ for science Science literacy	*Success stories/ anecdotes* *Evaluation study* Secondary statistics Survey

*All indicators will be depicted as level, %, # and / or will state actual vs. planed or vs. benchmark expectations.

The questions addressed through the indicators for the students and public inspiration pathway relate to the level, extent and continuity of engagement. It may also include (somewhat anecdotally) more extensive tracking of inspired students.

5.2.6 Contribute Directly to Local and High Tech Economic Activity

Table 7: Performance Measurement and Evaluation Matrix for Contribute Directly to Local and High Tech Economic Activity

Topic	Questions	Indicators	Assessment Approach Data Sources
Direct Economic Effects (apply to all LSRI, relatively easy to quantify, is fairly credible)			
Direct Benefits from construction, operations	Solicitations for facility upgrades? Level of technical challenge in upgrade? Total spending by LSRI and visitors in local economy? Effects on high tech firms? Effects on economy?	# jobs, $ of contracts Spending of visitors Multiplier effects Effects on firm's capabilities, products, competitiveness, exports	*Data analysis, Economic analysis* Administrative data Secondary statistics Interviews with firms Trade journals
Estimated, Quantified Economic (and sometimes social) Effects (difficult, requires many assumptions and modeling that can be questioned)			
Estimated Direct and/or Indirect Benefits	Individual cases What are the new technology sales, effects of its use that can be traced back to LSRI action? Has a research result had effects on health, energy, national security? Can these effects be quantified? What are the economic cluster effects, that is, has the LSRI drawn support companies and other organization to its vicinity and, if so, what have been the effects? What government policies or technical standards have been implemented that can be traced back to LSRI action? What are the effects of implementation on markets, the economy?	Quantified benefits (cost savings, energy savings, accidents prevented, etc.) Benefit-cost ratio, ROI Direct GDP impact ($) Total GDP impact Provincial impact These benefits are generally accompanied by qualitative evidence that helps explain how the benefits occurred.	*Cost benefits analysis* (Generally a mixed method case study) Historical tracing Econometric modeling (e.g. Statistics Canada economic impact models) Interviews Expert judgment

The indicators and assessment appropriate for these questions are best addressed by drawing on case areas showing impact in key pathways and then applying a cost-benefit approach to areas of the analysis which can be monetized. Note that for this reason, cost-benefit analysis may be more appropriate for some LSRI (i.e. those with easily monetized benefits) than for others.

6.0 Conclusions and Recommendations

Sections 3, 4 and 5 established that LSRI performance is strongly affected by context and that there are a number of valuable impacts possible in different permutations and combinations. These findings suggest that LSRI operate as part of complex, dynamic systems. Such systems are extremely difficult to benchmark and assess. The approach to developing performance questions, indicators and methods of assessment must be duly considered in this context. For these reasons simple quantitative 'summary' indicators, reported without contextualized description are probably the least valid approach to take for the assessment of LSRI outcomes and impact.

This report concluded that there are three major limitations in current practice and if these are filled, wider impacts can be assessed earlier in the process and these can increase understanding of the contributions of LSRI as well as inform improvements.

1. The wide variance in the nature and context of LSRI conditions for assessment of performance is not described. The variety of purposes, target audiences, activities and how these activities will lead to desired results given specific existing conditions are seldom explicit. Logic models with underlying causal assumptions have not been done.

2. Related to the first, while LSRI typically consider a number of important quantitative indicators for management, not all of the pathways by which outcomes occur have been clearly explained and pursued. Current practice rarely considers public policy and effects of government-funded technical infrastructure. Also detailed descriptions of common sequences of outcomes currently used in assessment of more ordinary S&T organizations are not used for LSRI. Examples are stages of building trusted relationships with partners or stages of technology development and commercialization.

3. Monitoring and evaluation approaches were found to be somewhat limited and inconsistent when it came to outcomes (see section 4.4). Opportunities exist to monitor and track more impact pathways, document sequences of outcomes, and improve on both success stories and periodic assessment with careful case studies building on monitoring data and assessment of progress along impact pathways. In particular, the relationships among the metrics should be assessed to paint a broader picture of impacts.

6.1 Recommendations

Given the findings and conclusions of this study, three fundamental activities are recommended as means to improve LSRI impact assessment as follows:

1. At the beginning of any assessment period each LSRI should systematically define specific aspects of the logic of that LSRI and the conditions underlying its performance. The essence of this is that assessment should answer questions about program/organization logic: Why is the program important to science and society? What sequence of results occur from which target audiences? and How did the program influence that through its activities, taking into account differing conditions? Many important characteristics are described in Section 2.1 (Table 1) and in Template 1 in this section.

2. LSRI need to define specific impact pathways that apply in their case. This report suggests six that apply to LSRI. The LSRI should, for those that apply to them, describe the logic of each in some detail (this could include a narrative as well as a graphic). The categories in the generic models and indicator tables shown in section 5 can serve as a guide.

3. In order to appropriately assess performance it is recommended that each LSRI implement a multi-year assessment plan that has three levels of analysis integrated over a period of time from four to ten years depending on funding cycles and LSRI context (including the LSRI lifecycle).

Section 6.2 explains the proposed assessment process in more detail.

6.2 Assess LSRI with Three Levels and Time Frames for Analysis

In order to appropriately assess performance it is recommended that each LSRI implement a multi-year assessment plan that has three levels of analysis integrated over a period of time from four to ten years depending on funding cycles and LSRI context. The foundation of this plan is development of a logic model and impact pathways specific to that LSRI. This model shows the important performance questions to ask and data to collect and analyze. The specific logic model and assessment plan can adapt the generic assessment framework proposed here to the specifics of the particular LSRI.

The collection of assessment questions and associated indicators at a lower level are linked to and support the questions at the higher level. For example, the quality of facility operations affects the level of user satisfaction and the significance of the results of supported research. Careful tracking of co-investment and major events over time helps to explain the contribution of the LSRI to an outcome related to that activity. This integrated approach lowers the cost and improves the effectiveness of impact assessment.

The three levels of analysis are:

1. Routinely collect data on inputs, activities, outputs, and engagement (users, partners, others involved), This includes looking at the quality efficiency and productivity of facility operations, relevance of research options provided, people trained, reach, and research projects enabled. It also includes documenting major external influences (e.g., funding for the whole field, breakthroughs elsewhere). The information collected depends on the LSRI logic model and context.

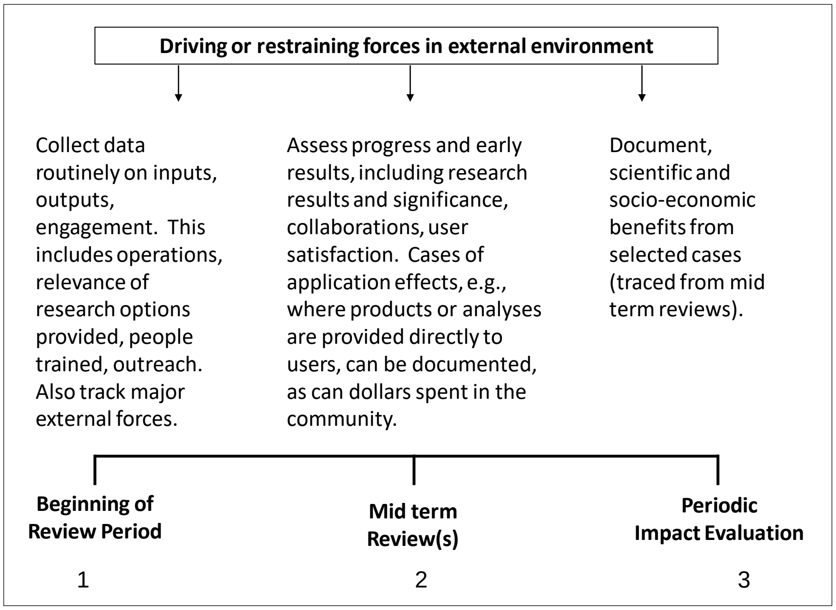
2. As required, mid-term reviews (at most every other year) will assess progress and early effects. This includes looking at the continued relevance of research options provided, research results as measured by publications, collaborations, significance of that research, various means of 'take-up' and user satisfaction. Where products or analyses are provided directly to users, the early effects of use of those can be documented, as can be other cases of application of LSRI outputs including potential socio-economic benefits[18]. The relatively simple calculation of dollars into the local community from payroll, visitors and operations spending can be calculated.

3. Periodically, larger assessment efforts can determine longer-term S&T leadership, policy and socio-economic impacts, both qualitative and quantitative, by using, updating, and expanding on the data and analysis completed in levels 1 and 2. Scientific impact and continued relevance can be investigated with bibliometric analysis and expert panel review. Success stories and anecdotes can be expanded and verified in case studies. Ideally, the contribution of the LSRI to the documented outcomes can be demonstrated by showing movement along the program logic and impact pathways, accounting for plausible alternative explanations external to the LSRI activities and results.

The following figure summarizes the three level strategy.[19].

[18] For example direct economic impacts may occur from the immediate take-up of information or products such as research productivity gains, reduced information search times, ship navigation support benefits, isotope use, technical dispute resolution, or standards implementation, and other immediate 'problem solving' impacts.
[19] Note: This contextual information will change over time (e.g., as LSRI need new operating funding, they will seek out new funding sources, which often come with new stakeholders and activities. These may shift the broader focus of the LSRI over time.)

Figure 7: Assess LSRI with Three Levels



The following sections provide guidance for developing a multi-year impact assessment plan using the generic impact pathways. These can be adapted and made specific to the LSRI and its context. First there is a template for developing an impact pathway. Then there are six tables of generic assessment questions, indicators, and assessment approach and data sources for each performance category. These are sorted by the six impact pathways and include all three levels of analysis.

6.3 A Template for Describing the Program and Its Context

The following are questions that LSRI staff can answer to develop a description of the program logic or impact pathways. These follow our earlier description of the LSRI context in terms of Why? What? Who? and How? (See **Table 1.**) Answering these questions is an iterative process. They are listed below.

Questions to Help Define Logic and Existing Conditions
The "Why" (Ultimate Outcomes) 1. In what area(s) will this LSRI enhance Canadian S&T leadership (fields, enabling services)? What broad strategic needs are served? How does it fit into national S&T policy? Why is it significant? 2. What role does this LSRI play in the international arena? How does it benefit S&T in this era of globalization? 3. Can it bring together and focus S&T communities? 4. What are likely and/or possible socio-economic impacts from use of the outputs of the LSRI and its users in further S&T or new commercial products or processes or policies or practices? **The "What" (Early Outcomes and Outputs) Consider all of the impact pathways that might apply.** 5. What user needs are satisfied by this facility? 6. What specific results can be expected in five years in the arenas the LSRI operates in? For example, these may be big science questions (define these) or options for solutions to an identified problem, or development of a new technology? 7. What are the possible application areas for those results, scientific and otherwise? 8. What, if any, results are expected from education and outreach activities?

Questions to Help Define Logic and Existing Conditions
The "Who" (Engagement of Partners and Users) 9. What S&T communities support and/or use this LSRI, national and international? 10. How multi or interdisciplinary or trans-sectoral is the work supported? 11. How much development and innovation is currently occurring in equipment, techniques, data, technical standards, and/or services of the facility? Might this spill over to firms or other applications? Where? 12. To what extent and how are industry and government policy makers directly involved, if at all?
The "How" (Activities and Inputs) 13. What are the unique capabilities of the LSRI that enable research? What kinds of research? 14. Does LSRI staff perform research and development as well as host these activities? 15. What activities does the LSRI perform in addition to research and research support? These could include providing products, data sets, analyses, protocols, public outreach. 16. What else, if anything, distinguishes this LSRI: Capital intensity, economics of scale or scope?
External Influences (Driving and Restraining Forces for Success) 17. What big picture changes might influence your success, such as global S&T spending, serendipitous discovery, or economic or societal conditions? 18. What circumstances - anticipated or not - might slow or hasten your progress toward outcomes, such as pace of evolution in S&T, S&T progress made elsewhere, technology readiness[20] of those who would apply S&T to problem areas? 19. What external events could influence collaborators or user groups, such as global trends, capacity, or political influences in collaboration or expenditures on S&T? 20. What external influences, that you have not already accounted for in LSRI design, might affect how the LSRI operates (e.g. national strengths or strategy, scientific or technical change, or governing authorities and management norms)?

The questions noted above should be used to establish impact pathways. The preliminary models for these pathways are shown in section 5 of this report.

6.4 Template for Describing Impact Pathways

The following template distils the pathways shown in section 5 into a simplified table to describe the impact pathways. Note that contextual influences (see **Table 1**) will still need to be considered at various points along each pathway.

Description of Impact Pathways (Not All Will Apply)						
Area	**Pathway**					
	Research Structure	**Knowledge, People, Tools**	**Competitive Industry**	**Policy Effects**	**Appreciation of Science**	**Local Benefits**
Emphasis on Pathway (High - Low)						
Activities						
Outputs						

[20] See technology readiness definition at https://buyandsell.gc.ca/initiatives-and-programs/build-in-canada-innovation-program-bcip/program-specifics/technology-readiness-levels

Description of Impact Pathways (Not All Will Apply)						
Area	**Pathway**					
	Research Structure	**Knowledge, People, Tools**	**Competitive Industry**	**Policy Effects**	**Appreciation of Science**	**Local Benefits**
Who is engaged						
Other major influences						
Early Outcomes (1-2 years)						
Intermediate Outcomes (3-5 years)						
Longer-term Outcomes (6-10 years)						
Ultimate Outcomes (10+ years)						

6.5 Templates for Levels of Assessment / Performance Measurement

Once LSRI have developed a logic model with impact pathways, they are ready to develop tables of the performance questions that the LSRI and its stakeholders need to answer. Then for each of these, LSRI management must define indicators in order to collect data to answer these questions, as well as the approaches to analyze this data, and likely data sources.

A template for these tables is provided below. The generic tables can be used to create impact pathways in monitoring and evaluation (M&E) strategies. Note that distinct questions, indicators, approaches and sources should be considered for each important impact pathway. The examples in section 5 can serve as a starting guide, but it is important to tailor the M&E strategies to the specific context of each LSRI.

Data Collection and Analysis – LSRI Monitoring, Operational Management and Impact Assessment				
Topic	**Questions**	**Indicators**	**Assessment Approach**	**Data Sources**
Level 1: ***Complete the following for each relevant impact pathway***				
Activities / Outputs				
Engagement				
Short –term Outcomes				
Key External Influences (see Table **1**, Section **3.2**)				
Level 2: ***Build on the data and analysis developed for each impact pathway completed in level 1***				
Intermediate-term Outcomes				
Key External Influences (see Table **1**, Section **3.2**)				
Level 3: ***For each impact pathway that applies – tailor the following – building on analysis completed for levels 1 and 2***				

Data Collection and Analysis – LSRI Monitoring, Operational Management and Impact Assessment				
Topic	**Questions**	**Indicators**	**Assessment Approach**	**Data Sources**
Contribution to S&T Leadership and National Goals				
Contribution to Research Capacity				
Contribution to Socio-economic impacts through Industry				
Contribution to Socio-economic impacts through Policy, Government				
Contribution through Outreach to Students and Public				
Estimated Direct and/or Indirect Benefits -Facility Expenditures, include upgrades -Take up by Industry or Governments				
Key External Influences (see Table **1**, Section **3.2**)				

Note that the three templates offered here may be considered starting points, subject to refinement, trial and adjustment over time. As such, these proposed frameworks are not meant to be prescriptive and comprehensive but rather indicative and inspirational.

6.6 Proposed Approach to Implementation

In summary, a three step process for LSRI monitoring and evaluation is proposed. The three steps are:

i. Consider key characteristics of LSRI relating to why they exist, what outcomes they produce, who they work with and how they operate.

ii. Outline impact pathways so that performance can be assessed by referring to a sequence of results and appropriate indicators linked to those pathways.

iii. Establish monitoring and evaluation strategies which operate at the level of on-going operations, a mid-term review related to early and intermediate results and periodic impact evaluations.

Canadian funding authorities should consider a pilot trial or trials for this approach for upcoming LSRI assessments. Through such pilots and early trials, the detailed elements of context, impact pathways, indicators and approaches can be refined, improved and shared over time to allow for generative learning in this emerging area.

Appendix A: List of References for Literature Review

KPMG. Assessment of the Sudbury Neutrino Observatory and SNOLAB Final Report February 27, 2012.
Australian Government, Department of Innovation, Industry, Science and Research, Strategic Roadmap for Australian Research Infrastructure, August 2008.
Autio, Erkko, Imperial College London. UK Department for Business Innovation and Skills, *Innovations from Big Science: Enhancing Big Science Impact Agenda.* March 2014.
Barker, K., Sveinsdottir, T., & Cox, D. 2013. The "Innovation Turn" in Policy for Large Scientific Facilities: Reflections on introducing innovation support dimensions to the operation of scientific research infrastructure. Paper presented at the 2013 EU-SPRI Forum Conference, Madrid.
Blamey, A. and M. Mackenzie (2007). *Theories of Change and Realistic Evaluation*. Evaluation 13(4): 439-455.
Boisot, M., Nordberg, M., Yami, S., & Nicquevert, B. (Eds.). 2011. *Collisions and collaboration: The organization of learning in the ATLAS experiment at the LHC*. Oxford: Oxford University Press.
Bressan, B., Kurki-Suonio, K., Lavonen, J., Nordberg, M., Saarikko, H., & Streit-Bianchi, M. 2009. Knowledge creation and management in the five LHC experiments at CERN: Implications for technology innovation and transfer. Geneva: CERN-Atlas.
Bronfenbrenner, U. (1979) The Ecology of Human Development: Experiments by Nature and Design. Cambridge, MA: Harvard University Press.
Canada Foundation for Innovation (CFI) Outcome Measurement Study: Instructions for the Institution
Canada Foundation for Innovation (CFI), Major Science Initiatives Oversight Framework, 2012.
Canada France Hawaii Telescope. Annual Report 2011
Canada France Hawaii Telescope. Annual Report 2012
Canada France Hawaii Telescope. Annual Report 2013
Canadian Light Source. Making Positive Economic, Scientific Impacts for Canada. (Background)
Canadian Light Source. Building on Success, 2011-2012 Annual Report
Canadian Light Source. Economic, Societal and technological impact of large scale research facilities: a view from Canada Prepared for European Research Facilities Workshop at DESY Hamburg May 31, 2012.
Committee on Science, Engineering, and Public Policy (COSEPUP), National Academy of Sciences, National Academy of Engineering, Institute of Medicine Experiments in International Benchmarking of U.S. Research Fields, National Academies Press, Feb 28, 2000
Compute Canada. Strategic Plan 2010-2020.
Compute Canada. Strategic Plan 2014-2019.
COST. 2010. Benefits of Research Infrastructures beyond Science: The Example of the Square Kilometre Array (SKA) In COST (Ed.). Brussels: COST.
Council of Canadian Academies, Expert Panel on the Socio-economic Impacts of Innovation Investments, "Innovation Impacts: Measurement and Assessment", Ottawa, Ontario, 2013
Council of Canadian Academies, Informing Research Choices: Indicators and Judgment, The Expert Panel on Science Performance and Research Funding, 2012

Crabtree, D.R., Bryson, E.P. 2001. *The Effectiveness of the Canada-France-Hawaii Telescope*. Journal of the Royal Astronomical Society of Canada 95, 259. (updated in 2014) http://cds.cern.ch/record/515349/files/0108430.pdf
Douthwaite, B., Alvarez, B.S., Cook, S., Davies, R., George, P., Howell, J., Mackay, R. and Rubiano, J. 2007a. *Participatory Impact Pathways Analysis: a practical application of program theory in research-for-development*, Canadian Journal of Program Evaluation, 22(2): 127-159. http://www.evaluationcanada.ca/secure/22-2-127.pdf
European Commission, Cost control and management issues of global research infrastructures, October 2010.
European Strategy Forum on Research Infrastructures, Strategy Report on Research Infrastructures, Roadmap 2010. (Luxembourg 2011).
Evaluation of NRC's Contribution to TRIUMF http://www.nrc-cnrc.gc.ca/eng/about/planning_reporting/evaluation/2013_2014/triumf.html
Evaluation of the NRC Technology Cluster Initiatives (2012). http://www.nrc-cnrc.gc.ca/eng/about/planning_reporting/evaluation/2009_2010/technology_cluster_initiatives.html
Finland's Strategy and Roadmap for Research Infrastructures 2014–2020, December 2013
Finnish Ministry of Education, National-Level Research Infrastructures: Present State and Roadmap, Helsinki 2009.
Hickling Arthurs Low, *Astronomy in Canada*, 2011
Hickling Arthurs Low. "Return on Investment in Large Scale Research Infrastructure", Ottawa, Ontario, May 2013 (TRIUMF is case study)
Insightrix Research Inc., "Economic & Social Impact Final Report", prepared for The Canadian Light Source Inc., Saskatoon, Saskatchewan, August 2011
Integrated Marine Observing System Strategy 2015-25, May 2014
Integrated Marine Observing System, 2013-14 Annual Progress Report and Final Report to 30 June 2014 for National Collaborative Research Infrastructure Strategy
Integrated Marine Observing System, Highlights 2012-2013
International Space Exploration Coordination Group (ISECG), "White Paper: Benefits Stemming from Space Exploration", IOWG Draft, August 2012
Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. American Economic Review, 76 (5), 984–1001.
Jordan, G. B., Hage, J., & Mote, J. 2008. A theories-based systemic framework for evaluating diverse portfolios of scientific work, part 1: Micro and meso indicators. In C.L.S. Coryn & Michael Scriven (Eds.), *Reforming the evaluation of research. New Directions for Evaluation, 118,* 7–24.
Kelly, Kofner, Norling *et al*, A Review of Reports on Selected Large Federal Science Facilities: Management and Life-cycle Issues, prepared for the Office of Science and technology Policy, Rand Science and Technology Policy Institute 2003. See http://www.rand.org/pubs/monograph_reports/MR1728.html#toc
Kirkpatrick, D. L. (1959). *Techniques for evaluating training programs*. Journal of ASTD, 11, 1–13.
Kline, S. J. and N. Rosenberg (1986). An overview of innovation. The Positive Sum Strategy: Harnessing Technology for Economic Growth. R. Landau and N. R. (Eds). Washington, DC, National Academy Press: 275-305.

Learning about Theories of Change for the Monitoring and Evaluation of Research Uptake, IDS Practice Paper In Brief 14, September 2013 http://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/2995/PP%20InBrief%2014%20FINAL.pdf;jsessionid=549510C35842BD8473D4E17344845D4A?sequence=1
Leeuw, F. (2012). *Linking theory-based evaluation and contribution analysis: Three problems and a few solutions*. Evaluation 18(3): 348-363.
Link, A. N. (1996) Evaluating Public Sector Research and Development Praeger Publishers
Lipsey, R. G. and C. Beckar, (1995). A Structuralist View of Technical Change and Economic Growth. In T. J. Courchene (ed.) Technology, Information and Public Policy, 9-75. The Bell Canada Papers on Economic and Public Policy, vol. 3. Kingston: John Deutch Institute for the study of Economic Policy.
Makarov, Marja. Finland's Roadmap for Research Infrastructures 2014–2020, Academy of Finland, March 2014.
Mansfield, E., John, R., Anthony, R., Samuel, W., George, B., 1977. *Social and private rates of return from industrial innovations*. The Quarterly Journal of Economics 91 (2), 221–240
Mayne, J. and E. Stern (2013). *Impact evaluation of natural resource research programs: a broader view*, ACIAR Impact Assessment Series Report No. 84. Canberra: Australian Centre for International Agriculture Research (ACIAR). Available athttp://aciar.gov.au/files/ias84.pdf.
Montague, S. and Teather, G. (1997) *Performance Measurement, Management, and Reporting for S&T Organizations – An Overview*, Journal of Technology Transfer, Vol. 22, Summer 1997, pp 5-12
Moltmann, Tim *Performance measurement for an ocean observing system – perspectives from Australia's IMOS* (University of Tasmania), November 2014
Mote, J., G. Jordan, and J. Hage. 2007. New Directions in the Use of Network Analysis in R&D Evaluation. *Research Evaluation* 16(3): 191-203
National Academies, Setting Priorities for Large Research Facility Projects supported by the National Science Foundation, the National Academies Press, 2004.
National Science Foundation, Large Facilities Manual, 2011.
NRC Evaluation of the International Telescope Agreements Program
Ocean Networks Canada Annual Report 2013
Ocean Networks Canada Strategic Plan 2013-2018
OECD Global Science Forum, Large Research Infrastructures, Report on Roadmapping of Large Research Infrastructures (2008); Report on Establishing Large International Research Infrastructures: Issues and Options (2010)
OECD Global Science Forum, Report on The Impacts of Large Research Infrastructures on Economic Innovation and on Society: Case studies at CERN, 2 October 2013 DRAFT.
Pahl, Wilson, Evans, Kowalski, Vickery, and Costa. Research, policy, and evaluation: systematic interaction informs air quality decisions. *Research Evaluation* 17(4), December 2008, pages 251-263.
Pentland, A. (2014) Social Physics: How Good Ideas Spread—The Lessons from a New Science. Penguin Press USA
Prochaska, J.; DiClemente, C., and Norcross, J. (1992) In search of how people change: applications to addictive behaviors. American Psychologist, 47(9), 1102-1114
Research Councils UK, Large Facilities Roadmap, November 2005.

Rogers, E. (2003, 1995) Diffusion of Innovations New York: Free Press
Rogers, Juan D., Jan Youtie, and Luciano Kay. "Program-level assessment of research centers: Contribution of Nanoscale Science and Engineering Centers to US Nanotechnology National Initiative goals." *Research Evaluation* (2012): rvs028.
Rogers, P. (2008). *Using programme theory to evaluate complicated and complex aspects of interventions*. Evaluation 14(1): 29-48.
Ruegg, R. (2000) Delivering Public Benefits with Private-Sector Efficiency. In Advanced Technology Program: Assessing Outcomes, edited by C. W. Wessner. Washington, DC: National Academy Press.
Science and Technology Facilities Council *The Socio-Economic Relevance of Research Infrastructures* Views from the UK John Womersley, 31 May 2012
Science and Technology Facilities Council, Terms of Reference for Large Facilities Subgroup. See https://www.stfc.ac.uk/resources/PDF/LFSGTORx.pdf
Science and Technology Facilities Council. STFC Impact Report 2011. In STFC (Ed.). Swindon: STFC. Seer http://www.stfc.ac.uk/resources/pdf/stfcimpactreport2011.pdf
Simon Fraser University, Thinking of the World, Centre for Policy Research on Science and Technology (CPROST), Research team of J.A.D. Holbrook, B.Wixted, F.Chee, M.Klingbeil, G.Shaw-Garlock, "Measuring the Return on Investment in Research in Universities: The Value of the Human Capital Produced by these Programs", Vancouver, British Columbia, 2008
SNOLAB Strategic Plan 2012-2017
Society for Science-Science for Society, Vision 2030, "The ESS in Lund-its effects on regional development", European Spallation Source (ESS), 2009
Steven, D. (2007) 'Evaluation and the New Public Diplomacy', presentation to the Future of Public Diplomacy, 842nd Wilton Park Conference, River Path Associates
Strategic Framework for Research Infrastructure Investment, Discussion Paper, December 2010, Government of Australia
Sumner, A.; Ishmael-Perkins, N. and Lindstrom, J. (2009) Making Science of Influencing: Assessing the Impact of Development Research, IDS Working Paper 335, Brighton: IDS
Tassey, G. (2007) *The technology imperative*. Edward Elgar Publishing
Tassey, G. (2008) The Roles and Economic Impacts of Technology Infrastructure http://www.nist.gov/director/planning/upload/Measurement_Infrastr_Roles_Impacts_v3.pdf
Technopolis Group, "The role and added value of large-scale research facilities", final report, Brighton, UK, 10 February, 2011
Simmons *et al* (2013), (Technopolis Group) Big Science and Innovation, July 2013.
The Thirty Meter Telescope and Astronomy in Canada, June 30, 2013
Tremblay, G.,Zohar, S., Bravo, J., Potsepp, P., and Barker, M. The Canada Foundation for Innovation's outcome measurement study: a pioneering approach to research evaluation Research Evaluation, 19(5), December 2010, pages 333–345
U.S. EPA. (2009). *Integrated Science Assessment for Particulate Matter (Final Report).* Washington, DC: Report No.: EPA/600/R-08/139F. Retrieved from: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546

United States Department of Energy, Office of Energy Efficiency & Renewable Energy, Jeff Dowd, "Return on Investment (ROI) impact evaluation in DOE EERE", Washington, District of Columbia, Federal Evaluators Conference November 1, 2012. (Ruegg R, O'Connor A, Lomis R. (2013). *Evaluating realized impacts of DOE/EERE R&D programs.* Report No.: DOE/EE-1025
United States Department of Energy, Office of Science, Four Years Later: An Interim Report on *Facilities for the Future of Science: A Twenty Year Outlook*, August 2007.
University of Arizona, Eller College of Management, "Astronomy, Planetary and Space Sciences Research in Arizona", an economic and tax revenue impact study, prepared for the Arizona Arts, Sciences and Technology Academy, Tucson, Arizona, October 2007
University of British Columbia, Planning and Institutional Research, Walter Sudmant, "The Economic Impact of the University of British Columbia", Vancouver, British Columbia, September 2009
Vancouver Board of Trade, "Economic Importance of Ocean Networks Canada", prepared for Ocean Networks Canada, Vancouver, British Colombia, January 2012
Vuola, Olli, and Ari-Pekka Hameri. "Mutually benefiting joint innovation process between industry and big-science." *Technovation* 26.1 (2006): 3-12.
Weinberg, A. (1967) Reflections on Big Science (Cambridge, Mass.: MIT Press, 1967).
Weiss, C. (1995) Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families In New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts , ed. James Connell *et al*. Washington, DC: Aspen Institute
Wessex Group Ltd, "Economic Impact of the Thomas Jefferson National Accelerator Facility Fiscal Year 2010", prepared for Jefferson Science Associates, Jefferson Lab, Newport News, Williamsburg, Virginia, January 2011
Wimbush, E., Montague, S. and Mulherin, T. (2012). *Applications of Contribution Analysis to Outcome Planning and Impact Evaluation* Evaluation Vol. 18, No. 3 July 2012, Sage Publications http://evi.sagepub.com/content/18/3/310.abstract

Appendix B: List of Persons and Institutions Consulted

	Canadian Facilities	
1	Henry Chen, Chief Financial Officer	TRIUMF
2	Doug Simons, Executive Director	Canada-France-Hawaii Telescope
3	Nigel Smith, Director	SNOLAB
4	Louis Fortier, Project Leader	CCGS *Amundsen*
5	Keith Levesque, Marine Research Manager	CCGS *Amundsen*
6	Mark Dietrich, President and Chief Executive Officer	Compute Canada
7	Nikki Macdonald, Executive Director, Corporate Operations	Ocean Networks Canada
8	Maia Hoeberechts, Associate Director	Ocean Networks Canada
9	Benoit Pirenne, Director, User Engagement	Ocean Networks Canada
10	Adrian Round, Director, Observatory Operations	Ocean Networks Canada
11	Tom Ellis, Director of Research	Canadian Light Source
	International Facilities	
12	Tim Moltman, Director	IMOS, Australia
13	Ilja Bohnet, Commissioner for Research Field Structure of Matter	Helmholtz, Germany
14	W.G. Stirling, Director	Institut Laue-Langevin, France
15	Timothy Meyer, Chief Operating Officer	Fermilab, USA