Impact Analysis and Evaluation at NPL

1. Introduction

The reliability and acceptability of measurements made in the UK is maintained by a series of specialist laboratories funded in part by the Department for Business, Energy and Industrial Strategy (BEIS). This note summarises an evaluation framework for assessing whether this programme - known as the National Measurement System (NMS) - is generating outcomes as anticipated and that it constitutes value of money. This evaluation work is necessary for accountability and to ensure that future funding is allocated on the basis of what worked well in the past.

The NMS laboratories: In the UK, the National Physical Laboratory (NPL) is the UK's NMI and works in partnership with five designated institutes:

- LGC (formerly Laboratory of the Government Chemist) – designated for chemical and biometrology
- NEL (National Engineering Laboratory) – designated for fluid flow metrology
- BEIS - Regulatory Delivery (Department for Business, Energy & Industrial Strategy) – designated for legal metrology and static flow
- NGML (National Gear Metrology Laboratory) – designated for gears metrology
- NIBSC (National Institute for Biological Standards and Control) – designated for bioactivity metrology

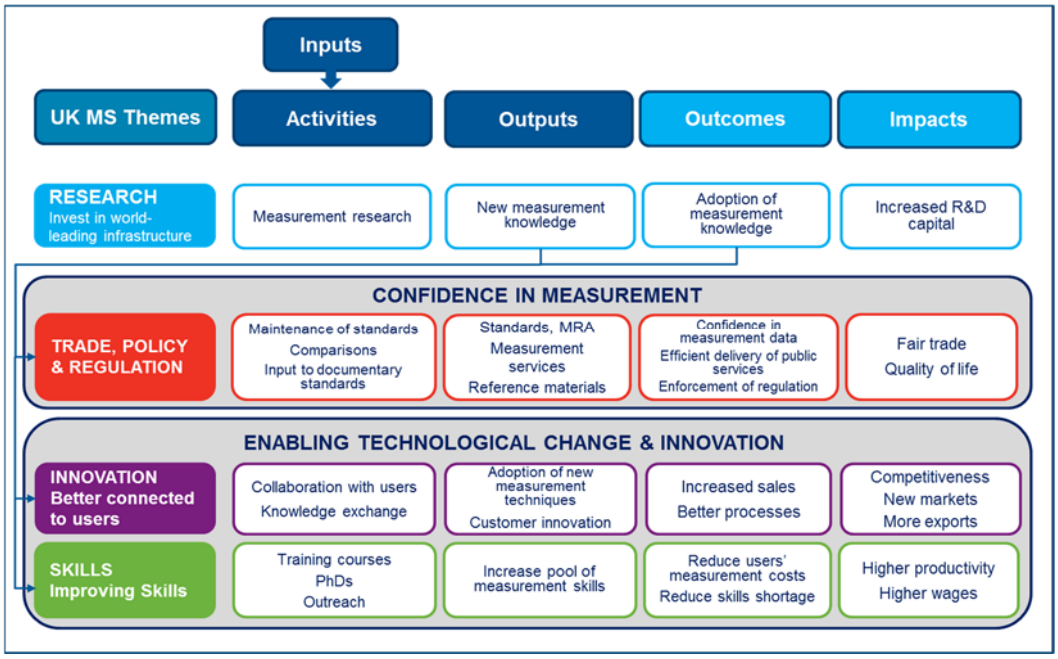
The role of the NMS: The current NMS strategy is based around a series of themes that emerged from a stakeholder consultation. To a first approximation these themes can be put into two broad categories:

- Enabling technological change and innovation:
  - Performing applied research in order to develop measurement tools, reference data, and protocols.
  - Helping organisations to innovate by supplying consultancy and training services.
- Underpinning trade and regulation:
  - Maintaining confidence in the nation's technical infrastructure
  - Supporting local authorities to enforce weights and measures legislation.

2. Logic Model and V-Diagram

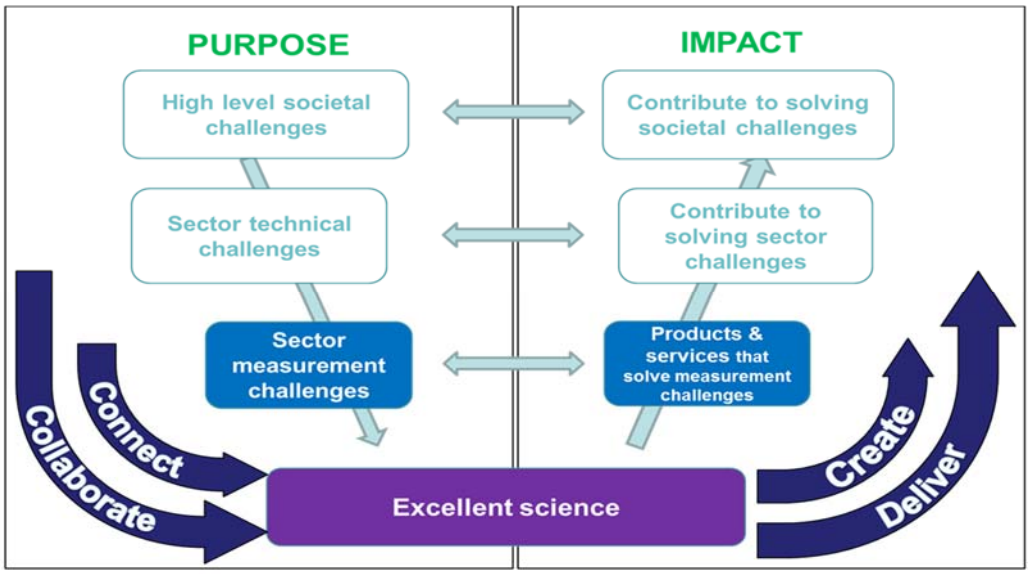
A logic model is a stylised way of describing how the programme operates and the anticipated pathways to impact. Such a diagram shows links in the causal chain and helps to identify the data required to assess whether the programme is functioning as expected. The logic model that follows (figure 1) provides a stylised representation of how the activities of the NMS laboratories leads to economic and social impacts.

Figure 1: Logic Model for the NMS



NPL has also developed another diagram that helps its staff to articulate how a specific piece of scientific work will create impact. That is, it enables NPL's staff to articulate the purpose of a scientific project and the benefit that it will create. This diagram is based on the 'V-Diagram' that helps engineers ensure they are both 'building the right thing' and 'building it right'. The first step is to identify a sector's measurement challenges that, when solved, will contribute to solving some higher level challenge. To do this NPL's scientists **CONNECT** to ensure they understand a sector's challenges and have identified the most important ones to focus on. And, wherever possible, they **COLLABORATE** to develop the right solutions. NPL's connections and collaborations with measurement users ensures its staff understand the route to impact so as to can **CREATE** and **DELIVER** the required measurement solutions.

Figure 2: NPL's V-Diagram



3. Contribution Analysis

Contribution analysis is a theory based approach that builds on the theory of change embodied in a logic model. Contribution analysis is conducted through a series of steps which are progressed through with appropriate iteration:

- Set out the expected attribution of impact to be considered
- Develop a theory of change, based on the programme's logic model
- Gather existing evidence on the theory of change
- Assemble and assess the contribution story and any challenges to it
- Seek out additional evidence, including through the evaluation methods set out above
- Revise and refine the contribution story

Mapping Methods to the Logic Model: The current evaluation plan is based on a mixed-methods approach, collecting both quantitative and qualitative information against a range of activity, output, outcome, and impact categories. The role of these methods in assessing the different elements of the logic model is presented in figure 3.

Figure 3: Mapping of evaluation methods against the logic model

Themes	Activities	Outputs	Outcomes	Impact
Investing in a world-leading measurement infrastructure. (Research)	Indicators	Indicators Case studies Peer Review	Indicators Case studies	
Ensuring good policy, standards, and regulations (Trade and Regulation)	Indicators	Indicators Case studies	Case Studies Surveys Economic Modelling	Economic Modelling
Getting better connected to our end-users to deliver impact (Innovation)	Indicators	Indicators Surveys Case studies	Case Studies Survey Econometrics	Econometrics
Improving the UK's measurement skills (Skills)		Indicators	Case Studies Surveys	

Qualitative Approaches: Qualitative methods can be used to explain 'why' and 'how' a given impact was generated. That is, case studies are useful for testing the hypothesised theory of change embodied in the logic model.

- We will perform case studies of specific outputs from a given element (theme) of the logic model. These case studies will identify the activities that led to a given output and investigate whether it led to clear outcomes.
- We will interview the direct users of the measurement services supplied by the NMS laboratories, primarily, the large calibration labs. The focus of this work would be on quantifying the cost and disruption of them having to go overseas for high-level calibration services if the NMS laboratories were to withdraw from certain areas.

- Peer Review: The International Science Review was commissioned by NPL's Science and Technology Advisory Council (STAC) to assess the quality and relevance of the science carried out at NPL.

Quantitative Approaches: It is important to have objective statistical evidence for the causal effect of the NMS laboratories on economic outcomes. The following methodologies include both established techniques and some more experimental methods.

- Indicators: Administrative data will be used to construct indicators for the activities and outputs of the NMS laboratories. The indicators have been designed to align with the logic model and can be reported annually.
- Surveys: The NMS has experience with customer surveys going back as far as 2005. The total database of both users and customers contains around 7,000 companies and, typically, a sample of 400- 600 telephone interviews is conducted by a market research company.
- Modelling: Models will be used to estimate impact in situations where it's unrealistic to pursue an empirical evaluation. That is, modelling is useful in cases where the distance between an NMS laboratory and the ultimate beneficiary is large and there is a range of factors that would confound an empirical evaluation. The intention is to develop studies that blend models from engineering (or medicine) with economic thinking.
- Event Studies: We will use stock market data to perform event studies based on the introduction of a particular standard or protocol. The basic idea is to find the abnormal return attributable to the introduction of a new standard by adjusting for the return that stems from the general price fluctuations in the stock market as a whole.
- Econometrics: In 2015, economists from NPL worked with Frontier Economics to quantitatively assess the effect of direct support from the NMS laboratories on the survival and growth of firms that used their services.

Annex 1 lists some of the challenges that are encountered when performing impact analysis. The challenges listed in this annex are particularly relevant when considering embarking on an empirical evaluation.

Annex 1: Evaluation Challenges

The attribution of impact is difficult because most outcomes involve inputs by the firms themselves, as well as, inputs from other public agencies or institutions. Furthermore, outcomes can also be effected by a range of external factors. All this makes it difficult to attribute a particular outcome to support from one specific laboratory. This subsection lists the challenges to consider when applying empirical methods:

- **Variation in firms' inventiveness**: Firms differ in terms of the ability of their management and staff, which means that some firms would be more inventive than others regardless of whether they are supported by the labs. Furthermore, the more inventive firms will probably have more need for the leading edge services of the NMS laboratories than the less inventive firms. If so, customers would have been more inventive than non-customers even if they had been denied access to the NMS laboratories.
- **Unique Primes**: The NMS laboratories work closely with a collection of very large multinational companies (primes) that dominate high technology sectors, such as, aerospace. It's difficult to find a convincing control group for the primes because all such firms use the laboratories and, in any case, each prime is unique.

- **Time lags and the skewed distribution of benefits**: It may take a number of years before the benefit of working with one of the NMS laboratories shows up in a firm's performance. Moreover, it could be that although innovation is generally beneficial, a small number of firms receive very large benefits while most firms receive small benefits. It follows that unless these fortunate firms are included in the sample, we will underestimate the average benefit.
- **No Experiments:** It is sometimes suggested that such programmes should be evaluated by performing experiments. While this type of approach works well elsewhere it is generally not going to be appropriate for evaluating the NMS. For example, it's not really practical to stop performing certain key comparisons in order to see if trade is damaged by a failure to maintain mutual recognition agreements.
- **Can't track indirect benefits:** Sometimes an organisation or individual can benefit without directly using one of the NMS laboratories. In particular, firms that take traceability from commercial calibration labs will benefit from the maintenance of primary standards by the NMS laboratories. Moreover, firms that engage in manufacturing or production benefit from the traceable calibration services and mutual recognition arrangements (MRAs).
- **Very few big changes:** The NMS laboratories have existed in a similar form for at least thirty years and in some cases much longer. Furthermore, most of the capabilities have evolved gradually over a large number of years. And, there are very few examples where the laboratories have shut down a capability or withdrawn a popular calibration service. This high degree of continuity makes it difficult to perform studies that use a major policy change that separates the time before from the time after.