A Review of Research and Development Evaluation Across the Globe

Nor Hasni Osman *, Abu Bakar Allumi Nura

School of Technology Management and Logistics, College of Business, Universiti Utara Malaysia, Sintok, Kedah, Malaysia

Abstract

In view of the present rapid changing of technology and economic landscapes, the Research and Development (R&D) field has garnered more attention than before throughout the world. Industrial innovation in this knowledge-driven era demands R&D as the undisputable key for improvement and to maintain competitiveness. Parallel to the industrial and economic growths, more and more researchers and skilled workers are required to fill in the specific fields all over the globe. The growing numbers of R&D facilities and personnel, however, is unable to provide assurance that improvement would take place after a particular research result is implemented. R&D laboratories themselves are systems that process inputs and generate outputs, where in the process, inevitable challenges are encountered to ensure that the task is not only capable of solving particular problems, but also to ensure positive and innovative outcomes. R&D is also subject to performance analysis to measure the effectiveness, accomplishments, and productivity amongst others, because performance is an importance element in grant disbursement. In doing so, R&D evaluation is an important aspect and practised in most parts of the world with different methods and strategies. This paper reports an R&D overview and explores the need of the global community for post decision evaluation to ensure that the decision is sustainable in the long run.

Keywords: R&D, post decision evaluation, R&D lab, performance analysis.

1.0 Introduction

Researchers have the task to utilise systematic and intensive studies to resolve specific problems from different angles and perspectives in order to embrace innovation, increase productivity, and profitability. Realising that knowledge has no boundaries and roams perpetually, many past and current theories that have been applied in research are easily busted, annulled, overruled, out-dated, and superseded in this challenging and competitive world. It is important to keep track of the ever changing scenarios, since it was revealed that even though a lot of R&D undertakings have been done, generally only a small amount or at some instances there is no research evaluation that follows

* Corresponding Author: Tel: +60194164666

E-mail Address: has1218@uum.edu.my
afterwards. Most research studies conclude once they are implemented without due consideration given to the potentially infinite innovation that emerges every single day where knowledge is expanding and old ideas are often replaced with new ideas. Apart from China, Asian countries, particularly in Malaysia, Singapore, Japan and India, have recently seen an upsurge in R&D innovation where research has crossed borders, making them well known in the global R&D scene (Krishna, Patra, & Bhattacharya, 2012). The increasing numbers of R&D worldwide reveal that even though different methods and strategies are employed, there is still room for R&D improvement to enhance performance by measuring the effectiveness, efficiency, accomplishments, and productivity. As such R&D evaluation is an important aspect.

2.0 Research and development concept

R&D concept has been developed decades ago with various definition compilations subject to whatever backgrounds or purposes of the R&D. Generally, the character of R&D has been conceptualised as a systematic study. According to Godin (2001), National Science Board (2001), and UNESCO (1978), R&D is defined as an innovative and ingenious task that is carried out by employing systematic instruments to enrich piles of knowledge in devising latest applications. As for this study, R&D refers to any probe or work performed on a methodical and proficient approach to magnify the piles of knowledge in scientific, technological, managerial, and socio-political perspectives, with the objective of devising new concepts.

3.0 Post decision evaluation of R&D

A job might have been finished, but it has not necessarily been completed since finishing and completing are two different words with different meanings, purposes, and terms. R&D practices should be familiar with the term before, during, and after in the decision making process because measures should be taken before implementation, during implementation, and what happens after the implementation to avoid “fishing in muddy water”, as suggested by Kralj (2011). It was reported by Nill, Landwehr, Carabias, and Gérard (2009) that about 40% or more of the respondents declared not having their R&D evaluated and informed that the evaluation was in fact carried out before the implementation of the results. In some instances, most of the evaluation was performed internally. It is a rare phenomenon when they would evaluate the results after the implementation. If any, the evaluations mostly focused on the number of researchers, and the allocated funds and state revenues. Schainblatt (1982) indicated that in spite of billions of American dollars spent in R&D projects, only 20% of the R&D managers of major companies performed productivity appraisal of their R&D operations, and only a few number of them made assessment of any kind on profits generated from R&D investments. Certain engineers and scientists are cynical and afraid that their low and insufficient productivity will be revealed if the performance measure is done (Brown &
Svenson, 1998), leaving a delusional thought that all systems of performance measures are impractical, thus affiliating the practices of R&D as a “Lab”, which comprises input, process, and output elements.

4.0 The R&D lab as a system

Input, process, and output are elements that work in a system in an active organisation. Funds, manpower, data, concepts, instruments, amenities, and specific instructions are considered as the basic input. The R&D lab is the processing mechanism which converts the input into output; from preliminary draft proposals through performing research, hypotheses testing, until reaching the final revision. The new breakthrough whether it is a new concept, theory or idea, will later transpire as the output. The output ought to be measured and processed in terms of cost, quality, and quantity to assess its worthiness. Three foremost points in evaluating a decision as stated by Brumm (1986) are sustainability, holistic approach, and self-assurance stage. Sustainability is important since the finest concepts and proposals are destined to fail ultimately if the motivating forces are inconsistent as time goes by. It is only a premature success if no follow-up steps are taken. A holistic approach, on the other hand, is an art of expecting the unexpected by thinking outside the box. The last point is the self-assurance stage where the researchers must ensure that the R&D output resolves the related issues and confirms that the rendered decisions are either correct or reasonable decisions.

Despite the fact that some of the ultimate obstacles encountered in measuring and evaluating R&D performance have been reviewed in some previous research, the most challenging task is to choose the best relevant set of correct measures that is coherent with the right subjects (Ojanen & Vuola, 2005). There are other problems encountered in measuring and evaluating R&D that need to be taken into consideration, such as contribution objective and expected period of completion. Brown and Svenson (1988) itemised four elements that instigate failure in R&D evaluation, i.e., extreme weight on internal measure, too much attention on behaviour, employing complicated measure procedure, and dearth of objective. Later, Kerssens-van Drongelen (1999) included acceptability problem into the list when he realised that experts argued the evaluations to be anti-productive while other argued that the measures are inappropriate or insufficient. Pappas and Remer (1985) added that this dissent is a setback to compete with.

R&D post decision reviews are often observed as an extra bureaucratic task where the willingness to cooperate with the application for revaluation work is usually at the bare minimum. If the required data are inconsistent with the data gathered at preceding checkpoints of the research, it will depreciate the project team’s trust in the company’s project management. As such, it is recommended that there should be a platform to ensure that effective post decision measurement systems exist to endorse R&D performance analysis so that the practice could be conducted at ease.
5.0 Rationale for R&D performance analysis

A measurement is only of value if utilised in a proper manner when making a decision. Lee, Son, and Lee (1996) claimed that evaluating R&D effectiveness is an undisputed factor in establishing whether such venture is tolerable or otherwise, and whether its highest output is accomplished and able to cope with the competition from new discoveries. Kerssens-van Drongelen and Cooke (1997) outlined two purposes of performance measurement; firstly, it motivates people, and secondly, to diagnose undertakings (e.g., projects) and units in an organisation. Furthermore, the studies by Kerssens-van Drongelen and Cooke (1997), Gold (1989), Meyer, Tertzakian, and Utterback (1997), and Loch and Tapper (2002) had classified R&D metrics according to the purpose of performance analysis in R&D.

6.0 Dimensions of R&D performance analysis

Although the number and the allocated amount of government grants are amongst the most common measure, what matters in the end is the accomplishment of the scientific objective. Berven (2011) insinuated that if the allocated amount of grants is treated to measure a performance, a decrease in the amount issued and the revaluation process might delay existing performance. It is necessary however to justify the actual allocated amount of grants and which specific field the R&D is designed. Chiesa, Coughlan, and Voss (1996) pointed out that when R&Ds begin to be responsible in their effectiveness and efficiency, as well as the likelihood for after decision assessment, it will uplift the awareness of both practitioners and academics (Brown & Svenson, 1988; Chiesa & Masella, 1996; Driva, Pawar, & Menon, 2000; Godener & Söderquist, 2004; Hauser, 1998; Ojanen & Vuola, 2005; Pappas & Remer, 1985; Pawar & Driva, 1999; Pillai & Rao, 1996; Poh, Ang, & Bai, 2001; Werner & Souder, 1997). Consequently, in view of the amplified attention in R&D performance measure, intellectuals began to explore the subject matter from different angles while clients focused more on planning in corporate and business strategies (Kumpe & Bolwijn, 1994; Pearson, Nixon, & Kerssens-van Drongelen, 2000). Kim and Oh (2002) considered the performance measure ought to be dealt in two separate perspectives, i.e., who is the authorised person to conduct the task and how to deliver the task effectively. Cordero (1990) suggested the need to establish a practical affiliation between R&D performance and performance measure system that works in tandem. This would form the fundamentals that R&D productivity require as an efficient system to measure its performance.

Many researchers, found to concentrate on one analysis unit (Collier, DeMarco, & Fearney, 1996), were amongst the pioneers who assumed that performance could be measured by project-level metrics. Schainblatt (1981), Cooper and Kleinschmidt (1988), and Patterson (1983) were in agreement with this idea. Furthermore, Schainblatt (1981) and Wilson et al. (1994) recommended individual-level metrics, while Stahl and Steger (1977), and Moser (1985) recommended team-level metrics. Meanwhile, Cordero (1990) and Wacker (1998) suggested from-level metrics. In respect to “who is the authorised person to conduct the measurement”, Frost and Whitley (1971) established that it is the responsibility of the researcher himself or herself, together with his or
her peers to perform the evaluation procedure, resolving that it is essential to engage several evaluation personnel instead of depending on the project directors and R&D heads. Stressing self-esteem as a factor, the process of measuring R&D performance is characteristically fragile (Keller & Holland, 1982). Wacker (1998) also agreed to engage several evaluators which include the researcher as well, external stakeholders, and his or her colleagues. With respect to “how to deliver the task effectively”, multiple citations are available that elaborate on the effect of more qualitative assessments on fostering an appropriate system in measuring performance. Barnowe (1975) recommended leadership as one of important criteria, while Wacker (1998) believed that subjective measures could be performed if the task is carried out by a group of evaluators comprising the researcher, his or her associates, and external auditors. Srinivas (2010) pointed out that when and how the evaluation takes place as deciding factors by indicating that the most preferable time to undertake the measure is when the research team remembers the most, i.e., soon after the delivery of the project and when all glitches have been taken care of. While still fresh, good supervision will generate new fruitful ideas. Nevertheless, it needs a stretched time frame along the process to justify the quality of the implementation before the result can materialise. This could be the adjusting period to make changes and to encounter challenges before the solution is finally found. Srinivas (2010) later explained further in “what to measure” factor by listing out the following terms:

- **Openness** – honesty and transparent.
- **Objectivity** – stick to the main objective and avoid prejudice.
- **Document success** – proper document organisation.
- **Look with hindsight** – take heed on unknown risk.
- **Being futuristic** – concentrate on the prospect of the future.
- **Reading between the lines** – recognising pros and cons, and learn the
  positive and negative aspects.

The methods have been classified into quantitative-objective metrics, quantitative-subjective, and qualitative-subjective metrics (Wacker, 1998). These methods depend on the type of measure, either numerical or non-numerical, and whether it is based on objective information or subjective judgments. It was found that, integrated metric is the most fruitful method which combines multiple objective and subjective methods. Integrated metric, that consists of an interrelated but divisible set of qualitative and quantitative techniques, can be adaptably employed throughout all R&D categories. Along with Werner and Souder’s literature, there are some other research classifying R&D metrics by grouping them into the R&D type (Brown & Gobeli, 1992; Hauser & Zettelmeyer, 1997; Kim & Oh, 2002; Loch & Tapper, 2002; Werner & Souder, 1997). There are only about 20% of R&D research revaluated after termination, making it not of any importance (Zedtwitz, 2002). As a matter of fact, R&D organisations are generally convinced that there are vast advantages in post-decision revaluation, but apparently they are not making good use of the educational prospect. It was explained further that Menke (1997), who conducted a benchmark study of 79 premier R&D
organisations, found less than 25% out of 79 organisations were fully engaged in post-project revaluations.

Table 1

**Global R&D Evaluation Strategy**

<table>
<thead>
<tr>
<th>Region</th>
<th>Country</th>
<th>Evaluation Strategy</th>
<th>Weakness of The Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>China</td>
<td>Associates review panel</td>
<td>Institutionalisation weakness and imbalance in review capacity</td>
</tr>
<tr>
<td></td>
<td>Singapore</td>
<td>Comprehensive variable evaluation</td>
<td>The variables are associated to the environment and consistently differ</td>
</tr>
<tr>
<td></td>
<td>Korea</td>
<td>Meta-analysis and in-depth review</td>
<td>Too rigid without contingency plans to accommodate incidental changes</td>
</tr>
<tr>
<td>Europe</td>
<td>United Kingdom</td>
<td>Target recognition by government institutions</td>
<td>Too bureaucratic and procedural</td>
</tr>
<tr>
<td></td>
<td>Turkey</td>
<td>Phrase anchored rating scale</td>
<td>Misleading, time consuming, and occasionally unreliable</td>
</tr>
<tr>
<td>South America</td>
<td>Venezuela</td>
<td>Annual review using national innovation metrics</td>
<td>Disconnection between the industry and the universities in terms of R&amp;D effort</td>
</tr>
<tr>
<td>Africa</td>
<td>Egypt</td>
<td>Task selection and generation</td>
<td>Optimisation is constrained and only solves non-complicated issues</td>
</tr>
<tr>
<td></td>
<td>South Africa</td>
<td>Treatment effect analysis</td>
<td>Too many fabricated positive rates</td>
</tr>
</tbody>
</table>

6.0 R&D evaluation status across the globe

R&D evaluations by all walks of professionals in different parts of the world have employed different methods and strategies. This paper provides a general outlook on how the performance evaluation is practised around the globe, which includes some Asian countries, two European nations, a South American nation, and two African countries. Table 1 illustrates the key points of evaluation strategy and the elaboration will follow afterwards.

6.1 Asia

**China**

A comparatively comprehensive framework is devised for R&D evaluation in China. Generally it is a desk study with the framework dimensions comprising objectives, goals, administration, implementation, effectiveness, and impact. This strategy is more
qualitative than quantitative. The evaluation is reviewed by the associates of the peer review panel to determine the project approval. The review panel consists of a group of experts in numerous fields and disciplines. The weakness of this strategy is that although this approach is very meticulous and intricate; institutionalisation weakness causes ambiguity on which project should be reviewed and how; and low guidelines for specific implementation and imbalance in review capacity. These weaknesses indicate shortage of capable staff to man the evaluation which makes the review more concentrated upon the evaluation task itself than the results of the evaluation.

**Singapore**

R&D evaluation in Singapore is performed mostly by embarking on a complete evaluation of a list of variables that is believed will potentially impact the selection and management of R&D. The variables are risk and return, strategic consideration, market exposure, resource and time based competitiveness constraint, and higher added value. The risk and return is an imperative variable since the real success and outcome of R&D remain unknown until they finally come into existence; when maximum return is the priority motive in decision making. As such, one of the best ways to gain competitive advantage is by reducing the allocated time from research to development. The next variable is strategic consideration which is founded on the strategic thinking skill because it has impact on funds and time obligation (Danila, 1989; Liberatore, 1987). The other stipulated variable is market exposure due to the fact that the prevailing upsurge productivity demands of high quality standard and ever increasing product innovation. The other variable refers to resources and time based competitiveness constraints, which refers to the scarcity of resources that will affect R&D decisions, progress, and etc. Higher added value is the other justified variable which refers to the engagement of unique technology in R&D innovation with added value, which will have a positive impact on the processing of products and the general development. In terms of weakness of the strategy, the main weakness is possibly the shortage of worldwide applicability because the variables are strongly attached to the environment, strategy, policy, and economic development that may not be suitable to other parts of the world (Liao & Greenfield, 2000).

**Korea**

Korea exercises R&D evaluation on two bases, i.e., the self-meta-analysis and in depth evaluation. The self-meta-analysis is a procedural strategy that evaluates the accomplishment of R&D corresponding to the performance scheme. The in depth analysis is used to identify problems confronted during implementation and concurrently rectify the problems to ensure effectiveness and efficiency. These strategies work in tandem. However, the weakness of this strategy is that it is too rigid and has no contingency plans where no account is taken on natural disaster, economic upturn and downturn, social upheaval, and political changes.
6.2 Europe

United Kingdom

In United Kingdom, the assessment of R&D is performed by setting the objectives that have to be achieved within the specified time to validate the fund amount taken. The government designs a speeding plan to be adhered to by all R&D organisations. The evaluation project is firstly selected by UK research council via traditional *ex ante* evaluations ahead of the handing of grants to safeguard a systematic review process. Upon approval, the R&D report is later submitted with matters ranging from the content of the research, income of staff members, publications, and etc. In spite of being very transparent and sequentially interlocked, this strategy is found to have too much procedure and characteristically too bureaucratic.

Turkey

Evaluating R&D in Turkey is conducted based on the classification of the research into three categories; curiosity driven (academic), customer driven (applied), and technology and innovation driven research. Each category is evaluated separately by five to eight reviewers who later will congregate as a panel to reach a final consensus. During evaluation, a phrase anchored rating scale is used. Even though the review process is simplified and reduces prejudice and variability of judgements, the reliability of the strategy is questionable. The apparent weakness is that it demands an exceptional observation talent and sufficient determination of analytical subjects. The strategy is also occasionally misleading and consumes too much time.

6.3 South America

Venezuela

In Venezuela, R&D evaluation is performed annually where the government draws up a list of national innovation metrics to assess the standard and competitiveness of R&D projects effectively. The evaluation is carried out in three different stages; innovation inputs, innovation outputs, and innovation outcomes. Innovation inputs are the deciding factor in the amount of funds and other support invested, while the innovation outputs observe whether the investment generates result and increases performance. Innovation outcomes focus on how many jobs and how much profits are created. The weakness of the strategy is that it does not perform as expected. The industry and university are not unified in the R&D effort whereas the assisted projects are not increasing as predicted. There is also lacking in funds for skills upgrade and training, and private companies are not interested in financing R&D projects.
6.4 Africa

Egypt

In Egypt, the main focus is project selection and generation to guarantee improved project development. The emphasis is to stimulate project selection models by using an effective communication system that is inter-connected to corporate priorities, implementation, strategy execution, and technical related issues. A project that is potentially viable technically and commercially will be projected and quantified by the prospective implementers and technologists, with the projecting and quantifying appraisals done from both the financial and marketing viewpoints. Schmidt and Freeland (1992) revealed that although this strategy was criticised to have constrained optimisation problem or as pointed out by Mitroff (1997) that it only solves simple R&D management matters, project selection strategy is one of the evaluation key elements and designs.

South Africa

As for R&D evaluation in South Africa, more consideration is given to treatment effect analysis based on harmonised data using non-parametric nearest-neighbour matching analysis. The weakness of this strategy is that the collected data depend on either sufficient or not the available nearest-neighbours, because it would result in fabricated statistics and false positive rates.

7.0 Concluding remarks

R&D needs a constant evaluation and post decision evaluation to measure its performance. Successfully implemented research tasks just end there and have never been kept track or re-evaluated, although there will always be space for improvement. This paper has elaborated on several evaluation strategies that have carried out around the world and explained some of these strategies and weaknesses. There are strategies that compare events before and after a change in policy is implemented, some strategies use parametric methods, and others employ non-parametric methods. Progressive R&D is actually functioning in a continuous circle and exposed in an ever-changing environment with constant innovation updates. The R&D “lab” itself is a system that processes inputs and generate outputs, where in the process inevitable challenges are encountered to ensure that the task not only solves particular problems, but also to ensure positive and innovative outcomes.
References


Kumpe, T., and Bolwijn, P. T., (1994). Toward the innovative firm-Challenge for R&D
management. *Research technology management, 37*(1), 38.


