

OECD Education Working Papers No. 122

Evaluative thinking for successful educational innovation

Lorna Earl,

Helen Timperley

https://dx.doi.org/10.1787/5jrxtk1jtdwf-en

OECD

Unclassified **EDU/WKP(2015)11**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **16-Jun-2015**

English - Or. English

DIRECTORATE FOR EDUCATION AND SKILLS

EDU/WKP(2015)11
Unclassified

EVALUATIVE THINKING FOR SUCCESSFUL EDUCATIONAL INNOVATION

Education Working Paper No.122

By Lorna Earl and Helen Timperley

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

Lorna Earl & Associates, Toronto, Canada (lorna@lornaearl.com)
Helen Timperley, University of Auckland (h.timperley@auckland.ac.nz)
David Istance, Senior Policy Analyst, EDU/IMEP (david.istance@oecd.org)

JT03378768

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

OECD EDUCATION WORKING PAPERS SERIES

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcome, and may be sent to the Directorate for Education and Skills, OECD, 2 rue André-Pascal,75775 Paris Cedex 16, France.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

---

www.oecd.org/edu/workingpapers

Copyright OECD 2015

TABLE OF CONTENTS

INTRODUCTION 6
WHY EVALUATION AND INNOVATION? 7
Evaluative Thinking 8
SETTING THE STAGE 9
What is Educational Innovation? 9
What is Educational Evaluation? 10
THE NEXUS BETWEEN INNOVATION AND EVALUATION IN EDUCATION 16
Who are the Innovators? 16
Who are the Evaluators? 17
Defining the Innovation 18
Multiple Stakeholders; Multiple Interests 20
Recognising Contexts 21
Identifying the Purpose(s) of Evaluation within the Innovation 21
GETTING ON WITH IT: APPROACHES AND METHODS 23
Framing Evaluation Questions 23
Collecting Evidence, Timelines and Timing 24
Organising and Analysing the Evidence 25
MAKING SENSE OF IT ALL 27
Interpretation as Building Knowledge 27
Capturing and Mobilising the New Knowledge 30
Sustainability 32
INNOVATION AND EVALUATION: SYNERGIES AND TENSIONS 34
Being Open to Improving Ideas 34
Being Pragmatic 35
Negotiating, Renegotiating and Renegotiating 36
CAPACITY FOR EVALUATIVE THINKING: PULLING IT ALL TOGETHER IN THE NEXUS 38
REFERENCES 39

Boxes

Box 1. Examples from the OECD Innovative Learning Environments 10
Box 2. Types of Evaluation 11
Box 3. Program Evaluation Standards 12
Box 4. Six Steps of Program Evaluation 13
Box 5. Next Generation Evaluation: Embracing Complexity, Connectivity, and Change 14
Box 6. Cognitive Biases 17
Box 7. Evaluator Competencies 18
Box 8. The Theory of Action Defined 19
Box 9. Monitoring a Theory of Action 20

Box 10. Who Cares to Know? 21
Box 11. Just In Time Evidence 24
Box 12. Analytic Tool in Knowledge Forum 25
Box 13. Using Vignettes To Understand Innovation 26
Box 14. Multiple Perspectives in Interpretation 28
Box 15. Dynamic Spiral Critical For Knowledge Creation And Sharing 29
Box 16. Knowledge Building Across Projects 30
Box 17. Knowledge Animation 31
Box 18. Habits of Mind 33
Box 19. Sustainability 33
Box 20. Embedding Evaluative Thinking 36
Box 21. Capacity 38
Box 22. Diverse Perspectives in Creative Nexus 38

ABSTRACT

In this working paper, Earl and Timperley argue that evaluative thinking is a necessary component of successful innovation and involves more than measurement and quantification. Combining evaluation with innovation requires discipline in the innovation and flexibility in the evaluation. The knowledge bases for both innovation and evaluation have advanced dramatically in recent years in ways that have allowed synergies to develop between them; the different stakeholders can bring evaluative thinking into innovation in ways that capitalise on these synergies. Evaluative thinking contributes to new learning by providing evidence to chronicle, map and monitor the progress, successes, failures and roadblocks in the innovation as it unfolds. It involves thinking about what evidence will be useful during the course of the innovation activities, establishing the range of objectives and targets that make sense to determine their progress, and building knowledge and developing practical uses for the new information, throughout the trajectory of the innovation. Having a continuous cycle of generating hypotheses, collecting evidence, and reflecting on progress, allows the stakeholders (e.g., innovation leaders, policymakers, funders, participants in innovation) an opportunity to try things, experiment, make mistakes and consider where they are, what went right and what went wrong, through a fresh and independent review of the course and the effects of the innovation. This paper describes issues and approaches to each phase of the cycle. It concludes by outlining the synergies to be made, building capacity for evaluative thinking, as well as possible tensions to be addressed.

RESUME

Dans ce document de travail, Earl et Timperley mettent en avant l'argument que la pensée évaluative est un élément indispensable à une innovation réussie, et qu'il ne s'agit pas seulement de méthodes de mesure et de quantification. Combiner évaluation avec innovation exige de la discipline dans l'innovation et de la souplesse dans l'évaluation. Les bases de connaissances pour l'innovation comme pour l'évaluation ont vu une évolution importante ces dernières années, permettant le développement de synergies entre ces deux domaines ; les différentes parties prenantes peuvent apporter la pensée évaluative à l'innovation, en tirant parti de ces synergies. La pensée évaluative contribue aux nouveautés en matière d'apprentissage en fournissant des preuves pour documenter, recenser et mesurer le progrès, les succès, les échecs et les obstacles dans l'innovation en cours. Il s'agit de réfléchir aux preuves qui seraient utiles au cours des activités de l'innovation, et donc d'établir un champ d'objectifs et de cibles propices à déterminer le progrès de ces activités, acquérir des connaissances et développer des usages pratiques des nouvelles informations tout au long de la trajectoire de l'innovation. La génération d'hypothèses en cycle continu, le recueil de preuves, et la réflexion sur le progrès permettent aux parties prenantes (par exemple, les leaders de l'innovation, les responsables politiques, les bailleurs de fonds, et les personnes prenant part à l'innovation) d'essayer, d'expérimenter, de faire des erreurs et de considérer où sont ces erreurs, ce qui s'est bien passé ou ce qui a mal tourné, grâce à un bilan nouveau et indépendant du déroulement et des effets de l'innovation. Ce document décrit les enjeux et les approches de chacune des phases du cycle. Il conclut en indiquant les synergies qu'il reste à accomplir, ouvrant le champ à la pensée évaluative, ainsi que des tensions éventuelles à traiter.

INTRODUCTION

Evaluating innovation is a perplexing topic, with innovation and evaluation often perceived as fundamentally incompatible. Bernholz (2011) captures this tension very well when she says that: "Many people's first response to the challenge of measuring innovation is to declare the intention oxymoronic. Innovation is by definition amorphous, full of unintended consequences, and a creative, unpredictable process…Measurements, assessments, evaluation are — also by most definitions — about quantifying activities and products" (p.1).

Although innovation and evaluation in education do not always live together comfortably, they are not mutually exclusive. As will become clear, we argue that evaluative thinking is a necessary component of successful innovation and involves more than measurement and quantification. Although they have developed separately, the knowledge bases for both innovation and evaluation have advanced dramatically in recent years in ways that have allowed synergies to develop between them.

In this working paper, we describe how innovators, funders and evaluators can bring evaluative thinking into innovation in ways that capitalise on these synergies, while acknowledging the challenges inherent in doing so.

WHY EVALUATION AND INNOVATION?

Innovation has become the watch word of education around the globe. The opening line in the OECD report *Innovative Learning Environments* (2013) says it succinctly: "Innovation is a key element of today's societies and economies, and that includes how we learn" (p.11).

Education is seen as the foundation of prosperous societies and there is considerable concern that traditional education will not suffice to prepare students for the future. Innovation is necessary in education to meet the needs of the students and the societies they serve and to cope with the various challenges of the contemporary world: "Innovation drives economic influence; economic influence underpins global leadership; and global leadership requires innovation to solve the many problems facing humanity in the next half century. If this is correct, and innovation is the key, then even the best education systems in the world need to radically rethink what they offer every student" (Barber, Donnelly, & Rizvi, 2012, p. 1).

The case for innovation is well made by Mulgan (2007): "A contented and stable world might have little need for innovation. Innovation becomes an imperative when problems are getting worse, when systems are not working, or when institutions reflect past rather than present problems" (p. 9). However, as Mulgan and Leadbeater (2013) caution in relation to the innovation process: "You can't plan for a breakthrough! (p.15)…There needs to be humility to learn from experience and the inevitable surprises" (p. 17).

Because innovations matter, there is considerable interest in how things are going, with a wide range of agendas at play. Certainly funders and policy makers are interested in the success of innovation initiatives. Even though they see educational innovation as necessary, they often want a sense of security about their investment, so the call for innovation is often accompanied by a demand for accountability. Bernholz (2011) captures this sentiment in the epigraph to a publication for the MacArthur Foundation: "Give me something new and prove that it works" (p.1).

Although evaluation may provide funders and policy makers with confidence that the innovation is proceeding in productive ways and that their financial support remains warranted, we argue that evaluation has a much more powerful role within innovation when it is positioned as an integral part of the innovation process, contributing to the development and evolution of the innovation, with milestones of success to be tracked along the way emerging and being established and negotiated as part of the process. As Drucker (1985) wrote in his landmark paper: "Innovation is work rather than genius. It requires knowledge. It often requires ingenuity. And it requires focus…It may be difficult, but knowledge-based innovation can be managed. Success requires a commitment to the systematic practice of innovation" (p. 8).

Successful innovation may be rapidly changing in response to uncertainty and complexity, but the changes are not random. Leaders of innovation draw on a blend of creativity and discipline that allows them to react effectively in diverse and changing conditions. Rather than being unstructured, disciplined innovation involves constant problem definition, horizon scanning, situation analysis, monitoring of progress, creation of contingency plans, and feedback for improvement throughout the innovation process.

Evaluative Thinking

Evaluation methods and evaluative thinking provide the tools for systematically gathering and interpreting evidence that can be used to provide information about progress and provide feedback loops for refinement, adjustment, abandonment, extension and new learning.

The essence of evaluative thinking is expressed in a recent publication from the International Development Research Centre (Bennett & Jessani, 2011):

> Evaluative thinking is a means of thinking, of viewing the world, an ongoing process of questioning, reflecting, learning and modifying. Evaluative thinking is an inherently reflective process, a means of resolving the "creative tension" between our current and desired levels of performance. It allows us to define the lessons we want to learn, to determine the means for capturing those lessons, and to design systems to apply them in improving our performance. By going beyond the more time- and activity-bound processes of monitoring and evaluation, evaluative thinking is learning for change (p. 24).

Evaluative thinking contributes to new learning by providing evidence to chronicle, map and monitor the progress, successes, failures and roadblocks in the innovation as it unfolds. It involves thinking about what evidence will be useful during the course of the innovation activities, establishing the range of objectives and targets that make sense to determine their progress, and building knowledge and developing practical uses for the new information, throughout the trajectory of the innovation. Having a continuous cycle of generating hypotheses, collecting evidence, and reflecting on progress, allows the stakeholders (e.g., innovation leaders, policymakers, funders, participants in innovation) an opportunity to try things, experiment, make mistakes and consider where they are, what went right and what went wrong, through a fresh and independent review of the course and the effects of the innovation.

SETTING THE STAGE

Although there is increasing consensus about the importance and value of evaluating educational innovation, there are many issues involved in turning this idea into a productive reality. It helps to understand something about innovation and evaluation, as a backdrop to understanding their relationship. Both fields are themselves complex and evolving, with a wide range of definitions and representations, often contested. Before considering the nexus between them, it is important to consider the complexity of each of them individually.

What is Educational Innovation?

The word innovation is somewhat hackneyed and has come to have many meanings. This is true more broadly and within education. It is not possible in this paper to detail all of the ways that educational innovation is characterised and understood. Just as a taster, it can be evolutionary or revolutionary. It can address a single project or programme or be large-scale encompassing a complete system. It can be initiated from within or from outside. It can be simple, complicated or complex. It can involve products, content, resources, processes, people and organisational arrangements. It can include technical, economic, social and educational aspects.

The European Commission's Green Paper on Innovation (1995) defines it as: "…the successful production, assimilation and exploitation of novelty in the economic and social spheres" (p. 1).

Some authors suggest that to be called an innovation it must be radical and disruptive: "Innovation must not simply be another name for change, or for improvement, or even for doing something new, lest almost anything qualifies as innovation. Innovation is properly defined as an original, disruptive, and fundamental transformation of an organization's core tasks. Innovation changes deep structures and changes them permanently" (Lyn, 1997).

Although there are many ways to describe innovation, it is important to remember that innovation is an idea, but in the enactment each innovation is unique, with its own development, trajectory, and personalities. Each one is also anchored in the particular context from which it has emerged. In reality, educational innovation ranges from relatively straightforward (but not simple) school improvement activities, often in difficult contexts, to transformational approaches that disrupt the way that "schooling" and "learning" happen. In some contexts, the idea of school improvement in itself can be considered innovative because of the starting points and nature of the problems. In others, the innovation is intended to totally transform education systems and the nature of schooling. Between these two extremes lie a wide range of innovation models and paradigms that bring theories and initiatives together in a multitude of ways. Box 1 provides some examples of diverse innovations within the OECD Innovative Learning Environment programme.

Most educational innovations are made up of multiple players, working across a number of locations and guided by a wide range of theories, some of which have been tested in other contexts. The innovations themselves typically roll out in unpredictable ways and change along the way with everyone learning as they go.

Innovation in education is not a new phenomenon. Education has been involved in a cyclical process of reform for many decades, with attention to school effectiveness and school improvement. It has largely focused on evolutionary incremental change in schools in which effort is concentrated on trying to make changes within the existing context of schools.

Box 1. Examples from the OECD Innovative Learning Environments

A learning community in Andalucía, Spain is introducing diversified resources, teaching methodologies and moreover, the school builds its own *classroom curriculum,* counting on the democratic participation of students, and taking their cultural reality (Gypsy culture) into account.

In Victoria, Australia, a community college is focused on creating seamless cohesion between the vision, architecture, social environment and pedagogical approaches throughout the school that is the result of the depth of attention given to every aspect of planning and operations. Listening to the entire school community has led to the development and enactment of a shared vision around the wellbeing and education of the children.

A school community of educators in British Columbia, Canada is focused on creating learning environments for themselves and their students that are steeped in inquiry mindsets and a value system that honours the self, while recognizing the innate need to belong in a community. Through a focus on belonging, support, interdependence and respect for diversity. Students are immersed in an environment that offers a balance of structure and autonomy so that they experience a combination of safety and accountability as they take risks in their own learning.

Source: OECD, 2013

Proponents of radical innovation in the context of schooling suggest that it is necessary to go beyond improvement and to challenge long-standing and deeply held beliefs about what schooling is for, with fundamental shifts in the way people think about the nature of learning and the rhythm of interactions in learning environments. It cannot simply be the linear application of innovative ideas to defined educational problems and existing processes but is concerned with the social, personal, institutional and cultural process of change, as well as changes to education processes. This sentiment is reinforced by Hannon (2009) who argues that: "current 'school reform' and 'improvement' efforts are wholly inadequate to the scale of the challenge to prepare young people to live well and sustainably on this planet in the new century and that any new paradigm must entail 'a holistic transformational shift towards connected institutions and processes, at a whole set of levels'" (p.1).

This perspective means that, to be innovative, change must go beyond a single project or programme. Rather it encompasses interconnected parts of a system for the purpose of creating radical change that recognises the complexity of modern systems: "Most modern systems are both hideously complicated and bewilderingly complex and innovations within them are likely to be an interconnected set of innovations, where each influences the other, with innovation both in the parts of the system and in the ways in which they interconnect. Education is no exception" (Mulgan & Leadbeater, 2013, p.43).

In this paper, we are not advocating for a particular definition of innovation. Instead, we point out that the definition is contentious and provide some sense of the range of possibilities that have emerged in education that are called innovation, with a related influence on the kind of evaluation that is required.

What is Educational Evaluation?

The idea of educational evaluation is deceptively simple. It involves the systematic collection and analysis of data needed to make decisions and identify effects of educational initiatives or, as the American Evaluation Association describes it: "Evaluation involves assessing the strengths and weaknesses of programmes, policies, personnel, products, and organizations to improve their effectiveness" (AEA, ND).

In practice, evaluation in education is not a singular thing and it has always been a contentious and challenging domain. There may be agreement about the need to assess and improve educational change initiatives but there have been continuous debates in the field about the purpose of evaluation, the methods

that are used, what counts as worthy outcomes, how to measure important concepts, whether the evaluation should be defined by programme or by theory, and the list goes on. These debates arise, in part, because of the many different reasons for evaluation. The diversity is exemplified in a recently published comprehensive analysis of how different approaches to evaluation and assessment are being used around the world to support effective teaching and learning in schools. The results of the OECD Reviews of Evaluation and Assessment in Education, a major cross-country project involving the participation of 29 education systems are summarised in the comparative international report Synergies for Better Learning (OECD, 2013).

Although we address the issue of evaluation purpose in detail later in this paper, it is important to raise the issue here because the purposes have a major influence on the way any evaluation is planned. Why? Because evaluations are driven by what people want and need to know about and different people have different interests and perspectives. Evaluation in education was initially formulated around summative and formative purposes, with formative evaluations being conducted during programme development and implementation to provide direction on how to best achieve the goals or improve the programme. Summative evaluations were completed once the programmes were well established to determine the extent to which the programme achieved its goals. Both of these models were premised on a belief that the programme was relatively static and could then be "scaled up" by replicating it elsewhere. Box 2 gives a wide range of different types of evaluation, any of which might apply within a particular innovation.

Box 2. Types of Evaluation

There are many different types of evaluations depending on the object being evaluated and the purpose of the evaluation. Perhaps the most important basic distinction in evaluation types is that between *formative* and *summative* evaluation. Formative evaluations strengthen or improve the object being evaluated -- they help form it by examining the delivery of the program or technology, the quality of its implementation, and the assessment of the organizational context, personnel, procedures, inputs, and so on. Summative evaluations, in contrast, examine the effects or outcomes of some object -- they summarize it by describing what happens subsequent to delivery of the program or technology; assessing whether the object can be said to have caused the outcome; determining the overall impact of the causal factor beyond only the immediate target outcomes; and, estimating the relative costs associated with the object.

Formative evaluation includes several evaluation types:

- ***needs assessment*** determines who needs the program, how great the need is, and what might work to meet the need
- ***evaluability assessment*** determines whether an evaluation is feasible and how stakeholders can help shape its usefulness
- ***structured conceptualization*** helps stakeholders define the program or technology, the target population, and the possible outcomes
- ***implementation evaluation*** monitors the fidelity of the program or technology delivery
- ***process evaluation*** investigates the process of delivering the program or technology, including alternative delivery procedures

Summative evaluation can also be subdivided:

- ***outcome evaluations*** investigate whether the program or technology caused demonstrable effects on specifically defined target outcomes

- ***impact evaluation*** is broader and assesses the overall or net effects -- intended or unintended -- of the program or technology as a whole
- ***cost-effectiveness and cost-benefit analysis*** address questions of efficiency by standardizing outcomes in terms of their dollar costs and values
- ***secondary analysis*** re-examines existing data to address new questions or use methods not previously employed
- ***meta-analysis*** integrates the outcome estimates from multiple studies to arrive at an overall or summary judgement on an evaluation question

Source: Trochim, 2006

Programme leaders have tended to be interested in formative evaluation for internal accountability, with policy makers and funders wanting summative evaluation for external accountability. As the call for evaluation grew and "high stakes" decisions have been based on the results of summative evaluations, there has been concern about the pervasiveness of naïve and often shoddy evaluation practices that were being used.

These concerns about the quality of evaluations resulted in a set of Standards for Program Evaluation, published by The Joint Committee on Standards for Educational Evaluation in the US to provide guidance to programme evaluators (summarised in Box 3). This was produced first in 1981, revised in 2004 and again in 2013[1], and provides a broad base of expectations for high quality and ethical evaluation procedures.

Box 3. Program Evaluation Standards

Utility Standards: The utility standards are intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs.

Feasibility Standards: The feasibility standards are intended to increase evaluation effectiveness and efficiency.

Propriety Standards: The propriety standards support what is proper, fair, legal, right and just in evaluations.

Accuracy Standards: The accuracy standards are intended to increase the dependability and truthfulness of evaluation representations, propositions, and findings, especially those that support interpretations and judgments about quality.

Evaluation Accountability Standards: The evaluation accountability standards encourage adequate documentation of evaluations and a meta-evaluative perspective focused on improvement and accountability for evaluation processes and products.

External Meta-evaluation: Program evaluation sponsors, clients, evaluators, and other stakeholders should encourage the conduct of external meta-evaluations using these and other applicable standards.

Source: The Joint Committee on Standards for Educational Evaluation, 2013

[1] For details about the standards, see Program Evaluation Standards at http://www.jcsee.org/program-evaluation-standards-statements

Educational evaluation has typically taken a linear approach of describing the programme and using the programme definition to determine the evaluation design and methodology. Evaluators establish the range of stakeholders with an interest in the programme and in the evaluation and what they expect the evaluation to provide for them. They engage the programme team to get clarity about the nature of the programme, its structure, the underlying assumptions and the theory of action that underpin the expectation for success of the programme. They get information about the goals and outcomes for the programme and develop indicators and measures to provide evidence related to these outcomes. And, they determine an evaluation design that is consistent with the programme goals and purposes. Once they have this information, they choose and devise data collection procedures keyed to the theory of action, collect the pertinent data, analyse the data, draw conclusions and prepare a report for the programme team. The final step in the process is presenting the findings and recommendations for use by the stakeholders.

Many guides have been prepared to help programme leaders engage with evaluation and evaluators[2]. The example in Box 4 from the University of Washington gives a good overview of the nature of the process.

Box 4. Six Steps of Program Evaluation

- **Step 1: Define your stakeholders**: Your stakeholders are supporters, implementers, recipients, and decision-makers related to your program. Getting them involved early on will help you get different perspectives on the program and establish common expectations. This helps to clarify goals and objectives of the program you'll evaluate, so everyone understands its purpose.

- **Step 2: Describe the program**: Taking the time to articulate what your program does and what you want to accomplish is essential to establishing your evaluation plan. Your descriptions should answer questions like: What is the goal of our program? Which activities will we pursue to reach our goal? How will we do it? What are our resources? How many people do we expect to serve? Articulating the answers to those questions will not only help with accountability and quality improvement, but it will also help you promote the program to its beneficiaries.

- **Step 3: Focus the design of your evaluation**: Evaluations can focus on process, means, resources, activities, and outputs. They can focus on outcomes or how well you achieved your goal. You may also choose to evaluate both process and outcomes.

- **Step 4: Gather evidence**: Qualitative and quantitative data are the two main forms of data you may collect. Qualitative data offers descriptive information that may capture experience, behavior, opinion, value, feeling, knowledge, sensory response, or observable phenomena. Three commonly used methods used for gathering qualitative evaluation data are: key informant interviews, focus groups, and participant observation. Quantitative methods refer to information that may be measured by numbers or tallies. Methods for collecting quantitative data include counting systems, surveys, and questionnaires.

- **Step 5: Draw conclusions**: This is the step where you answer the bottom-line question: Are we getting better, getting worse, or staying the same? Data comparisons show trends, gaps, strengths, weaknesses. You can compare evaluation data with targets set for the program, against standards established by your stakeholders or funders, or make comparisons with other programs.

- **Step 6: Present findings and ensure use**: It is important that all the work you put into program evaluation gets used for quality improvement. When you present your findings and recommendations, it is important to know the values, beliefs, and perceptions of your group; build on the group's background and build on common ground; and state the underlying purpose for your recommendations before you get to the details

Source: Northwest Public Health Center, University of Washington

[2] See http://www.cdc.gov/EVAL/resources/index.htm for a compilation of evaluation guides and resources.

This process is generally necessary but not sufficient for evaluating innovation. Innovation is rarely intended to produce replicable programmes and the rapidly changing context requires approaches that are fluid and responsive, as they explain in the example in Box 5 from the philanthropic sector.

The field of evaluation is vast and growing, with a robust literature and "on-the-ground" practices being developed to move beyond simple models and to provide processes that can be integrated into the development of innovation. Evaluators and researchers studying evaluation have continued to both extend and refine the nature and form of evaluation processes to adapt to the changing demands for having defensible evidence available for important decisions about policy and practice. New methods for data collection and analysis have emerged to extend the scope and breadth of evidence that can be considered.

Within the field of innovation, technological advances have expanded the capacity for addressing social problems, and have created more data and more sophisticated data that requires evaluative expertise in order to understand the evidence and use it within the innovation. Box 5 describes trends that are emerging in next generation evaluation.

Box 5. Next Generation Evaluation: Embracing Complexity, Connectivity, and Change

New ideas have expanded the ways in which organizations think about improving society and have opened up additional opportunities for collaboration. Simultaneously, technological innovation has expanded the sector's capacity for understanding and addressing social problems, and it has created more data that the social sector can leverage in its work.

We found three primary trends that are driving the need for evaluation to evolve:

1. New Philanthropic Innovations: There is an increasing realization that traditional philanthropic models have had limited success in curing chronic social ills. Even tested solutions are forced to experiment when faced with the challenge of scaling to new populations and geographies. In addition, several new and often untested approaches are coming to the fore, including social entrepreneurship, impact investing, social impact bonds, and others that do not lend themselves to traditional methods of evaluation. This growth of experimentation in the social sector demands that evaluation better capture learning in complex environments.

2. Different Rules of Interaction: Over the past few years, the pace of change in the sector has accelerated, increasing the number of solutions and approaches involving multiple interdependent actors. The environment has become more fluid and the solutions less predictable. Ideas such as collective impact have reinforced the need for organizations to work together across boundaries, share information, and build on the lessons each has learned. As rules of interaction between social sector organizations evolve, evaluation approaches and methods must adapt to provide relevant, credible, and useful feedback to social sector stakeholders.

3. Proliferation of Digital Infrastructure: Technology adoption is widespread and a "digital infrastructure" (Bernholz, 2013) is now emerging for philanthropy and the social sector. The explosion of social media and the use of handheld devices have rapidly reduced the length and duration of the feedback cycle between funders, non-profits, and end beneficiaries. As more data are created and analyzed, evaluation must expand to allow social sector stakeholders to better understand the nature of social issues and maximize the use and effectiveness of data to solve social problems.

Source: Gopalakrishnan, Preskill & Lu, 2013

Recent theorists have been concerned with creating frameworks that are malleable and designed to consider evidence in the context of innovation. For example, Patton (2011) describes developmental evaluation, intentionally directed at innovation projects as "an extension of the summative/formative

repertoire focused on using evaluation within the process of innovation in which both the path and the destination are evolving, as a mechanism for bringing rigorous inquiry to development by being intentional about using data in a meaningful way to inform innovation in progress."

Viewed this way, evaluation is itself a dynamic, flexible, process that is specific to context, actively involving the various communities represented in the project in an iterative and cyclical process to determine the nature of the evaluation, in the context of this particular innovation.

THE NEXUS BETWEEN INNOVATION AND EVALUATION IN EDUCATION

Both innovation in education and evaluation of educational innovation are striving to address the demands and complexity of the changing societal and global landscape. When innovation and evaluation come together, they can provide a powerful iterative process for addressing new ideas and engaging in inquiry and learning, as complementary and intertwined processes. Innovations and innovators bring new ideas and rationales for intervention, as well as new approaches to capabilities, behaviour, and institutions. Evaluation and evaluators provide mechanisms for disciplined collection and use of evidence to investigate, support, challenge and guide innovation. The conjoint power of innovation and evaluation comes from the depth of thinking that emerges from the interface of generative ideas and appeal to evidence, in a deliberate process of learning for change. They do not work as separate processes but are connected and reciprocal, with close working relationships among the key players (innovators, funders, participants, facilitators and evaluators) to understand and influence the innovation as it unfolds.

Bringing innovation and evaluation together, like any relationship, involves all groups becoming familiar with the strengths and belief systems of the others and then determining how they will create synergy as they inhabit the same space and work towards the mutual goal of successful innovation. Because of their different backgrounds, innovators, funders and evaluators have a lot to learn about one another. They may have different views of the world. They may come to their relationship with a range of misconceptions about the values, beliefs, personalities and working style of the others. Making the relationship work is an ongoing process of listening, questioning, and trying to understand the different perspectives, in order to profit from each other's expertise and insights.

In the remainder of this working paper, we unpack how this interface works by identifying important issues at key decision points where evaluative thinking can enhance the innovation process. Although we have organised the paper sequentially around these decision points, in reality the process is more iterative and evaluative thinking can come into play at any time. The tidiness of the order belies the usual complexity.

Who are the Innovators?

It is not always easy to identify the "innovators" in an innovation initiative. There are those who conceive of or invent the ideas, those who lead and shepherd it into the public realm, those who support it directly or indirectly and those whose lives are changed as a result of it. Because innovation in education is a public undertaking, the work happens in real time on a public stage, with a wide range of participants. First off are the initiators of the innovation, with plans and ideas. They are accompanied by the funders (often policy makers) who provide the wherewithal for the innovation to proceed. There is also an active group of "innovation facilitators", who provide innovation leaders with support in the "process" of being innovative. Not to be overlooked are the participants in the innovation – the people who are expected to change what they think and do (often school based educators, students, parents and the community). These groups and individuals can bring very different perspectives and expectations to the innovation but they all have a role to play in innovation and will be interested in learning about and from the innovation as it evolves.

Box 6. Cognitive Biases

Katz & Dack describe a series of cognitive biases that can cloud how people respond to new information when they are monitoring change initiatives:

- don't think through all possibilities
- focus on confirming existing hypotheses, not challenging them
- pay too much attention to things that are vivid
- consider the information to be an exception or an anomaly
- hesitate to take action in a new direction
- don't want to expose vulnerabilities

Source: Katz & Dack, 2013

All of the individuals with a stake in the innovation will bring their own biases, seek confirmatory evidence about success and overlook disconfirming evidence. Box 6 gives some examples of human cognitive biases that can influence understanding. The strength of combining perspectives is that it creates the space for new insights and breakthroughs in learning, as the collective interprets and brings evaluative thinking to the evidence at hand.

It is not unusual for different members of the innovator group to imagine that they are able to conduct their own internal evaluation, given that the innovation will often move quickly and they possess intimate knowledge of the theory and the enactment of innovation on the ground. These conditions make them essential to the evaluation but their inherent biases can often stop them from seeing outside their existing perspectives.

Who are the Evaluators?

Evaluation is a large field that has been claimed by many institutions and groups as its own. Defining and regulating the field has been a major undertaking since evaluation became a significant part of public policy that goes beyond the scope of this paper. However, it is important to recognise that evaluation is often seen as the "country cousin" of research and people who have done research in academic settings may see themselves as qualified to become evaluators, without the extensive expertise and skill required. There are also many entrepreneurs who see a "ripe" market and establish themselves as evaluators. However, evaluation of innovation is not a simple process that can be done by anyone. It is technical, ethical and creative. Box 7 describes a set of evaluator competencies.

Box 7. Evaluator Competencies

The Canadian Evaluation Society (CES, ND) has prepared a set of competencies in five domains for accrediting evaluators, as a mechanism for specifying "...the background, knowledge, skills, and dispositions program evaluators need to achieve standards that constitute sound evaluations".

- Reflective Practice competencies focus on the fundamental norms and values underlying evaluation practice and awareness of one's evaluation expertise and needs for growth.
- Technical Practice competencies focus on the specialized aspects of evaluation, such as design, data collection, analysis, interpretation and reporting.
- Situational Practice competencies focus on the application of evaluative thinking in analyzing and attending to the unique interests, issues, and contextual circumstances in which evaluation skills are being applied.
- Management Practice competencies focus on the process of managing a project/evaluation, such as budgeting, coordinating resources and supervising. Interpersonal Practice competencies focus on people skills, such as communication, negotiation, conflict resolution, collaboration, and diversity.

Source: CES, ND

As we mentioned earlier, evaluation is itself a complex science with its own expertise associated with understanding context, collecting and analysing data and connecting evidence to theory and to intentions. Operating within the innovation space requires comfort working with a range of theoretical stances and world views, maintaining high integrity and ethical standards and being adaptable in order to focus and refocus the evaluative activities within the evolving innovation.

Defining the Innovation

One of the first tasks in bringing evaluative thinking to an innovation is getting a detailed and comprehensive description of what the developers intend and are doing – defining the particular innovation – its roots, goals, theoretical underpinnings and philosophy, in order to formulate an efficient and workable evaluation approach that will contribute to ongoing decisions about the innovation and satisfy accountability requirements along the way. This process is particularly important in innovation because of the wide range of possible definitions of what constitutes an innovation and the likelihood that innovations will change and morph over time. This description typically forms the foundation for tracking development, determining progress, and deciding what evidence is important to support and assess the success of the innovation. Innovation, by its nature, is not (and cannot be) a defined programme. As Gamble (2008) says:

> Initiatives that are innovative are often in a state of continuous development and adaptation, and they frequently unfold in a changing and unpredictable environment. This intentional effort to innovate is a kind of organizational exploration. The destination is often a notion rather than a crisp image, and the path forward may be unclear. Much is in flux: the framing of the issue can change, how the problem is conceptualized evolves and various approaches are likely to be tested. Adaptations are largely driven by new learning and by changes in participants, partners and context (p. 13).

Although innovation does not follow anyone's blueprint, it can be misleading to see everything as "emergent", without any conscious shaping (Mulgan & Leadbeater, 2013). Complex innovations do not appear from nothing or by chance. They are shaped and designed by individuals and groups as goals and a

set of ideas to achieve them that form the basis for action. As the constellations of innovation emerge they require equally complex and bespoke evaluation approaches that bring evidence to bear throughout the innovation to complement and inform decisions.

Defining an innovation is not a one-time exercise. Although innovation is emergent, it begins with a clear formulation of a vision. What are the problems (sometimes intractable) to be addressed and the desired or preferable state for the future? From there, the innovation team can move from these "big picture" goals to defining more specific goals and actions. The process continues with a "best shot" theory of action, making the underlying assumptions, frameworks and activities designed to progress the goals clear to the innovators themselves, and to stakeholders.

Developing an initial theory of action points the direction and defines the thinking underpinning the innovation. This theory of action is not static or immutable. It is expected to change and transform as part of the innovation process. The activities that occur within an innovation are likely to have modest and measureable goals that are intended to contribute to the larger goals and can provide some obvious checkpoints in describing places to pause, consider feedback from evidence, examine emerging patterns, reflect and rethink the plan. As the innovation progresses, the outcomes and processes that emerge are more fully understood.

Box 8. The Theory of Action Defined

A Theory of Action is an organisation's "theory", or story of how it will make change in the world. A theory is an explanation of why certain things happen. The fundamental component of a theory of action is a diagram that maps, at the most basic level:

- The intended impact on the world and how communities will be different because of the work – called long-term outcomes.

What changes or actions are necessary now, in order for these impacts to come about and what is the logic of how these actions will contribute to short term or intermediate outcomes along the way. **A Theory of Action is best suited to:**

- Understanding the rationale/logic behind why the organisation does what it does (how and why certain changes are expected to come about)
- Tracking contributions to complex change processes and outcomes shared among multiple actors.
- Testing assumptions and thereby gaining a realistic understanding of the work and resources needed.
- Learning and building evaluative thinking within an organisation.
- Achieving clarity and transparency of means and purposes among stakeholders.
- Providing a means to communicate the compelling story of the initiative to funders, board members and outside constituents.

Source: Adapted from Borgman-Arboleda, C. (ND)

Evaluators, at this stage, can help clarify the description of the innovation, assist in identifying the goals for building the theory of action, capture evidence about a range of questions "just in time" to

contribute to the decisions as they are being made, and make the twists and turns within an innovation more deliberate and visible. Box 9 is an example of a process for monitoring a theory of action within an innovation.

Box 9. Monitoring a Theory of Action

At the centre of a state innovation was the idea that bringing families into the centre of schools would create better connections between home and school for learners. It was hoped that the professionals would become more responsive to families' aspirations and ways of learning, particularly in ethnically-diverse communities. The end game was to improve student outcomes. Schools and community representatives were encouraged to join one of many networks established for the purpose. Network development was facilitated by an expert to help them to undertake an analysis of the current situation and stay focused.

Evaluators worked alongside the practitioners in the networks so they could monitor their own progress. The evaluators developed a series of evaluative probes to be used by practitioner evaluators within each network. By using the same measures, the findings were able to be collated across the state.

The first probe was simple. It involved identifying the position of those who attended network meetings. If this was to be a joint effort between schools and their communities, the expectation was that community representatives would be at network meetings. It was very easy in the early stage to identify that school professionals were essentially planning things for absent parent and community members. They were consulted but they weren't involved in the decision-making processes.

The second probe was designed to be used when networks were formulating their action plans. The practitioner evaluators were asked to identify the position of those attending their meetings and survey the participants to identify how the community was involved in developing the network plan. It was still evident that parents were consulted but remained on the periphery when it came to decision-making. The consistency of these findings allowed those involved in the design of the networks to work with the networks to find out why the professionals preferred to consult rather than to involve parents and communities directly and reconsider ways to redesign the interaction.

Evaluators can help make the theory of action explicit by asking questions about goals and about anticipated outcomes (questions like: "What do you expect from this? For whom? When? What might it look like? How does it work?") as prompts to support strategic thinking along the way about how to adapt and adjust the process and as opportunities to provide insights about successes and challenges.

There are no ground rules about how the innovation should evolve but it is important to routinely and rigorously revisit the goals and the theory of action, and to chronicle, document and monitor the progress and decisions over time, as a backdrop for understanding what works, how and under what circumstances.

Multiple Stakeholders; Multiple Interests

As we mentioned earlier, there are many players in any social innovation and the various stakeholders are more than passive observers of innovation. A traditional "hands-off" evaluation approach limits the utility of the evaluative thinking in moving the innovation forward. All of the groups who are involved in the innovation should also be part of the evaluative thinking process. These groups extend far beyond the core leaders of an innovation, often including communities, parents and particularly students themselves as key participants and decision-makers. When all of the groups who have a commitment to and interest in the innovation bring their diverse perspectives and intentions to the evaluation, the evaluation is likely to be more authentic and all stakeholders are more likely to understand, share, and support decisions (Cousins & Earl, 1992). Conversely, without stakeholder involvement and support, the learning from evaluation (and from the innovation itself) does not necessarily address important questions and the findings may be ignored, criticised, resisted, or even sabotaged. Evaluators need to be very aware of the wide range of

stakeholders and focus attention on increasing understanding and credibility by engaging key stakeholders in evaluative thinking as the innovation progresses. Box 10 is an example of the different expectations of multiple stakeholders.

Box 10. Who Cares to Know?

A foundation funded innovation was designed to address the intractable problem of secondary school students who were not engaged in schooling. The foundation was particularly keen to have immediate confirmation that students were off the street and in school. The network leaders were particularly focused on learning why the students were disengaged. Both are laudable outcomes but reflect different perspectives and beliefs.

Recognising Contexts

The situations or systems that form the context for innovation are rarely stable. Innovations are more likely to be embedded in local, political, social, historical and economic realities (advantages and constraints) that often change and influence every decision about the innovation. Evaluation work in cross-cultural context has raised awareness of the need to understand the cultures and norms that surround any innovation and to situate an innovation in its context (Rog, Fitzpatrick & Connor 2012). How is the innovation situated in relation to prevailing cultural norms and practices? Will there be ramifications for the wider economic, political, social and/or educational systems?

Given the unpredictable nature of context, innovators and evaluators must be culturally responsive and aware of the prevailing context and to changes that are occurring, through careful attention to the progress of the innovation "in situ". Evaluators and evaluation designs, in this case, need to be flexible and adapt to emergent and dynamic realities in complex environments, through decisions about methods and approaches that are respectful of prevailing norms and coherent with specific evaluation situations and questions. Many evaluation models and approaches call for working closely with the innovators and stakeholders to learn more about their preferences and interests, their values and their culture. This is a dramatic shift from an evaluation framework that is "hands off" to one in which evaluators become partners in the interpretation of complex systems rather than measurers of specific outcomes. Rather than tell a simple black and white story, they need to engage with the innovators and other stakeholders in all phases, beginning with determining what evidence is required and collecting the data. Stakeholder involvement cannot stop here, however, because what is usually most contested and context-dependent is interpretation of the evidence. Interpretation becomes an iterative process of capturing the viewpoints of all the key participants, deciding together what the data really mean in terms of the progress of the innovation and deciding how to monitor the process of what is actually occurring.

Identifying the Purpose(s) of Evaluation within the Innovation

In an earlier section, we described the wide array of purposes for evaluation. When there are many stakeholders and the context is complex, there are often many different purposes for evaluation and audiences for the insights that evaluation can provide. This is most obvious in the competing space between evaluation for accountability (external accountability) and evaluation for feedback and improvement (internal accountability and improvement). Both of these purposes are legitimate and require attention. The challenge is to get clarity about what questions will be addressed within the innovation/evaluation process and, with particular attention to when they are appropriate.

Defining and negotiating the purposes of evaluation is a critical and ongoing issue for attention by the key players that needs to be addressed directly, transparently and often. Even though policy makers and

funders support innovation, they may also be "risk-averse" and want a sense of security about their investment in innovation. This can lead them to overlook the fact that innovation, by definition, is risky and many innovations should "fail" otherwise they are likely to be safe, rather than truly innovative (Perrin, 2002). They may look for confirmation and success early, and in simple ways.

Innovators, on the other hand, may agree with Bill Gates, who said in his 2013 letter (Rosen, 2010, cited in Gates, 2013): "Without feedback from precise measurement, invention is doomed to be rare and erratic" (p.1).

They often accept that evaluative thinking has inherent value for development of the innovation, with the main purpose being to bring an ongoing evidentiary basis to the feedback, reflection and decision-making processes as successive iterations of innovations are planned, implemented and reviewed. However, they are wary of being expected to produce outcomes or looking for indicators of success too soon.

Negotiating this territory is not simple or straightforward and often requires the key players to spend considerable time clarifying their assumptions and establishing shared evaluation purposes, not once, but at the points in the innovation process where the expectations diverge.

GETTING ON WITH IT: APPROACHES AND METHODS

Evaluating innovation can usefully draw on the rich knowledge base already available from the evaluation domain. The evaluation literature has highlighted the importance of developing theories of action, identifying specific evaluation questions, developing methods to answer particular evaluation questions, bringing independence and rigour to the process and focusing on interpretation. Many of the principles underpinning this knowledge base are directly applicable to evaluating innovations, although they can require adaptation and new methodologies.

One of the main ways in which evaluating innovation is different from many traditional approaches is the need to be flexible and open to the emergent and dynamic reality of innovations. Decisions about methods and approaches will seldom be guided by and directly built on established evaluation models and the approaches will change throughout the innovation as explorations of data lead to particular interpretations with new issues to be explored.

Having flexibility and openness does not suggest an unplanned process. Rather it means a systematic and iterative process of both looking forward and looking back with intentionality. Looking forward involves formulating evaluative questions and collaboratively planning what evidence to collect to answer them. Looking back means considering evidence and deciding what analyses and consideration of the evidence are most useful to examine what is happening in the innovation, and the extent to which the innovation is progressing towards its goals or resulting in unintended consequences.

Framing Evaluation Questions

Evaluation questions are formulated by asking "What do we need to know?" These questions shape how the evaluation and the innovation will evolve, with new questions emerging as the innovation evolves. Thinking about what questions should take precedence means pausing and reflecting on the current status of the innovation to establish the kind of evidence that would be useful now and in the future. It is essential that the key players spend time identifying the right questions and return to this conversation regularly. In fact, these may be among the most important conversations because the questions that arise internally may be quite different from the ones that are posed for external accountability and the evaluation process must address both along the way.

Identifying the questions that warrant and require an appeal to evidence is part of the intellectual challenge of flexible and ongoing evaluation. Certainly, innovators, funders, policy makers and evaluators will have lots of ideas, but not all of them can be addressed in the evaluation. Lengrand et al., (2006) caution that there can be over elaborate efforts to understand everything that is going on at the expense of providing timely or even relevant information. The skill is to identify questions that provide the most relevant information, at the particular time, in the particular context, and balance stakeholder needs and the intended short, medium and long-term outcomes. Innovators, funders, policy makers and evaluators all play a part in shaping these pivotal decisions.

Typically there are two orders of questions related to what is happening in the short term and the bigger picture.

- What do we need to know now about what is happening for immediate feedback and decision-making? These questions are quite specific. They focus on the kinds of details that stakeholders agree are important to find out to give an indication that the innovation is moving in the right direction.

- What do we need to know about what is happening in terms of the bigger picture of intended outcomes for the innovation locally and systemically? These questions are more general and ensure that evidence is collected "along the way" that will contribute to understanding of the innovation and its impact more broadly.

Given the unpredictability of innovation, both orders of questions need to include the possibility of unintended consequences because it is sometimes these unintended consequences that prove to be the breakthrough, or the stumbling block.

Deciding the right questions is a science in itself. It requires careful consideration and negotiation of the needs and timelines of all the stakeholders. Otherwise, there's a risk that the evidence will not contribute to building new knowledge and moving the innovation forward.

Collecting Evidence, Timelines and Timing

The systematic collection of evidence provides the platform for answering the evaluation questions. There are innumerable texts on the important qualities of evidence, with the technical aspects beyond the scope of this paper.

Box 11. Just In Time Evidence

An innovative arts programme focused on increasing student engagement in their learning partnered artists with teachers in classrooms to deliver elements of the curriculum through arts-related activities. To gather data related to the students' engagement with their learning during these activities, the evaluators designed surveys to be completed by the students. These were delivered to the students through an "experience sampling" process (Hektner, Schmidt & Csikszentmihalyi, 2007) in which students received the survey via their cellphones at selected times during their classroom activities, coinciding with their participation in the arts-related activities and in their other classroom experiences, several times during the school year. This produced "just in time" data rather than asking students to give global statements related to their engagement, after the fact.

In essence, the evidence must be fit-for-purpose, of sufficient quality to form an accurate representation of the situation being evaluated and be available when decisions are being made. These criteria come together in different ways when deciding what evidence counts as fit-for-purpose, when to collect it and what level of attention needs to be given to quality. Sometimes there will be a need to engage in fast cycles of evolving strategies to react to opportunity windows and respond to external factors, and provide rapid feedback to inform immediate decisions or concerns. Some evidence will become a routine part of checking progress along the way. At points in time, there will be interest in stepping back and determining progress towards the ultimate goals. In situations where rapid feedback is needed to inform an imminent but low-stakes decision, demands for quality will be less than in situations where evidence is needed to inform a high-stakes decision or for accountability purposes. Similarly, requirements for representativeness will be more stringent.

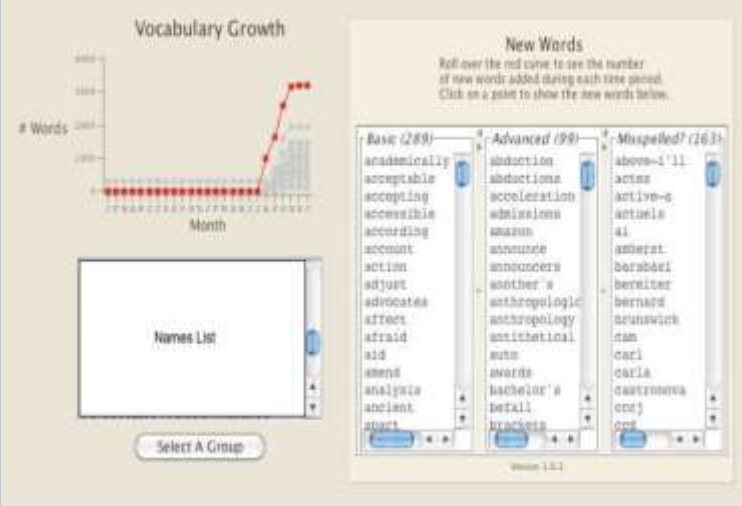
Over the years, the evaluation community has developed innumerable mechanisms for collecting information that are potentially applicable to evaluating innovation. These methods can range from document analysis; narrative, stories and vignettes; surveys, focus groups and interviews; to students' just-in-time responses using digital technologies and social media; and more recently the advent of analysis of "big data". Boxes 11 and 12 are examples of alternative ways to collect data. All have the potential to transform how organizations will engage in evaluative thinking in order to design, organize, and manage change. "Big data" makes it possible to capture huge amounts of information about employees, students, and operations through the:

> …millions of networked sensors that are being embedded in the physical world in devices such as mobile phones and automobiles, sensing, creating, and communicating data. Multimedia and individuals with smartphones and on social network sites will continue to fuel exponential growth. Big data—large pools of data that can be captured, communicated, aggregated, stored, and analyzed—is now part of every sector and function of the global economy. (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh & Byers, 2011, p.1)

All of these methods (and more) may have applicability to inform decisions within an innovation. Being "fit for purpose" means selecting methods that take into account the evaluation purposes and practicalities in a particular context. Methods for collecting evidence, in an imminent but low-stakes decision context, might best capitalise on the speed of digital technologies to sample relevant activities and participants' responses to them in real time. In the higher-stakes accountability context, multiple sources of evidence collected over a longer period of time are more appropriate.

Box 12. Analytic Tool in Knowledge Forum

Knowledge Forum, the learning platform developed by Marlene Scardamalia and her colleagues to promote knowledge building has a built in suite of analytic tools that allow teachers and students to analyse their work in ways that have not been possible before. For example, one tool analyses the words that students use in their posts and charts vocabulary growth.



Source: Marlene Scardamalia

Organising and Analysing the Evidence

Whenever innovators and evaluators are involved in the collection of evidence, the next step is one of deciding how to analyse and organise it so that it sheds light on the questions at hand. There are no standard analyses in evaluation contexts, and when it is associated with innovation, this is a collaborative task that is determined by the questions that have focused the investigation. And, not surprisingly, these questions are rarely simple. Innovation is multi-faceted and the hunches and quandaries that prompt an appeal to data are almost always complex and nuanced. Complex questions require complex analyses, geared to providing insights and clarifications.

The advent of computer programmes for analysis of quantitative data (everything from Survey Monkey to SPSS to big data analytics) gives many people ready access to statistical analysis. All too often the analysis becomes a "fishing expedition" where routine analyses and typical reports create over-confidence and drive simplistic understanding, based on stand-alone statistics, without the aid of context, a range of perspectives and other data to help assess the meaning. Statistics provide a numerical language in which ideas are quantified and organized using mathematical algorithms. Analysing quantitative data is fundamentally a thinking process, using statistics, to understand some phenomenon better. Statistics provide tools for developing and challenging hypotheses in a process of deepening understanding and raising questions.

As Abelson (1995), in his book "Statistics as Principled Argument" maintains: 'It is essential to argue flexibly and in detail for a particular case when you use statistics. Data analysis should not be pointlessly formal. It should make an interesting claim; it should tell a story that an informed audience will care about, and it should do so by intelligent interpretation of appropriate evidence from empirical measurements or observation" (p.2).

This approach requires a deep understanding of the way that statistics work, along with expertise and experience in using evidence and statistics as tools for thinking. Understanding statistics, data analysis and determining reasonable boundaries on interpretation are even more critical when considering "big data". The McKinsey Report on Big Data (Manyika, et al., 2011) highlights the problem: "Companies and other organizations and policy makers need to address considerable challenges if they are to capture the full potential of big data. A shortage of the analytical and managerial talent necessary to make the most of big data is a significant and pressing challenge and one that companies and policy makers can begin to address in the near term" (p. 3).

Analysing qualitative evidence requires its own expertise and experience. Studying education often means investigating the "everyday activities". When the reality of classroom or school life is captured through interviews, observations, blogs, videos, audiotapes, etc., the resulting evidence can be almost overwhelming. At first glance, it can often lead to the response "so what" or "what really matters here". It can be tempting to look for ideas or stories that confirm existing beliefs or to form categories of responses based on intuitive responses to the evidence. However, analysing qualitative data is not a clear-cut procedure. It is a very time consuming process that requires meticulous attention to detail through an active interrogation of the data in order to elucidate the significance of seemingly every day talk or behaviour and moves the conversation from specific incidents to larger themes and ideas. Box 13 gives a good example of the complexity of qualitative analysis. The evaluators' role is to bring organisers and analysis procedures to bear on the data to help the team approach the data with sensitivity to the culture in which it was produced, while at the same time, standing outside to see the larger themes, issues, ideas, actions.

Box 13. Using Vignettes To Understand Innovation

The OECD Innovation Learning Environments project in Austria is using vignette research as a data collection and analysis approach to gain access to learning experiences in the midst of the pedagogic situation, *in medias res.* As Westfall-Greiter& Schratz (2014) describe them, vignettes are a dense narrative of a poignant moment, capturing a lived experience as it occurs – a form of literary non-fiction that stems from researchers co-experiencing the lived experience of students by staying open and particularly attentive to elements such as atmosphere, facial and bodily expressions and tone of voice while co-experiencing. Vignettes are intended to initiate an experience in the reader which is as close as possible to that of the researcher's experience of the experience of students experiencing school. Vignettes captured in an innovative learning environment become an evaluation tool when the innovators engage with the vignettes, as a vehicle for vignette-driven conversations in which the evaluator facilitates the innovators in engaging with the vignette, peeling off and adding layers to come to enable exploration of "what I think I understand" and revealing layers of meaning and new insights about the impact of the innovation. These insights form the basis for such things as giving student teachers feedback after the lesson identifying pedagogical issues for planning.

MAKING SENSE OF IT ALL

Insights and observations that arise from looking at evidence can be interesting and compelling, but they really are not worth much unless they can be converted into useful knowledge that can inform the stakeholders and influence and guide the innovation. All too often, much attention and expense is attached to collecting evidence, and interpretation is hurried and superficial, when what matters are the insights that accrue from the evidence. Digging jewels out of the evidence is at the core of evaluative thinking – a process of considering the evidence, in relation to the questions that prompted its collection and engaging in careful inquiry and interpretation. This is a process of building and capturing knowledge, within the context of multiple stakeholders and multiple interests.

Having the evidence is the beginning. Analysing and organising it provides the next stage and structure for interpretation, but the real "work" of using evidence comes in the thinking process that occurs when all of the people who care about the innovation engage in making sense of what it means. Evidence does not have a particular "value valence" associated with it and data, by themselves, are benign, or at least neutral. It is the interaction between evidence and people that results in decisions that creates harm or beneficial effects. Whether we view this process as organised and strategic, or a consequence of an unpredictable confluence of people, problems, data and decision opportunities (March & Olsen, 1986) makes a difference. However, both strategic and non-rational models assume that using data is a thinking activity that draws on personal views but also on capturing and organising ideas in some systematic way, turning the information into meaningful actions and making the interpretation public and transparent (Senge, 1990).

Interpretation as Building Knowledge

When evaluation operates as a separate process from innovation, data analysis and interpretation are typically situated in the domain of the evaluator. When it is situated as part of innovation, interpretation becomes a shared and ongoing process of inquiry, using the discipline of appealing to evidence to deepen understanding within and about the innovation. It occurs through a cycle of collaborative knowledge building to improve ideas:

> "Knowledge building is the creation or modification of public knowledge that lives "in the world" and is available to be worked on and used by other people. That goal is to advance the frontiers of knowledge as they perceive them.
>
> […]
>
> In knowledge building, ideas are treated as real things, as objects of inquiry and improvement in their own right. Knowledge building environments enable ideas to get out into the world and onto a path of continual improvement. This means not only preserving them but making them available to the whole community in a form that allows them to be discussed, interconnected, revised, and superseded" (Scardamalia & Bereiter, 2003).

This notion of knowledge building is a sophisticated one that advances many notions about how learning happens for individuals and how it is transformed in social contexts to become part of what a culture holds as knowledge.

As Nonaka, Toyama & Konno (2000) indicate: "Knowledge is dynamic, since it is created in social interactions amongst individuals and organisations. Knowledge is context specific, as it depends on a particular time and space. Without being put into context, it is just information, not knowledge.

Information becomes knowledge when it is interpreted by individuals and given a context and anchored in the beliefs and commitments of individuals" (p.7).

Knowledge building can happen within individual projects and across projects. The process is the same; only the data to be considered and the players differ. Within projects, it occurs whenever the team are trying to achieve greater understanding and clarity about how things are developing and changing, as they evolve. At points in time, it may be worthwhile to look across a number of projects to synthesize the learning into a coherent and defensible set of interpretations, investigate trends and look for synergies in order to address large and important questions that can be useful to all and to inform the larger field of innovation.

In any complex situation what is seen to be important depends on what is valued, which in turn, is influenced by the various roles in the innovation/evaluation. Multiple stakeholders will attend to the evidence most relevant to them and they will interpret it through different lenses. Different players will bring different theories to the table and understand the evidence in different ways. This is a major advantage of the merger between innovation and evaluation because including the multiple perspectives pushes thinking and challenges "taken for granted" ideas. Funders and policy makers will inevitably want to know the extent to which the innovation is progressing towards the intended outcomes, what is working as intended and what is not. Innovators often want to know similar things but from their perspective are more likely to be focused on the shorter term goals and how evidence can inform the next decision. Practitioners involved in implementing the innovation usually want to know more directly about how the innovation is impacting on the day-to-day aspects of their work and likely immediate goals. Making meaning of evidence together is a delicate balance of listening to different perspectives, negotiating the issues for consideration and addressing questions of interest in order to deepen and extend the understanding and engagement for everyone. The example in Box 14 shows the invaluable contribution of students to the evaluative thinking and interpretation process.

Box 14. Multiple Perspectives in Interpretation

In a large secondary school in a very multi-cultural area, a group of students were concerned that the programmes being offered in the school for the community were not serving the community needs. They approached a steering committee of school and community agencies and were asked to join the committee. The expanded committee decided to hire a group of students representing the diverse nature of the school to interview parents and community members within their own language and culture group about the best use of the school facilities. The committee formulated questions; the students conducted the interviews in their home languages and came together as a group to translate, transcribe and create initial themes for discussion by the steering committee. The steering committee morphed into an interpretation and planning committee to prepare a facilities use plan. The plan was deeply influenced by the insights that emerged from the data and from the students' intimate understanding of the needs of the many different cultural groups.

Intentional knowledge building experiences can interrupt the status quo and create the space for multiple alternative views to emerge, through discussing what the evidence means, generating hypotheses, and establishing a range of possible interpretations. This process sets the stage for new knowledge to surface as the participants encounter new ideas or discover that ideas they have held as "truth" do not hold up under scrutiny and they use the recognition as an opportunity to rethink what they know and what they do (Earl & Katz, 2010).

The example in Box 15 describes a useful inquiry approach for systematic analysis of the situation evaluative thinking and knowledge building. In the inquiry process, the players use the evaluation evidence, along with the tacit knowledge and experiences of the participants as the fodder for inquiry and

knowledge building, as innovators and stakeholders consider complex educational issues from a range of vantage points.

Box 15. Dynamic Spiral Critical For Knowledge Creation And Sharing

- ***Creating Context:*** Through explicitly creating a context, the issues being investigated are connected with deep principles of the knowledge domain in question, and anchored in authentic, practical, and complex problems of the external world, or issues that the participants generally care about.

- ***Engaging in Question-Driven Inquiry:*** An essential aspect of progressive inquiry is generating one's own problems and questions to guide the inquiry; without questions generated by the participants themselves there cannot be a genuine process of inquiry. Questions that arise from one's own need to understand have a special value in the process of inquiry.

- ***Generating Working Theories:*** Construction of their own working theories guides inquirers to systematically use their background knowledge and become aware of their presuppositions. Progressive inquiry is aimed at the explication of these intuitive ideas.

- ***Critical Evaluation:*** Critical evaluation underscores the need to assess the strengths and weaknesses of the tentative theories (explanations) produced so as to direct and regulate the evolution of inquiry. It is essential to focus on constructively evaluating the advancement of the inquiry process itself, rather than simply an end result.

- ***Searching for New Information:*** Searching for and working with "research" is necessary for deepening one's understanding. New information can come from literature, consulting experts, or conducting one's own explorations. Explicit comparison of the intuitive working theories with the well-established ones makes the limitations of individual and collective knowledge apparent.

- ***Engagement in Deepening Inquiry:*** A critical condition for progress is that inquirers focus on improving their ideas by generating more specific questions and searching for new information. The dynamic nature of inquiry is supported by the fact that generating working theories and obtaining new research knowledge makes new guiding questions accessible.

- ***Shared Expertise:*** The agent of knowledge creation is not an isolated individual but an individual embedded in a community, or even the community itself. All of the preceding aspects of inquiry can be shared with other inquirers. Inquiry can advance substantially through relying on socially distributed cognitive resources and collaborative efforts to enhance shared understanding.

Source: Hakkarainen et al., 2004

The tension between those who see evaluation as a mechanism for external accountability and those who see it as an embedded part of the innovative process is often one of timing, not of kind. Engaging in routine evaluative thinking allows everyone with a stake in the innovation to gain a better understanding of the progress of the innovation as it develops and the extent to which it is meeting its intended or evolving goals. When the key players work together to clarify goals, recognise progress in smaller steps, and negotiate the nature of acceptable evidence, they are more likely to understand the process and become advocates for both the innovation and the evaluative thinking that is taking place. The timetable for gathering evidence will continue to be flexible, with all parties engaged in deciding what they want (and need) to know, what they will accept as evidence, and when it is appropriate and possible to try to address their purposes.

When the knowledge building spans a number of innovations, the process is somewhat like a meta-interpretation, where the interpreters work with information from a number of innovations to identify, describe and codify the collective learning so that it becomes explicit and can be shared. The TLRP project described in Box 16 is an example of a model for aggregating ideas across projects.

Box 16. Knowledge Building Across Projects

Although it is not evaluation, per se, the Teacher Learning and Research Project (TLRP) in England used a knowledge building process to learn from individual projects that were focused on different research questions and utilised a range of methods and theoretical resources. They used a cross-programme thematic seminar series and task. This process generated considerable discussion and the identification of a set of principles that transcended each of the individual studies to be communicated to wider audiences.

Source: James & Pollard, 2010

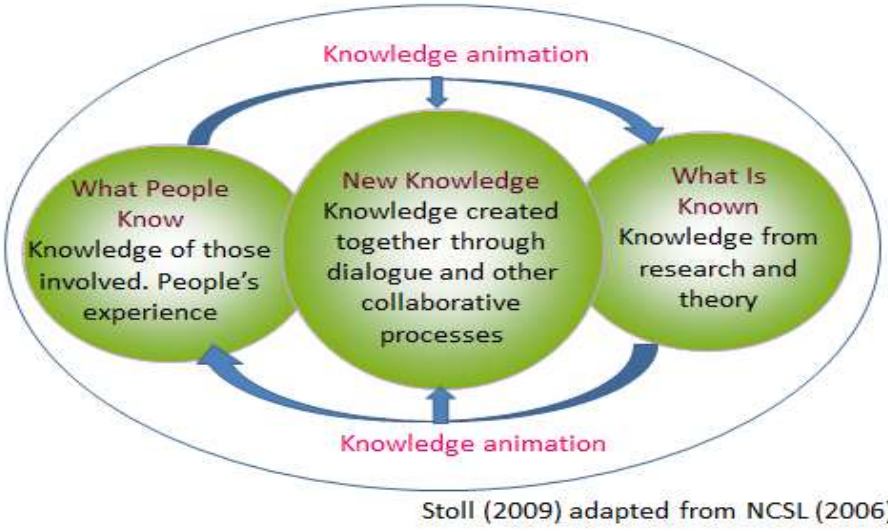
The nature of knowledge building is rarely predictable and the evidence from evaluation will be processed and interpreted in unexpected ways. It is as often the gems, surprises, outliers and snippets of evidence that give insight into what is happening for whom and why. Sometimes differences in interpretation will persist because of various players' competing theories. These differences need to be framed as theories in competition, not whether one is right or wrong. Adjudication of these theories comes from further testing, and possibly collecting more evidence (Timperley & Parr, 2005). Learning and change arises from this deep inquiry process. Through this iterative process, the team engages in interpretation, with inquiry and evaluative thinking as a way of doing business, rather than a discrete event.

Capturing and Mobilising the New Knowledge

Having worked through the interpretive and knowledge-building process with particular stakeholder groups, the learning needs to be visible and accessible to others both within and beyond the innovation, through processes of capturing it in some accessible and retrievable form (print, audio recording, video, translations, etc.) and intentional mechanisms for sharing the learning with people (internally and externally) who are not part of these regular episodes of making meaning.

Sharing and mobilising knowledge means creating new learning environments locally and beyond for others to engage with the ideas. This does not mean "going out and telling others" what has worked or not worked. Rather it means engaging in a wider process of evidence-informed inquiry with those not involved in the original interpretive knowledge building activities to activate existing knowledge, infuse the new knowledge from the innovation and socially construct new understanding.

Box 17. Knowledge Animation



There is an expanding literature on sharing knowledge and connecting it to practice and to policy that has grown from a range of disciplines and has evolved into a distinct scholarly field. Depending on the context, it is variably referred to as knowledge transfer, knowledge management, knowledge translation, knowledge mobilisation, knowledge animation or knowledge diffusion. Box 17 is a graphic presentation of knowledge animation as described by Stoll (2009):

> …a social process by which practitioners and policy makers make learning connections when engaging with research findings. Knowledge animation is about helping people to learn and use ideas generated elsewhere, and through this process create their own knowledge. It's concerned with finding ways of making knowledge accessible and mobile so that it stimulates dialogue that challenges people's thinking, promotes new understanding and helps them generate new knowledge that will enhance their practice and policy. (p. 1)

As is obvious in this paper, knowledge mobilisation in innovation is not a "one-off" activity. Knowledge mobilisation includes aggregation and packaging of explicit knowledge but it also requires knowledge work – direct and active engagement with the knowledge in ways that make tacit knowledge visible and allow the participants to draw on both tacit and explicit knowledge as they share the learning from their collective work. It is a deliberate process of stopping at various points in the innovation and evaluation process and asking, "What do we think we know that should be shared and checked?" "Who should we involve?" and "What process is best suited to sharing this stage of the learning?"

In essence, an effective knowledge mobilisation or animation process embodies the principles associated with any effective learning environment, as they have been described by the OECD Innovative Learning Environments initiative (OECD, 2013), things like active engagement, self-regulation, social nature of learning, attention to tacit knowledge, challenge, clarity of expectations, feedback, and horizontal connectedness.

Sustainability

As innovation progresses, there is always concern about whether it will be sustained. Sustainability, however, is another of those concepts that has multiple meanings in different contexts. According to the Oxford Dictionary, the definition is: "able to be maintained at a certain rate or level". The Merriam Webster Dictionary says: "able to last or continue for a long time".

So, what does it mean in the context of educational innovation? In our view, considering sustainability means addressing the questions: Sustain what? For what?

Even within the field of innovation, there are different expectations for sustainability. Christensen (1997) described innovations as sustaining or disruptive. "Sustaining innovations" are innovations that create improvements and can be incorporated into the existing practices of organizations. "Disruptive innovations" require different models to succeed, models that require new mindsets, and skillsets and cannot be integrated into existing structures and cultures. Because innovation in practice is local, each innovation needs to determine what sustainability means in the particular context.

In evaluation of educational innovation there is a wide range of conceptualisations of what sustainability might mean and are described below.

- **Sustainability as Fidelity to Programmes**: This notion of sustainability, which is closely connected to a search for methods and approaches that can be "scaled-up". Once something has been shown to work, the sustainability issue is fidelity to and spread or breadth of a defined programme, with serious attention to ensuring that resources are available, teachers are trained and that implementation follows the original plan – the closer, the better.

- **Sustainability as Maintenance of the integrity of Key Processes or Routines**: Sustainability can be considered as maintenance of things like routines, approaches, materials, etc. that are focused on a specific area. This often is associated with situations where a short-term influx of resources, professional development opportunities and other forms of assistance have been available and then withdrawn. The sustainability question becomes one of ascertaining if the injection of these resources leaves a footprint behind them. Are elements of the innovation visible after the passage of time and have they been transferred or passed on to others?

- **Sustainability of Theories, Principles or Ideas**: When innovation is not a single "thing" or project and is more focused on changing cultures or ways of thinking towards continuous improvement, the sustainability issues are likely to be related to ongoing improvement that is influencing success for students. This conception of sustainability relates closely to the ideas about capturing and mobilising new knowledge in the previous section. It takes what was learned within a particular context and connects it to underlying principles through unpacking the big ideas and using them to move to the next iteration of the innovation. Scardamalia & Berieter (2003) refer to this process as developing deeper learning in ways that advances the frontiers of knowledge. It may go beyond a single innovation as those involved tease out the big ideas and improve theories about approaches to innovation. Principled knowledge is the kind of knowledge that can be transferred and further developed across contexts (Pellegrino, 2006) whether it is an innovation planned for the future or for others that are developing alongside.

- **Sustainability as a Habit of Mind**: Although each innovation is local, there is a larger interest in the nature of innovation and sustainability of innovation as a "habit of mind" as described in Box 18 – a way of thinking and action that permeates the culture of the people involved and continues as a cyclical process of innovation that extends across projects and across time.

So where does evaluation fit in this picture? In the understanding of sustainability as programme fidelity, evaluation was focused on identifying whether the programme worked or not for the purpose of deciding whether it should be continued. This traditional summative evaluation was typically outcomes focused. Formative evaluation provided evidence along the way, as the programme was being developed in order to identify important processes and approaches essential to sustainability and anticipated programme spread.

Box 18. Habits of Mind

Habits of Mind is knowing how to behave intelligently when you DON'T know the answer. It means having a disposition toward behaving intelligently when confronted with problems, the answers to which are not immediately known: dichotomies, dilemmas, enigmas and uncertainties.

Source: Costa & Kallick (2000)

As the more principled view of sustainability has emerged, along with more developmental views of evaluation, evaluative thinking has become an integral part of identifying the big ideas within an innovation and linking them to evidence of how they have played out and what might or might not contribute to sustainability of process and outcomes that can contribute to system capacity as described by Fullan in Box 19.

Box 19. Sustainability

Sustainability is the capacity of a system to engage in the complexities of continuous improvement consistent with deep values of human purpose.

Source: Fullan (2004)

INNOVATION AND EVALUATION: SYNERGIES AND TENSIONS

Throughout this paper, we have tried to convey an image of a nexus at which innovative thinking, disciplined by evaluative thinking, work together as an intellectual disposition for guiding innovation in social contexts. Together they operate as a composite of imaginative innovation and reasoned evaluative critique to create new knowledge. This knowledge is developed through bringing together perspectives from innovation leaders, innovation participants, funders and evaluators as they struggle to move from ideas to actions and to bring relevant evidence to the table to address emerging questions. To make this combination work requires discipline in the innovation and flexibility in the evaluation. Cautions apply to each and there are inevitable tensions. One of the risks on the innovation side is the potential to slide from complexity to chaos. Evaluation designs can cope with complexity but not chaos, other than possibly to document its dysfunction. Quinn-Patton writes about how social innovations, like education, can become so complex at the extreme that they do not serve their purposes: "At the extreme of this continuum of uncertainty is chaos, intense conflict among the key stakeholders and extreme uncertainly about what to do to achieve desired outcomes" (Patton, 2011, p 93).

On the evaluation side, there is a potential for the evaluation to detract from, rather than to enhance an innovation, in ways best expressed by Lengrand et al., (2006): "Evaluation is not a panacea...It is a vital element of this [the innovation] process, but one element among others. It is not an end in itself… Especially if carried out in a rigid fashion, with strict adherence to a narrow set of targets and indicators, the evaluation process can come to limit the conduct of programmes" (p. 40).

There are processes that can help to reduce, and sometimes avoid these tensions. We now highlight some that we believe are useful and important – openness to improving ideas, being pragmatic, negotiating and re-negotiating meaning and relationships.

Being Open to Improving Ideas

Evaluative thinking within innovation assumes that continuous learning is intentionally embedded into the innovation process. As we described earlier, knowledge building is a collaborative process designed to improve ideas through additions, refinements, and adaptations. It requires space, time, and resources for dialogue, collecting evidence, reflection, posing questions, identifying and challenging values, beliefs and assumptions, and instituting feedback loops. As Preskill & Beer (2012) describe it:

> Those who are interested and willing to experiment with social innovations must be willing to take risks and accept missteps or failure. They must be willing to live with uncertainty and acknowledge that their plans, regardless of how well laid out, will likely shift as the circumstances around them change. With uncertainty and unpredictability comes an even greater need for strategic learning as an innovation is conceptualized, designed, and implemented. (p. 3)

Although most funders and innovators see themselves as open and flexible, many innovation initiatives are bounded by contracts and expectations that keep change and adaptation to a minimum. Innovations may also be the "brainchild" of individuals or groups who are strong advocates for particular directions, follow their instincts and are disdainful of appealing to evidence as a strategy for change and have little interest in alternative perspectives. Evaluative thinking is not consistent with only considering positive evidence or with approaches and direction that are already decided or entrenched to produce "good news".

At the same time, traditional evaluation models often do not lend themselves to the changeable and adaptive nature of working within innovations because they are not sufficiently flexible to capture the

ongoing development in innovation, nor are they responsive to the immediate information needs of innovators and policy makers.

> In many cases, traditional evaluation approaches fail to meet the fast-paced information needs of …decision makers and innovators in the midst of complex social change efforts. At worst, the application of traditional evaluation approaches to innovative change initiatives may even decrease the likelihood of success because they restrict implementers to pre-set plans that lose their relevance as the initiative unfolds. (Preskill & Beer, 2012, p. 1)

Although it seems simple, one of the conditions that we have identified is having an open, inquiry-based view of the innovation and the role of evaluation within it. The process of strategic learning that integrates evidence and evaluative thinking into innovation is a melding of innovators who engage in evaluative thinking and evaluators who engage in innovative thinking. This can even mean that funders and innovators need to negotiate among themselves to select evaluators carefully to ensure a match of purposes and perspectives. Evaluators may choose to leave the team if they feel they are being used to promote rather than improve the innovation.

Being Pragmatic

Clearly, introducing evaluative thinking into innovation is not a trivial undertaking. The role for evaluators can shift from periodic involvement, often on the margins of the innovation, to being central players in the innovation planning and developing evaluative thinking. Evaluator cost and availability are likely to be issues. It is rarely practical or cost-effective to have the evaluator join as a full member of the innovation team. On the other hand, it is important to allocate sufficient resources to engage the evaluator for key episodes of evaluative thinking, to share in leading particular processes related to evaluative activities, to provide evaluative expertise and to build evaluative capacity.

These occasions will vary within and across innovations and require that the evaluator be available during the different phases of the innovation. In the early stages, evaluators have a role in sketching the theory of action, identifying the important questions, determining methods to gather evidence and collecting it in partnership with the innovators and organising the data to be accessible for interpretation. A major role is leading the discussion around what the evidence means.

The need for just-in-time evidence, together with considered interpretation to inform ongoing decisions, places high demands on the evaluators' flexibility and availability. These roles have resourcing implications that cannot be fully anticipated or planned for because of the unpredictable nature of innovation and the integral nature of the approaches to evaluation within it.

Being pragmatic may also mean leaving some aspects of the innovation outside of the realm of evaluation. Focusing on key elements of the theory of action that are agreed by all stakeholders to be pivotal in achieving the purpose of the innovation may provide the greatest return on the evaluators' involvement. Not everything can or should be evaluated.

An alternative is to use the work of the evaluators in a few key areas to help innovators bring evaluative thinking into other areas, thus permeating evaluative thinking throughout the innovation, as is exemplified in Box 20. Becoming evaluative thinkers does not happen automatically. It must be an explicit purpose and be systematically addressed, with periodic involvement of evaluators, throughout the evaluation to ensure that innovators understand and internalise evaluative thinking and see how it can be applied across contexts.

Box 20. Embedding Evaluative Thinking

In a New Zealand initial teacher education programme, the goal was for graduates to develop adaptive expertise so they had the meta-cognitive and self-regulated learning skills to continue to learn as they entered the teaching profession. All teaching approaches and routines were to be introduced within this framing. The limited availability of funding for an evaluation led to the decision that the evaluators would focus on collecting evidence about the effectiveness of the teacher educators' feedback practices in developing the student teachers' adaptive expertise because major resources were going into this area and it was considered the optimal learning opportunity. The teacher educators were closely involved in designing the data collection of audio recordings of feedback practice and interviewing the student teachers. They also spent considerable time discussing the meaning of the data and the implications for future practice.

The teacher educators reported that the exercise had been so valuable in developing both their understanding of how to develop adaptive expertise and what it means to engage in evaluative thinking around their practice that they decided to investigate the effectiveness of the way they presented new ideas to the student teachers and follow-up group discussions. This time round, the evaluator acted more as a guide to answer their questions and providing prompts for engaging a range of views in the interpretation process.

Negotiating, Renegotiating and Renegotiating

As we have identified, the blend of disciplined innovation and evaluative thinking is essential for moving innovation forward and capitalising on the opportunities for knowledge building. The power of innovation often lies in crossing these boundaries:

> Interdisciplinary innovation arises from the positive effects that result when stepping across the social boundaries that we structure knowledge by. Those boundaries include academic disciplines, government departments, companies' internal functions, companies and sectors, and the boundaries between these domains. In the knowledge economy, it is often the case that the right knowledge to solve a problem is in a different place to the problem itself, so interdisciplinary innovation is an essential tool for the future. (Blackwell, Wilson, Street, Boulton, & Knell, 2009, p.1)

The challenge in the co-habitation of innovation and evaluation will be finding the time and space where those involved can come to a shared valuing of both the processes and the knowledge that emerges from their mutual efforts. Creating new knowledge involves negotiating, renegotiating and then again renegotiating meaning as problems become redefined and new evidence is brought to the table for consideration and interpretation. Negotiating meaning and creating new knowledge requires the mix of expertise of innovation leaders, innovation participants, funders and evaluators. Funders bring their unique view of where the innovation fits in the larger picture. Innovators bring a vision of innovation possibilities, expertise in education, and change management. Innovation participants know their own context with its strengths and limitations. Evaluators bring the technical expertise involved in defining questions and developing appropriate methodologies to answer them, together with analytical and interpretive expertise involved in the systematic reasoning from evidence.

The power of negotiating meaning in the nexus often lies in blurring the space between them, with innovators taking on evaluation roles and evaluators becoming involved in the innovation. This creates a space for productive dissonance as their backgrounds and beliefs bump into one another. As Blackwell, et al., (2009) have found, "it is often believed that people with different training have difficulty communicating because they have learned different specialist languages. But we have found a bigger

problem – that they are actually trying to achieve different things. Different disciplines often have different core values...In order for a new interdisciplinary team to become effective, that team must develop shared values and culture" (p.4).

The values and culture different players bring to any situation are sometimes referred to as personal theories. They consist of beliefs and values, together with the practices that follow from them and the outcomes considered desirable. In any situation, differing personal theories between innovation facilitators and participants, stakeholders and evaluators should be expected and accepted with the differences framed as different theories in competition with one another and in need of negotiation, rather than one being right and another wrong. In a theory competition approach all parties need to take responsibility for engaging with others' theories with mutual understanding at the nexus fundamental to successful negotiation of meaning and the development of new theories that become shared new knowledge (Senge, 1995; Timperley & Parr, 2005).

CAPACITY FOR EVALUATIVE THINKING: PULLING IT ALL TOGETHER IN THE NEXUS

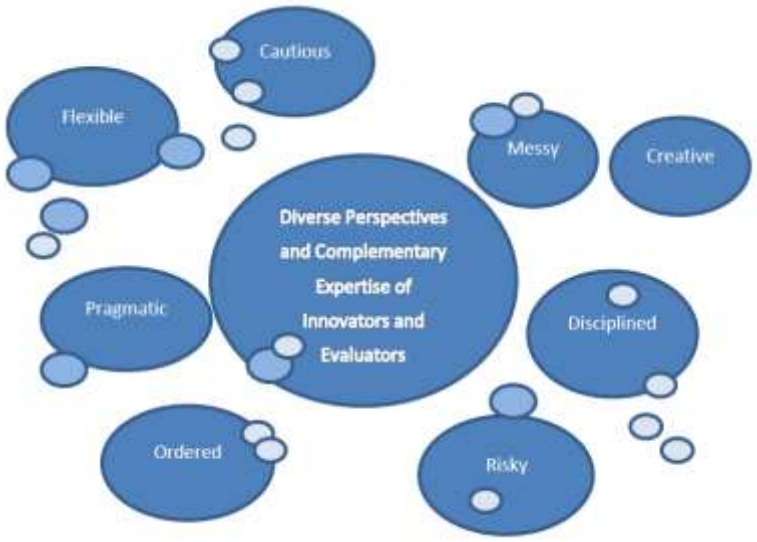
Box 21. Capacity

Capacity is a quality of people or organisations that allows them routinely to learn from the world around them and apply their learning to new and sometimes novel situations so that they continue on a path toward their goals, even though the context is ever-changing. It also helps them continuously to improve learning and progress at all levels.

Source: Stoll & Earl (2003).

Successful innovation evolves through building knowledge at the nexus between innovation and evaluation. The interaction of diverse perspectives and complementary expertise in this space creates the opportunities for evaluative thinking in innovation. The nexus is meant to be a place of creative dissonance and intentional interruption of "taken for granted" ideas using evidence; capitalising on a mix of expertise, theories about how the world works and the pragmatics of what is possible in a particular context. It can be messy and ordered, risky and disciplined. Negotiating this complexity is new territory for innovators and evaluators that requires having the capacity, as described in Box 21, for evaluative thinking to engage with the messiness of trying out new things, while bringing sufficient discipline and flexibility to the process. Embedding evaluation in innovation blends disciplined innovation and evaluative thinking at the nexus. In this space creativity combines with evidence and reasoned argument as innovators and evaluators work together to interpret evidence, challenge ideas and clarify their assumptions, beliefs and directions and the stage is set for new theories and improving ideas to move innovation forward, as a sustained process.

Box 22. Diverse Perspectives in Creative Nexus



REFERENCES

Abelson, R. (1995), *Statistics as Principled Argument*, Lawrence Erlbaum Associates Inc., Hillside, New Jersey.

American Evaluation Association (ND), www.eval.org/p/cm/ld/fid=1(accessed 11 November 2013).

Barber, M., Donnelly, K., and Rizvi, S. (2012), *Innovation: The Atlantic, The Pacific, Global Leadership And The Future Of Education*, Institute for Public Policy Research, London.

Bennett, G. and N. Jessani (2011), *The Knowledge Translation Toolkit, Bridging the Know–Do Gap: A Resource for Researchers*, International Development Research Centre, http://ajpp-online.org/resources/downloads/04-TheKnowledgeTranslationToolkit.pdf.

Bernholz, L. (2011), "Evaluating Innovation", MacArthur Series on Field Building, www.scribd.com/doc/57548064/Evaluating-Innovation (Accessed 06 October 2013).

Blackwell, A., et al. (2009), "Radical innovation: crossing knowledge boundaries with interdisciplinary teams", Technical Report No. 760, University of Cambridge Computer Laboratory, www.cl.cam.ac.uk/techreports/UCAM-CL-TR-760.pdf.

Borgman-Arboleda, C. (ND), Developing Your Theory of Action: A Facilitation Guide, Evaluation Collaborative www.seachangecop.org/node/1257.

Canadian Evaluation Society (ND), Evaluator Competencies, evaluationcanada.ca/txt/2_competencies_cdn_evaluation_practice.pdf.

Christensen, C. (1997), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business School Press, Boston, MA.

Cousins, B. and L. Earl (1992), "The case for participatory evaluation", *Educational Evaluation and Policy Analysis*, vol. 14, (4), Winter.

Costa, A. and B. Kallick (Eds.) (2000), *Habits of Mind: A Developmental Series*, Association for Supervision and Curriculum Development, Alexandria, VA.

Drucker, P. (1985), "The Discipline of Innovation", Reprinted 1998 Harvard Business Review, http://ogsp.typepad.com/focus_or_die_ogsp/files/drucker_1985_the_discipline_of_innovation.pdf .

Earl, L. and S. Katz (2010), "Creating a Culture of Inquiry: Harnessing Data for Professional Learning", in Blankstein, A., Houston, P. and Cole, R. (eds.) *Data Enhanced Leadership Series: The Soul Of Educational Leadership*, Corwin Press, Thousand Oaks.

European Commission (1995), Green Paper on Innovation, http://europa.eu/documents/comm/green_papers/pdf/com95_688_en.pdf .

Fullan, M. (2004), *Leadership and Sustainability: System Thinkers in Action*, Corwin Press, Thousand Oaks.

Gamble, J. (2008), *A Developmental Evaluation Primer*, The J.W. McConnell Family Foundation, www.mcconnellfoundation.ca/en/resources/publication/a-developmental-evaluation-primer (accessed November 2013).

Gates (2013), Why Does Measurement Matter?, Annual Letter, www.gatesfoundation.org/Who-We-Are/Resources-and-Media/Annual-Letters-List/Annual-Letter-2013.

Gopalakrishnan, S, Preskill, H. Lu, S. (2013), Next Generation Evaluation: Embracing Complexity, Connectivity, and Change, www.ssireview.org/nextgenevaluation.

Hakkarainen, K., Palonen, T., Paavola, S., and E. Lehtinen (2004), *Communities of networked expertise: Professional and educational perspectives*, Elsevier, Amsterdam.

Hannon, V. (2009), "'Only Connect!': A new paradigm for learning innovation in the 21st century", Occasional Paper No. 112, Centre for Strategic Education.

Hektner, J., J. Schmidt and M. Csikszentmihalyi (2007), *Experience Sampling Method: Measuring the Quality of Everyday Life*, Sage Publications Inc., Thousand Oaks.

James, M. and A. Pollard (2011), "TLRP's ten principles for effective pedagogy: rationale, development, evidence, argument and impact", in *Special Issue: Principles for Effective Pedagogy: International responses to evidence from the UK's Teaching and Learning Research Programme Research Papers in Education*, Vol. 26/3.

Joint Committee on Standards for Educational Evaluation (2013), Program Evaluation Standards, Sage Publishing.

Katz, S. and L. Dack (2013), *Intentional Interruption: Breaking Down Learning Barriers to Transform Professional Practice*, Corwin Press, Thousand Oaks, CA.

Lengrand, L., and Associés (2006), *Smart innovation: A practical guide to evaluating innovation programmes*, DG Enterprise and Industry, ECSC-EC-EAEC, Brussels-Luxembourg, http://admin.interact-eu.net/downloads/2395/Smart%2520innovation:%2520A%2520practical%2520Guide%2520to%2520Evaluating%2520Innovation%2520Programmes.pdf .

Lyn, L. (1997), "Innovation And The Public Interest", in Altshuler & Behn, D. (eds) *Innovation Challenges, Opportunities, and Dilemmas in American Government*, Brookings Institute.

Manyika, J., et al. (2011), Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute.

Mulgan, G. (2007), "Social Innovation: What it is, why it matters and how it can be accelerated", Working Paper, Skoll Centre for Social Entrepreneurship, Oxford Business School. www.politicadeinnovacionsocial.co/documents/250640/254377/Social-Innovation-what-it-is-why-it-matters-how-it-can-be-accelerated-March-2007.pdf/4aa7263e-571c-48c5-a14b-493b66c064a0.

Mulgan, G. and C. Leadbeater (2013), "Systems Innovation", Discussion Paper, NESTA, London.

Nonaka, I., Toyama, R. and N. Konno (2000), "SECI, Ba, and leadership: a unified model of dynamic knowledge creation", *Long Range Planning*, Vol. 33, No. 1, pp. 5-34.

Northwest Public Health Center (ND), Six Steps of Program Evaluation, Northwest Public Health Center, University of Washington, www.nwcphp.org/evaluation/tools-resources/program-evaluation-tips.

OECD (2013), *Innovative Learning Environments*, Educational Research and Innovation, OECD Publishing, Paris, www.oecd-ilibrary.org/education/innovative-learning-environments_9789264203488-en.

OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, www.oecd-ilibrary.org/content/book/9789264190658-en.

Patton, M. Q. (2011), *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Press, New York, NY.

Perrin, B. (2002), "How to – and how not to – evaluate innovation", *Evaluation*, Vol. 8(1), SAGE Publications, pp.13-28.

Preskill, H. and T. Beer (2012), *Evaluating Social Innovation*, Centre for Evaluation Innovation, www.fsg.org/publications/evaluating-social-innovation.

Rog, D., J. Fitzpatrick and R. Connor (Eds) (2012), "Context: A framework for its influence on evaluation practice", *New Directions for Evaluation*, No. 135, John Wiley & Sons, Hoboken, NJ.

Scardamalia, M. and C. Bereiter (2003), "Knowledge building", in *Encyclopaedia of Education*, Gale, Farmington Hills, MI, pp. 1371-3.

Schratz, M. and T. Westfall-Greiter (2014), "Beyond the Reach of Teaching: Learning as Experience", Paper presented at ICSEI, Indonesia.

Senge, P. M. (1990), *The Fifth Discipline: The Art and Practice of the Learning Organization*, Doubleday/Currency, New York.

Stoll, L. and L. Earl (2003), "Making it Last: Building Capacity for Sustainability", in B. Davies and J. West-Burnham (eds), *The Handbook of Educational Leadership and Management*, Pearson Education, London.

Stoll, L. (2009), "Knowledge Animation in Policy and Practice: Making Connections", Paper presented at the Annual Meeting of the American Educational Research Association as part of the symposium Using Knowledge to Change Policy and Practice.

Timperley, H.S. and J.M. Parr (2005), "Theory competition and the process of change", *Journal of Educational Change*, Vol. 6(3), pp. 227-252.

Trochim, W., *Introduction to Evaluation*, www.socialresearchmethods.net/kb/intreval.php.