

Developmental Evaluation: Applying complexity concepts to enhance innovation and use

Report from an Expert Seminar
with Dr. Michael Quinn Patton
March 22, 2012

Irene Guijt
Cecile Kusters
Hotze Lont
Irene Visser



Wageningen UR Centre for Development Innovation (CDI) works on processes of innovation and change in the areas of secure and healthy food, adaptive agriculture, sustainable markets and ecosystem governance. It is an interdisciplinary and internationally focused unit of Wageningen University & Research centre within the Social Sciences Group.

Through facilitating innovation, brokering knowledge and supporting capacity development, our group of 60 staff help to link Wageningen UR's expertise to the global challenges of sustainable and equitable development. CDI works to inspire new forms of learning and collaboration between citizens, governments, businesses, NGOs and the scientific community.

More information: www.cdi.wur.nl



Innovation & Change



Ecosystem Governance



Adaptive Agriculture



Sustainable Markets



Secure & Healthy Food

Context, international cooperation (Utrecht, the Netherlands) is a development organisation understanding itself as a social enterprise. We are innovators who use business principles that aim to achieve fundamental social change and generate revenue. Currently Context has 14 staff members, in addition to a number of external Programme Associates. We position ourselves between the development and academic communities, and try to facilitate dialogue between policymakers, academics and development practitioners. Context staff members and Associates have a sound academic background as well as working experience in development practice. For more information: www.developmenttraining.org.

Learning by Design is the company name under which Dr. Irene Guijt operates. She works independently on assessing results in social change, such as through the [BetterEvaluation](#) initiative that is being beta-tested and the [Big Push Forward](#). She advises and researches on, and facilitates learning processes and systems in rural development and natural resource management, particularly where this involves collective action. Work over the past twenty years with multilateral organisations and many development NGOs has focused largely on strengthening critical reflective thinking to strengthen pro-poor development.

Vis-à-vis advies is the company name owned by Irene Visser. She works as a consultant in the management of natural resources, climate change, landscape architecture and Monitoring and Evaluation.

Context,
international cooperation



Vis-à-Vis Advies



Learning by Design

Developmental Evaluation: Applying complexity concepts to enhance innovation and use

Report from an Expert Seminar with
Dr. Michael Quinn Patton

Irene Guijt (organiser, editor)

Cecile Kusters (organiser)

Hotze Lont (organiser)

Irene Visser (documentation)

March 2012

Project code 8141200503

Report number CDI-12-015

Wageningen UR Centre for Development Innovation

Developmental Evaluation: Applying complexity concepts to enhance innovation and use
Report from an Expert Seminar with Dr. Michael Quinn Patton

Guijt, I.
Kusters, C.S.L.
Lont, H.
Visser, I.

March 2012
Centre for Development Innovation, Wageningen University & Research centre

This report summarises the discussions and presentations of the Expert Seminar 'Developmental Evaluation', which took place in Wageningen on March 22, 2012. The Expert Seminar was organised by the Wageningen UR Centre for Development Innovation in collaboration with Learning by Design and Context, international cooperation.

Cover photo
Guy Ackermans
www.ackermans.net

Orders
+ 31 (0) 317 486800
info.cdi@wur.nl

Preface

Dr. Patton is one of the founding fathers of the evaluation profession in the 1970s to which he has dedicated his life. He is a former president of the American Evaluation Association and author of *Utilization-Focused Evaluation*, now in its fourth edition, and *Qualitative Research and Evaluative Methods*, now in its third edition. His 2006 book *Getting To Maybe*, with Frances Westley and Brenda Zimmerman, introduces the idea of Developmental Evaluation. His latest book is *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, Guilford Press, 2011.

Monitoring and evaluation in development cooperation has a crucial role to play in deeper learning and understanding to tackle fundamental issues. For example, estimates are that in 2050 we will have a global population of nine billion, and we will need twice as much food and resources to reduce hunger and poverty. What does this mean for the bottom four billion people, also in the context of climate change? For this, simple quantitative assessments are insufficient; we need to bring science and emotion together as we search for new ways of knowing, evaluating and adapting to dynamic contexts.

Evaluations inevitably highlight problem areas and weaknesses. This touches individuals' emotions, often deeply. Who likes to be wrong? As discussed in *Being Wrong* (Schulz 2011), many of us find this an unpleasant emotional state and we find ways to tackle it. A key challenge for evaluation is how to contribute to people achieving results on topics that are crucial for the planet and its people. 'Developmental evaluation' offers significant insights for evaluative processes that support the evolution of innovations and solutions for our deep-rooted problems.

Dr. Jim Woodhill
Director Wageningen UR Centre for Development Innovation

Acknowledgements

We are very grateful for the opportunity to having experienced a seminar on Developmental Evaluation in the Netherlands with Dr. Michael Quinn Patton. We deeply appreciate his wisdom and experience and take a range of insights forward in our work. We would also like to thank our partners who provided funds to make this event possible: GIZ (Martina Valhaus, Sabine Dinges), Hivos (Karel Chambille), Cordaid (Rens Rutten) and ICCO (Dieneke de Groot). We are also very grateful to our organisations that have allowed us time, money and operational support to prepare this important event: Centre for Development Innovation and Context, international cooperation. The participants' active energy was highly appreciated. Last but not least we would like to thank the documenter Irene Visser for writing up this report (with guidance and editing by Irene Guijt).

The organising team

Cecile Kusters, Wageningen UR Centre for Development Innovation

Irene Guijt, Learning by Design

Hotze Lont, Context, international cooperation

Table of contents

Preface	iii
Acknowledgements	iv
Executive summary	vi
List of abbreviations and acronyms	vii
1 What is developmental evaluation?	1
1.1 Introduction	1
1.2 From classic types of evaluation to developmental evaluation	1
1.2.1 Formative and summative evaluation: the search for the “it”	1
1.2.2 Developmental evaluation: there is no stable consistent ‘it’	2
1.2.3 Definition of Developmental Evaluation and Core Questions	2
1.2.4 Key DE Characteristics	3
1.2.5 High degree of process use(s)	3
1.2.6 Tracking strategies	3
1.2.7 What is the essential difference between DE and other evaluation forms?	4
1.2.8 Fundamental issue: development vs. improvement	5
1.2.9 And what about the difference between DE and monitoring?	5
1.3 Complexity concepts and evaluation	5
2 The utility of developmental evaluation	9
2.1 The evolution of evaluation practice: From best practices and evidenced-based practices (models) to adaptation (principles)	9
2.2 Under what conditions is DE useful and when is DE not suitable?	9
2.2.1 The sacred baseline	10
2.2.2 Going beyond static accountability models	10
2.2.3 Matching the evaluation process and design to the nature of the situation: Contingency-based evaluation	10
2.2.4 Not all evaluation should be DE	11
2.3 As everything in international development is complex sooner or later, why should we not always undertake DE?	11
3 Practical implications and ideas	15
3.1 Programme design and implications for evaluation & programming	15
3.1.1 Evaluability assessment: is there a programme?	15
3.1.2 Picking indicators - what get’s measured gets done	15
3.1.3 Arguments to use to create space for DE	15
3.1.4 Speaking truth to power	17
3.2 Implementing DE: Some practical and ‘attitude’ issues	17
3.2.1 DE without an external evaluator?	17
3.2.2 Implications for roles: build trust, share values and not beliefs	18
3.2.3 Monitor the early steps and have conversations	18
References and resources	19
Appendix 1 – Biography	21

Executive summary

This report summarises the discussions and presentations of the Expert Seminar 'Development Evaluation', with Dr. Michael Quinn Patton, which took place in Wageningen on March 22, 2012. The Expert Seminar was organised by the Wageningen UR Centre for Development Innovation in collaboration with Learning by Design and Context, international cooperation.

The report describes Developmental Evaluation (DE), which mainly focuses on the question: *'What is getting developed, what is not getting developed and what are the implications?'.* The accountability question in DE is *whether something is getting developed.* It also places complexity sensitive DE in the wider context of traditional evaluation approaches.

Key characteristics of DE are:

- Focus on *development* (versus improvement, accountability or summative judgment)
- Takes place in complex dynamic environments
- Feedback is rapid (as real time as possible).
- The evaluator works *collaboratively* with social innovators to conceptualize, design and test new approaches in a long-term, on-going process of adaptation, intentional change, and *development.*
- The DE evaluator can be part of the intervention team.
- The evaluator's primary functions are to elucidate the innovation and adaptation processes, track their implications and results, and facilitate on going, real-time, data based decision-making in the developmental process.
- DE becomes part of the intervention.

The report describes when and when not to use DE. It explains the relevance of DE for the development sector, and how to open up political spaces to do DE, and therefore to nurture innovation.

The report concludes with practical implications and ideas to take DE forward. Some lessons can be learnt from private industry: diversified programming, acceptance of failure, embracing failure as seeds for successful innovations, and setting aside budget to experiment. All material in this report is based on Dr. Patton's lectures and answers to participants' questions.

List of abbreviations and acronyms

AEA	American Evaluation Association
CDI	Wageningen UR Centre for Development Innovation
DE	Developmental Evaluation
IDRC	International Development and Research Centre
JCSEE	Joint Committee Standards on Educational Evaluation
M&E	Monitoring & Evaluation
NGO- IDEAS	NGO Impact on Development Empowerment and Action
RCT	Randomised Control Trial
UF-E	Utilization Focused-Evaluation
USAID	United States Agency for International Development
Wageningen UR	Wageningen University & Research centre

1 What is developmental evaluation?

This report summarises the seminar given by Dr. Patton. The content is a mix of quotes, summaries and interpretations. We hope that the essence of his ideas have been conveyed accurately.

1.1 Introduction

Development Evaluation (DE), and good evaluation in general, is an option for social innovators, deeply committed professionals, dealing with issues such as poverty, sexual violence, and HIV & Aids. Most people want to 'do good', do it well and believe that they are doing so. Therefore, evaluation is a highly emotional exercise. An intrusive external judgement can be painful. On the other hand, professionals do not want to waste time on things that do not work. Social innovators are at heart of the niche of innovation; they have a vision and are driven by values. Their motivation comes from the unacceptability of hunger, disease & violence, to which nobody has solutions.

The DE evaluator shares the commitment to make the world a better place, but he or she can offer a view and strengthen the internal commitment on finding out what is working or not. By asking questions about how do you know you are doing good? And analysing the many ways in which reality is distorted, which makes that professionals not really know what the effects are.

This way of evaluation has an added value compared to the less meaningful evaluations like what we often see, which is evaluation as part of the paperwork game of getting funding.

The evaluation version of creation stories is asking: How do you know that 'it' is good? The core of the Maori Creation story is that through innovation, outcomes emerge that one would never know that could exist before trying something out (Patton 2011, p. 338).

1.2 From classic types of evaluation to developmental evaluation

1.2.1 Formative and summative evaluation: the search for the "it"

DE can best be understood by first understanding the dominant evaluation model. This model encompasses a classic Terms of Reference (ToR), and Patton says is often a fairly mechanistic, easy to administer, contract approach to evaluation. At the centre of the dominant model is the pair of 'formative' and 'summative' evaluation, originally distinguished by Michael Scriven. Formally, the purpose of formative evaluation is to get ready for summative evaluation. A project needs time in the beginning to work out implementation difficulties, how an idea works in practice. The time to stabilise, forming and improving a model is the focus of formative evaluation. Once 'the model' is ready, only then should it be subjected to summative evaluation. The original distinction of purpose was important, as before that the overall effectiveness (i.e. summative evaluation) was evaluated prematurely.

Originally, formative equalled mid-term evaluation, and summative equalled end-of-project reviews that informed subsequent phases. However, in practice, mid-term evaluations are often more often summative, since it is then that decisions often need to be made about whether or not to fund a new cycle; then the end of project is formative of how to do a better job in the next cycle. Patton also clarified the confusion in the literature about summative being only focused on outcomes and formative being focused on process. Both summative and formative evaluations can and do deal with outcome and process.

So in dominant programme thinking, a stabilised prototype or model is effective and then taken to scale. For the last 50 years, evaluators and policy makers have been caught up in searching for an 'it', a stable consistent 'it'. This framed how people conceptualised programmes. In the evaluation of programme model, we have promulgated the 'journalisable it', the methods we have used, what our work is about, that is what is at stake, was and remains the template of how evaluation works.

1.2.2 Developmental evaluation: there is no stable consistent 'it'

Developmental Evaluation departs from a fundamentally different paradigm. DE is driven by the idea that there never will be a model; the model will never be fixed, it is constantly adapting. DE marks the difference from a classic *modelling* approach, to a *non-linearity unfolding* approach. In the classic modelling approach, we plan around impacts. In the non-linear version we look at the way projects unfold.

DE started from an interaction with a client that had a programme set out to develop leaders. The ever-changing world made it impossible for them to find a fixed curriculum for would-be leaders. But they did care about *developing* leaders. For the client, the traditional formative or summative evaluation was not useful. Dr. Patton was invited to provide feedback to the design team to support them in their on going, adaptive process of changing what they did.

DE focuses on evaluating something that keeps changing. There is no stable or consistent 'it'.

Hence DE originates from a critique on traditional evaluation approaches. The approaches were felt to be not just benign, but actively interfering with innovation. Social innovators made clear that traditional M&E evaluation was their biggest obstacle and not meeting their needs. Rather than 'doing good', the rigid, mechanical, narrowly accountability focused evaluation was doing harm. Innovators need approaches that work with them, and bring about the differences made; they do not need to be blamed by evaluators for not fitting in evaluation models.

1.2.3 Definition of Developmental Evaluation and Core Questions

The purpose of DE is to inform and to support innovative and adaptive *development* in complex dynamic environments. DE brings to innovation and adaptation the processes of asking evaluative questions, applying evaluation logic, and gathering and reporting evaluative data to support project, program, product, and/or organizational *development* with timely feedback. Other names for DE are: Real time evaluation, emergent evaluation, action evaluation, and adaptive evaluation. (Patton noted that definitions should always be made and understood within its context, including formative and summative, outcome, impact, and so forth. There is no universal agreement on any of these concepts.)

Central questions that a DE process seeks to answer – in an ongoing manner are:

1. What is getting developed in interaction with the complex environment?
2. What are the implications, the anticipated consequences of that being developed?

DE can also focus on what is *not* getting developed (in the case of no innovation). There is a focus on *development* and not on the improvement. This is an important but subtle difference. Whether something is developed is the accountability question in DE.

1.2.4 Key DE Characteristics

DE can be said to take place when the evaluation:

- Focuses on *development* (versus improvement, accountability or summative judgment)
- Takes place in complex dynamic environments;
- Ensures feedback is rapid (as real time as possible);
- Sees the evaluator works *collaboratively* with social innovators to conceptualize, design and test new approaches in a long-term, on-going process of adaptation, intentional change, and *development*;
- Has the DE evaluator as part of the intervention team;
- Sees the evaluator's primary functions as elucidating the innovation and adaptation processes, tracking their implications and results, and facilitating on going, real time, data-based decision making in the developmental process; and
- Becomes part of the intervention, supporting the innovation and adaptation as it evolves.

1.2.5 High degree of process use(s)

Dr. Patton's other body of work is Utilization-Focused evaluation (UF-E) (Patton, 2008). UF-E refers to evaluations undertaken for and with specific intended primary users for specific, intended uses. Dr. Patton mentioned that, for a long time, *use* was framed as what happens to findings and recommendations. However, around the time of the 3rd edition of the UF-E book, he realised that much of the evaluator's impact comes from engagement with people: process use.

Process use refers to and is indicated by individual changes in thinking and behaviour, and programme or organizational changes in procedures and culture, that occur among those involved in evaluation as a result of the learning that occurs *during* the evaluation process. Evidence of process use is represented by the following type of statement after an evaluation: "The impact on our programme came not so much from the findings but from going through the thinking process that the evaluation required."

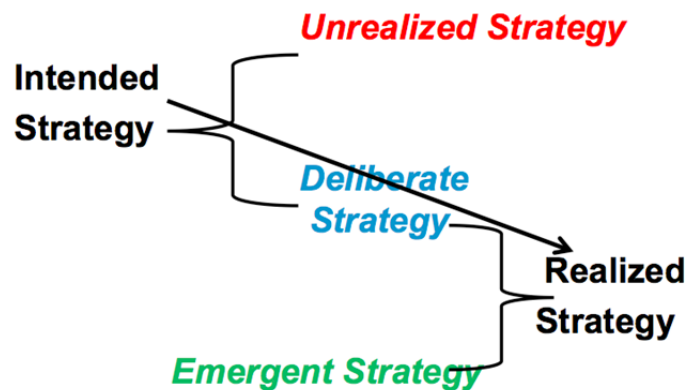
DE is on the extreme end of the spectrum of 'process use' – the questioning and reflecting on an evolving process is what is valuable. But it is not process evaluation, the evaluation of process – it is a process for evaluation for a specific purpose.

1.2.6 Tracking strategies

Henry Mintzberg, one of the world's leading thinkers and writers on management, Cleghorn Professor of Management Studies at McGill University, has spent several decades examining the processes by which strategies have formed in a variety of contexts. Defining *realized strategy* - the strategy an organization has actually pursued – can be viewed as a pattern in a stream of actions. This investigation tracked strategies in organizations over long periods of time. This revealed the patterns by which strategies form and change in organizations, the interplay of '*deliberate*' with '*emergent*' strategies, and the relationships between leadership, organization, and environment in the strategy formation process. Figure 1 illustrates the 'realized strategy' in relation to other pathways.

Figure 1

Mintzberg on strategic journeys



Source: Mintzberg, Tracking strategies, toward a general theory, 2007

Mintzberg found that no high performing organisation implements all they intended as the world changes continuously. The realized strategy is a combination of what was intended, left behind and picked up. DE tracks the forks on the road. Mintzberg's ideas are supported by other management scholars, such as Snowden and Boone (2007): "Wise executives tailor their approach to fit the complexity of the circumstances they face."

1.2.7 What is the essential difference between DE and other evaluation forms?

Many people who first come across DE wonder if it is about a different method or about a different role for the evaluator and a different relationship with the evaluand.

Patton explained that within a DE process, you can use a variety of methods and a variety of roles for evaluator and evaluand, and is methodologically agnostic – you can even use (quasi)experimental methods within a DE process (see Example 1 below).

He stressed that DE is, above all, a different *purpose* for evaluation, neither a role distinction nor a type of evaluation (see also Patton, 2011, p. 187). Hence the focus is on interviewing participants about what is emerging in programmes, and the ongoing development and adaptation in their complex environments. The niche for DE is one of helping to clarify by asking 'What is being developed (not just improved)?'

Example 1

Developmental messaging in presidential campaign

In 2008 in the presidential campaign, for Obama, social media was used for the first time. Four different versions of messages were used in events and fundraising. The content of messages would change, according to what different formats did they respond to. Randomised control trials were applied to four different messages; the design was changed every twice weeks. The campaign and messaging was developing, based on reflective practice (what was going with Hilary Clinton, with news stories etc.)

1.2.8 Fundamental issue: development vs. improvement

“Does it work?” is a common evaluation question but also very static due to the way it is framed. There is an ‘it’ which assumes predictable results as opposed to looking at a program or initiative in terms of what is effective for whom, under what conditions¹? In complex situations, larger systems perspective comes into play. In workshops with staff, DE facilitates reflection about changes, and whether the change is significant to develop a different way of going about the work.

Example 2

From improving a programme for men to developing a programme for men and women

The original design was an individual coaching model for people needing employment. After 3 months, feedback contributed to a job skills development model in the programme itself. However, in the middle of the programme period, the welfare system was changed, causing a huge demand for employment for women. Before they even noticed, half of the programme participants were women, with new issues related to day care and welfare to work laws. Due to contextual changes, it was no longer about improving a programme for men but a programme for men and women.

1.2.9 How DE and monitoring differ

Given the ongoing nature of a DE process, it is akin to how some programmes in international development undertake monitoring. So how is DE different from such monitoring (both operational as well as strategic)? And how is it different from action research? And from outcomes oriented monitoring? Wouldn't it be better to call it developmental monitoring or management?

Patton responds by referring to the standard use of the term monitoring. Monitoring data is based on routine recordkeeping against predetermined indicators. So the unintended is not investigated. But monitoring data can be a great source for reflection within DE, for example, if indicator levels suddenly change. DE would focus on understanding why it is changed and reflecting on implications. DE with a developmental and adaptation approach goes beyond monitoring, more of a problem solving approach (Patton, 2011, p 280).

Patton also referred to the work of Hans Rosling of the Karolinski institute, who brilliantly combines statistics with trends and is challenging common beliefs.

1.3 Complexity concepts and evaluation

Patton discussed the characteristics of complex adaptive systems (see Table 1 below). Complex situations are: highly emergent (difficult to plan and predict), highly dynamic (rapidly changing) and relationships are interdependent and non-linear (rather than simple and linear (cause-effect)). On a practical level this means that DE evaluators engage with programme staff about both what is known and controllable, and about what is unknown. In chapter 5 of his book (2011), Patton elaborates on the developmental evaluation implications of each of these dimensions of complexity in detail.

¹ This is the classical question that drives ‘realist evaluation’; see CDI seminar report 2011 based on Dr. Westthorp’s presentation.

Table 1	Characteristics of complex adaptive systems
	<p><u>Emergence</u>: Patterns emerge from self-organisation among interacting agents. What emerges is beyond, outside of, and oblivious to any notion of shared intentionality. Each agent or elements pursues its own path but as paths intersect and the elements interact, patterns of interaction emerge and the whole of the interactions becomes greater than the separate parts.</p> <p><u>Nonlinearity</u>: Sensitivity to initial conditions; small actions can stimulate large reactions, thus the <i>butterfly wings</i> (Gleick 1987) and <i>black swans</i> (Taleb, 2007) metaphors, in which highly improbable, unpredictable and unexpected events have huge impacts.</p> <p><u>Dynamical</u>: Interactions within, between, and among subsystems and parts within systems are volatile, turbulent, cascading rapidly and unpredictably</p> <p><u>Uncertainty</u>: Under conditions of complexity, processes and outcomes are unpredictable, uncontrollable, and unknowable in advance. The book 'Getting to Maybe' (Westley et al, 2006) captures the sense that interventions under conditions of complexity take place in a Maybe World.</p> <p><u>Co-evolutionary</u>: As interacting and adaptive agents self-organise, ongoing connections emerge that become co-evolutionary as the agents evolve together (coevolve) within and as part of the whole system, over time.</p> <p><u>Adaptive</u>: Interacting elements and agents respond and adapt to each other so that what emerges and evolves is a function of ongoing adaptation among both interacting elements and the responsive relationships interacting agents have with their environment.</p>
	<p><i>Source</i>: Patton, 2011, p.8</p>

Patterns emerge from self-organization among interacting agents. After many years of interviews, Patton found consistency in the answers to the question about the most important aspects for people: other people they met in the programme. People find each other. Evaluators randomize from individuals, in programmes of all kinds, but people are social animals and have experiences as groups. People do not experience a programme as individuals, social groupings are hugely important in terms of impact. If evaluators do not take into account subgroup differences, when people find each other, e.g. around gender, language, HIV or Aids, they will miss the point. Evaluators need to understand the social dynamics of self-organisation. Interaction of small effects can lead to patterns that we cannot predict or control.

For example, for the evaluation of a wilderness leadership programme, Patton chose for participant observation, instead of a survey in the end of the programme. Thirty university administrators in the USA went on 10-day hiking trips in the winter and summer. In the first hour, two groups formed, which were sustained throughout the entire year. The 'Truckers' were achievement people 'let's hit the trail and get there'. 'Turtles' barely got up in the morning and walked at a slower pace. Some people unsuccessfully tried to move from their initial place as a trucker or a turtle. This differentiation greatly influenced participants' experience of the programme and would be missed by surveys.

Some aspects of the world operate along Newtonian lines. But non-linearity sharpens our awareness of how critical events make a huge difference at macro and at micro level. Programmes, and therefore evaluations, need to cater for dynamical and non-linear relations/reactions. Recognising the dynamical nature of much of life, evaluators need to inquire into questions such as: What went as planned, what not,

what factors impinged upon the programme, process, project being evaluated? The question for evaluators is how to fit our skills and practices around more chaos, disorder, and not assume order.

Dr. Patton mentioned two examples of ripple effects at macro level, and at micro level. Ripple effects at macro level happened with SARs disease in Toronto, where 32 people died, which had some effects on the amount of client hotels, restaurants, night clubs and bookstores. In contrast, Mexico City shut down when there was one case of swine flue. The fear caused huge ripple effects, out of proportion to the actual threat.

An example of a ripple effect at local level is the following. Lives are chaotic and sometimes a small incident can have a huge effect and cause vicious circles of poverty, which was shown in an evaluation of a programme of poor people aiming at education in the evenings for a high school degree. For example the person would go to work by car, the car breaks down, there is no cash to pay the repair, when the public transport is used they come late for work, the person get fired, there is no way to pay the rent, and the family is homeless, let alone go to school and 'receive' the education.

The crucial DE phase is early on in a new initiative when ideas meet reality and everyone involved quickly finds out which assumptions need revisiting. For example, for IDRC Canada, Mitzen developed a mid-term evaluation model. She worked with development staff around the world: When does it make sense to undertake an evaluation? When does key learning and when do core developments emerge? Is there an ideal time or opportunity for learning in a five year programme? As in many international agencies, IDRC employs a 18-24 month process of developing an R&D programme. After approval of the Board, the implementation of the programme can start. Staff found that the first 6-9 months of the programme are crucial, because then all the assumptions are put to test.

Naturally, big changes, such as hurricanes or changing subsidy schemes, can have huge impacts. For example, in the Caribbean, repairing the damage after 3 hurricanes was more important than continuing the current extension programme. Also, the same programme, which included banana producers, was affected by a change in subsidy from the United Kingdom. Both events had unexpected effects on the infrastructure of the programme, causing the programme to adapt.

Often very minor changes can be huge in terms of the theory of change. An example is a coaching programme in Minneapolis that trained coaches to work with 'low income men', and using some incentives/stimuli such as bus passes. In the beginning some coaches gave bus passes to everyone, others had application procedures. In the end, the entire programme created a rule for bus passes to make sure every body got treated the same. Patton started a reflective practice process with the programme around the question, what has changed in the last three months? The coaches were surprised to find out that in a bigger picture, they had not noticed their shift from a flexible not standardized procedure, to a rigid fashion of a highly rules system. This shift influenced their theory of change.

2 The utility of developmental evaluation

2.1 The evolution of evaluation practice: From best practices and evidenced-based practices (models) to adaptation (principles)

Evaluation grew up in relation to 'project testing' models under a theory of change that said that pilot testing would lead to proven models that could be disseminated and taken to scale. This model is a very seductive mental model; find something that works, disseminate, and assume it is context free. Conditions that challenge this traditional model-testing evaluation are those of high innovation, development, high uncertainty, dynamic, emergent and systems change. Principles, not models, will provide you with directions on how to adapt to context. This has also implication for evaluation.

Taking a model to scale involves *fidelity-focused* evaluation, while adaptive management calls for *Developmental/Evaluation*. In fidelity- focused evaluation, the 'fidelity police' ask: "Are people implementing the model exactly as specified?" Sometimes with answers such as: "Please don't tell the funder we are adapting". For evaluators, this is often where it becomes interesting. DE is in the middle between *bottom up* adaptive management with proven principles and *top down* dissemination of proven models. DE harvests ideas about what works and what does not.

An example is McDonalds. As in any franchise operation, McDonalds takes pride in producing the same big Mac all over the world, regardless of geographical position. However, the experience differs among the world. The company adapts the 'experience package' around consuming to the situation. In an area where a lot of retired people live, employees are more senior than in suburban shopping malls, where music is louder. In France, the company will serve wine, in Latin America hot sauce.

So projects and programmes often use a mix of recipe based (quality control) and principle based programming (the unknown). A good example is immunisation campaigns in different political contexts. The dose of a polio vaccine for a child, drink or shot is known (recipe). But every campaign is also developmental; tracking outbreaks and what the impacts are. In India, after the last campaign, they had to revaccinate within 150 km of outbreak. Anthropological studies studied: How did we miss that child? The cause was that there was rumour in modern Hindu families, that the vaccine was a plot to sterilise Muslim children and they hid their children. Northern Nigeria stopped an immunisation campaign, believing it was a western plot.

2.2 Under what conditions is DE useful and when is DE not suitable?

Thematically, DE is well suited to the following types of applications, in which innovation is central:

- social movements and networks
- advocacy evaluation
- large-scale, cross-sector, collaborative initiatives
- R&D in public health, technological innovation, science
- public policy implementation.

But it is not just thematic appropriateness that must be kept in mind. Context-specific evaluation requirements and conditions will determine whether or not DE is appropriate. For example, DE is not suitable when the primary engine is traditional accountability with performance indicators that are highly

prescribed in the beginning. The evaluation mandate DE is counter to that, since indicators are allowed to evolve.

Within DE, the larger questions are asked about how short term results relate to outcomes and changes on system level. (Again, this is different from improvement, where a model is applied to make the situation better). Developmental work is about changing what you are doing in relation to context. Questions that are asked are: does it work, what works under what condition, what is getting developed in that response? What kind of system is emerging, and what is going to be left behind?

DE evaluators can have added value for managers by asking empirical questions, to help them think through what kind of data they need to inform their decision-making. One of the ironies is that very good managers already do DE, for bad managers DE is just an extra thing that needs to be managed.

2.2.1 The sacred baseline

In mainstream evaluation, nothing is more sacred than the baseline. However, the baseline is often inaccurate because we do not know in the beginning what is true or not or what information will be needed or not. More correct answers may be found in subsequent evaluations, after having built trust and relationships with staff and intended beneficiaries. Such early evaluation interviews themselves are part of DE as they help to define the problem and understand the context. In the beginning, at the time of baseline establishment, people often do not trust us, as evaluators, enough to tell what is going on. People lie about drug use history, violence history etc. Updating the baseline is often seen as corrupting the baseline, (what school of ethics did you go to?) but it is essential to get accuracy for subsequent comparisons.

2.2.2 Going beyond static accountability models

Patton distinguishes between rigid accountability and other forms of accountability. Referring to the Mintzberg diagram (Figure 1 in 1.2.6), the extent to which environments are complex and dynamic partly determine the extent to which results can take place. Rigid accountability holds people accountable for what they intended to do in the beginning. What DE requires is that people are honest about themselves. People can document that some assumptions were not true. This is a higher level of accountability, to one's strategy and broader intentions, rather than to the precision of a preconceived path. Hence redefinition is allowed, and is accompanied by systematic documentation.

An example of how IDRC is accountable in this way to their treasury board is as follows: "We are accountable for what we accomplished, and we justify what we did not do. We show how we picked up new opportunities; we are accountable for both outcomes and learning."

So do not abandon the accountability language entirely. The challenge is not to take a narrow form of accountability. It is also important to remind political leaders of the fact that the very rigid forms of accountability imposed on development work no business operation would tolerate.

2.2.3 Matching the evaluation process and design to the nature of the situation: Contingency-based evaluation

Participants wondered if DE is most suited where everybody is harmonious, focused on the same objective. How do you apply DE in conflicts of interest in many programmes, self-interest, and diverging groups interest?

Patton says that every process that is going to be collaborative is suitable, but that initial conditions will differ. It comes back to the idea of conflict management and working with people who might be able to recognise that they do not know what do. Conflict mitigation programmes are highly contextual, and impossible to predict: “We want to live in peace 3 years from now.” “This space will be a conflict-free area in 3 years time.” So it is contingent. Bring the highly conflicted groups together, and establish initial conditions and explain/agree that the path forward is unpredictable. An example is USAID, for their conflict management programme where there are no ‘knowns’, have invested in DE (and everybody can understand that in these kinds of programmes you cannot have indicators).

2.2.4 Not all evaluation should be DE

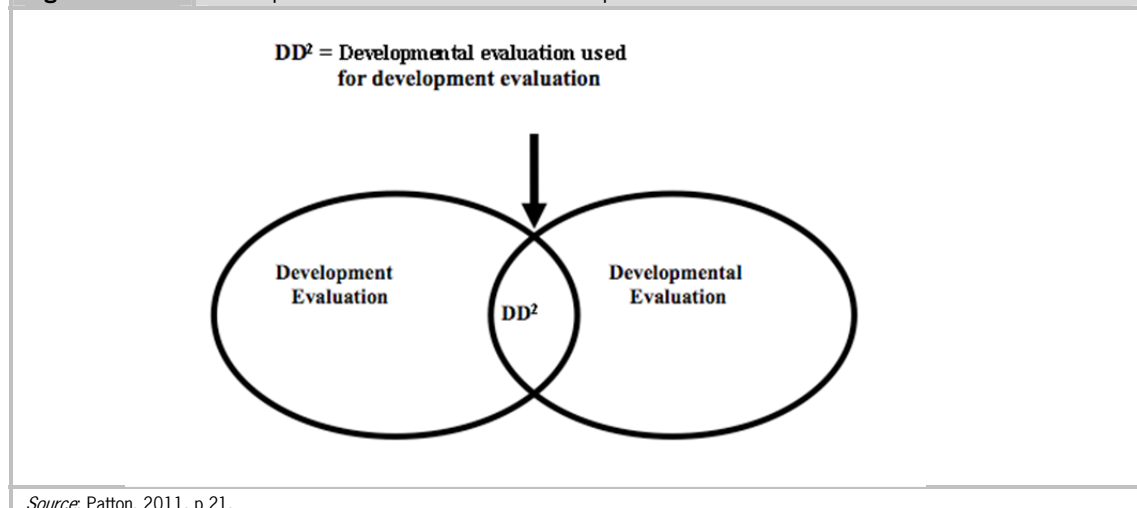
Patton is not suggesting that all evaluation should be DE. The niche for DE is quite small and particular. DE is suited for where people are innovating. Also, it is hard to do funding wise, and methodologically, therefore much misunderstanding is possible. Formative and summative should not go away. What is needed for DE to be of use is recognition that a particular situation is complex and that a rigid, preconceived design does not work. So when the very complex nature of the situation requires DE mind set and DE form of engagement.

2.3 As everything in international development is complex sooner or later, why should we not always undertake DE?

Participants wonder about why DE should not always be the default option, given that much of development is complex. Conversely, what is the risk that we are treating situations developmentally when they are not?

First, Patton stresses that ‘developmental evaluation’ does not equal ‘development evaluation’. Developmental evaluation can be used to evaluate development (see Figure 2 below), and is true to complex dynamic development, in which some parts are known are controllable and others are not. (For example the polio vaccine dose is simple and known, but a campaign for immunisation is complex.

Figure 2 Developmental evaluation used for development evaluation



Source: Patton, 2011, p 21.

Second, let's turn to the five purposes of developmental evaluation and how this might answer the question about 'when to DE and when not to DE'? Table 2 lists these purposes.

Regarding 'ongoing development', this refers to a particular community, with a mix of models and principles for that situation.

Regarding the second purpose, 'adapting principles in a new context', this is the case when the insights from this community would be applied in another community. (In complexity language, these are 'simple rules'). An example of an extension programme in the Caribbean: there were principles of farmer engagement, but the substance of what the extension delivered was subject to the context, and nature of the farm operation (e.g. coconuts, bananas, vegetables, animal husbandry). Another example are microfinance groups to create collateral, are highly varied, but microfinance principles are the same.

The third purpose, develop a rapid response in turbulent disaster, is more self-evident. An example is the earthquake in Haiti. Real time feedback was needed, not only classic monitoring to get information about where water is needed, but also more systems development questions like how to manage cholera outbreaks and insights about the consequences of resource investment and lack of donor collaboration. In this rapidly changing situation, solving one thing might create unanticipated consequences. A situation emerged in which NGOs hired the higher professionals and destroyed the existing medical system, by paying them more. What is getting developed: a humanitarian response. What kind of system is emerging, what is being left behind: the long term indigenous health care system. What is the role of donors and aid in that? Noticing this helps tie the way the short-term humanitarian aid is done. Very early on, will create opportunity to create insights into what is left behind by the way resources are coming in.

Table 2 Five purposes of developmental evaluation		
Primary developmental evaluation purpose	Complex system challenges	Implications
1. On going development	Being implemented in a complex and dynamic environment	Models may move into formative and summative evaluation, others remain in developmental mode Inform different potential scaling options
2. Adapting effective principles to a new context	Innovative initiative develops 'own' version based on adaption of effective principles and knowledge	Top-down—general principles knowledge disseminated Bottom-up—sensitivity to context, experience, capabilities and priorities Adaptation vs. Adoption
3. Developing a rapid response in turbulent disaster situations	Planned interventions must adapt and respond as conditions change suddenly	Planning, execution and evaluation occur simultaneously
4. Pre-formative development of potentially scalable innovative	Changing and dynamic situations require innovative solutions to worsening conditions Model needs to be developed/does not exist	Models may move into formative and summative evaluation, others remain in developmental mode Inform different potential scaling options
5. Major systems change and cross scale developmental evaluation	Disrupt existing system Taking an innovation to scale Major systems change and changing scale will add levels of complexity, new uncertainties and disagreements	System is the unit of change (and unit of analysis) Adaptive cross-scale systems change innovations assume complex, nonlinear dynamics requiring agility and responsiveness
<i>Source: Patton, 2011, note that originally only three purposes were distinguished, Patton included two other ones.</i>		

The fourth purpose is pre-formative development of a potentially scalable innovative. Examples are the 3 M post-its notes. The notes were developed in an innovative space, discovering a use for a glue that does not glue. The notes went to scale space and now post-its are all over the world.

Finally, DE is suited when dealing with systems change and systems as unit of analysis. In some cases there are common findings that parts of programmes are dysfunctional, organisations, and systems that were dysfunctional, or where there is no history or models on how to do things. For example, the big changes in philanthropic organisations, or Obama pleading for changing the health and education systems. An example is from the Harvard early childhood education centre, that undertook a very innovative venture based on the assertion that everything that we know about early childhood education was wrong. They convened 100 big thinkers in brain research, neurological researchers, early childhood, emotional cognitive research to begin remaking the entire field. They had no idea what they are going to come up with but they were supposed to develop something. There were a lot of unknowns but as they talked, they started to build a new field, starting from a vague fuzzy vision.

Even when there is no such thing as systems- they are, after all, mental constructions - a systems approach is introducing the nature of change, and understanding system sets of relationships. Projects are not self-contained and they depend on the system in which they are embedded.

Finally, the following table (3), in which traditional evaluation and complexity sensitive development evaluation are contrasted, helps to understand in another way, when it is appropriate to use a developmental approach.

Table 3 Value contrasts between traditional evaluation and complexity sensitive development evaluation	
Traditional evaluations...	Complexity-based, Developmental Evaluation...
1. Testing models	1. Supporting innovation and adaptation
2. Render definitive judgments of success or failure: Does the programme work?	2. Rendering nuanced, disaggregated feedback & generate learnings for adaptation & development: What works for whom in what ways under what conditions?
3. INDEPENDENCE Evaluator external, independent, objective	3. RELATIONSHIP-FOCUSED, COLLABORATIVE Evaluator a facilitator and learning coach bringing evaluative thinking to the table, supportive of innovator's vision
4. CONTROL Evaluator determines the design based on the evaluator's perspective about what is important. The evaluator controls the evaluation.	4. OPENNESS & NATURALISTIC INQUIRY Evaluator goes with the flow, watches for what emerges
5. CERTAINTY <ul style="list-style-type: none"> ✓ Predetermined outcomes ✓ Fix the design upfront ✓ Predetermined indicators ✓ Fixed questions 	5. FLEXIBILITY <ul style="list-style-type: none"> ✓ Emergent outcomes ✓ Flexible design ✓ Emergent indicators ✓ Dynamic questions
6. Linear cause-effect thinking and logic models	6. Systems and complexity thinking with attention to dynamics, permeable boundaries, interdependencies, and emergent interconnections
7. Value top-down change based on generalizable findings across time & space High fidelity, prescriptive "best practices" based on summative evaluation	7. Value bottom-up principles that provide direction but have to be adapted to context Aim to produce context-specific understandings that inform ongoing innovation and adaptation.

8. Accountability focused on and directed to external authorities and funders.	8. Accountability centered on the innovators' deep sense of fundamental values and commitments – and learning as accountability
9. Being outside the action, above the fray	9. Being part of the action, engaged in the fray
10. TRUTH Speaking truth to power	10. PERSPECTIVES Facilitating dialogue and engagement with complexity and shifting understandings
<i>Source: Powerpoint slide Patton Note: an elaborate version of this table can be found in Patton 2011, p23-26 (or annex B Kusters et al 2011)</i>	

3 Practical implications and ideas

3.1 Programme design and implications for evaluation & programming

3.1.1 Evaluability assessment: is there a programme?

Dr. Patton pointed at the notion of the link between evaluation, design and programming, which has evolved during the time frame of his career. In the 1960s, needs assessments were done in order to design a programme with clear goals and objectives. In the early 1970s, evaluators found out that in spite of a lot of activities of practitioners, often there was no programme or people working outcomes. As a reaction, a logic model for both programme and evaluation was developed, with activities and outcomes. Suddenly evaluators were at the front end of programme design, to determine the programme was ready to be evaluated.

Tools like LFA and logical models became dominant for programming. Although logic modelling is helpful to connect the dots, such models are not fit for everything. (Indeed Patton has looked at evaluation textbooks that use the story of Alice in Wonderland to make fun of people who do not have logical clear specific and measurable indicators outcomes at the start. Alice is lost and meets the Cheshire cat. Alice asks 'I am lost, can you help?' And the cat responds 'Where are you trying to get to?' On hearing her reply 'I don't know', he says 'then any road will take you there'.)

3.1.2 Picking indicators - what get's measured gets done

The Logical Framework Approach has huge impact in focusing resources. Our designs tell people what priorities are. But if you measure the wrong things, people do the wrong things. Indicators direct energy and resources, they are not benign.

It is closely tied to the question about where indicators come from. On the one hand, development is often seeking to empower people in bottom-up processes. Yet outside people develop these programmes ahead of time with indicators about what kind of development is desirable. This tension is problematic.

Within DE, indicators are allowed to evolve. This implies that often it is better to start the evaluation after the first meeting of stakeholders, because in that meeting stakeholder still might decide to change their theory of action.

3.1.3 Arguments to use when you want to create space for DE

Instead of tackling the entire accountability question, professionals working with DE can try to find cracks in the system, and work with willing colleagues. Often, less is more, so working within a smaller supportive niche is more effective. See for example the innovation model developed by Rogers (1962), in which he identifies the early adopters, who adopt an innovation until a tipping point, after which the innovation becomes common practice for a larger group of people. This 'early adopters' notion might be a starting point from which evaluators can work.

Rather than only focusing on predetermined outcomes, Patton encountered politicians that could identify: "I would get cover for that kind of cutting edge work, as long as there are other programmes that deliver programmes with predictable outcomes, so that I can show that public money is well used". They know

that anticipating the future, by allowing new ways of thinking about how to solve tough problems creates public support.

To open up political space to do evaluation, there are some 'take away lessons' from the private sector, from which we can learn to support our arguments.

Take away from private sector #1: diversified programme approach

One of the lessons from the corporations is a diversified programme approach: the combination of concrete deliverables with 'blue-sky' projects. Parts of the resources are invested into innovation. Businesses invest a proportion of their profits into R&D for the future, as innovation is needed in order to remain in business. This means that there is space to innovate, and an evaluator or manager needs to analyse how different projects with different timelines intersect. For evaluation this means that in some parts of the programme there are knowns and control, whereas other parts deal with the unknown, uncertain but potential successful innovations.

Another example of a diversified programme approach with surprising ripple effects comes from a leadership programme. After training, participants had to return to their community, with a diversified project portfolio. They had intermediate objectives to improve the park (paint, fix lights and swings), which balanced long term objectives like to create jobs to keep young people in the community. However, as soon as they repaired the lights, an unanticipated effect started. Drug dealers like to operate in the dark, and after lights were repaired, the dealers were shut out. The original design did not call for lights, but the park is now a demonstration project for new kinds of lighting and is a safer place to play. These are ripple effects in terms of systems analysis.

Take away from the private sector # 2: invest into innovations and accept failure

In a similar fashion, another lesson is to frame the concept of failure in a different way. The primary mind set in evaluation is a negative, intrusive accountability mind set, which interferes with innovation. More often than not, external evaluators are interested in paper trails and not in the rationale of how the programme evolved and adapted to the situation. Everything that deviates from a predetermined outcome is perceived as failure. It would be a shame to show an evaluator a list of failed projects. However, nobody ended up five years later where they thought they would be in the present.

In the private sector, R&D space is created to develop entire new products. There is greater risk, because it is unknown what works and does not work. An example of innovation space is the Bell labs where engineers were allowed to engage together, without even knowing what would come out of this. It was not an ordinary laboratory, but an innovation and creativity laboratory. In capitalism and private sector, high failure rates are accepted. However, strangely, in the public sector of the same capitalist systems, high returns on Investments (HROI) for every dollar are expected. In a political dialogue, it is important to remind officers about development and improvement, it is very difficult to find answer to complex problems.

Take away from business # 3: allow methodological combinations

The last lesson to be learned is about how to collect evidence. Business folks make decisions based on evidence from Randomised Control Trials (RCT). However, they might use RCT for products, but not for how they do their business, about relationships with people. This combination proves to be a struggle in bureaucracies, and almost no evaluation design, is allowing budget for open-ended fieldwork. This means that there is no way to find out what has emerged. Dr. Patton says that often there is no real room for innovation, even not in tenders explicitly asking for innovation proposals. Also, even when the mandate to innovate and to take risk is given by the board; a board might still think: "take risk, but just make sure it works." Fortunately, some philanthropic foundations are working on the cutting edge and are embracing risk as part of their mission.

3.1.4 Speaking truth to power

We can learn something from the role that was played by the court jester in medieval times to serve the monarch with important information. The court jester spoke truth to power in a way that could be heard by the monarch, typically using humour and wit.

In modern times, an evaluator can play/embrace an important positive role to leaders to whom they can offer real-time information to support informed decisions. Surplus negative baggage around compliance can be left behind.

Figure 3 Speaking truth to power



3.2 Implementing DE: Some practical and ‘attitude’ issues

3.2.1 DE without an external evaluator

Internally located DE evaluators have the advantage that they can participate in important decision-making events. Often a combination of internal and external is best to support the credibility of the internal one. The internal keeps the books, the external validates/audits the book. The internal person can be present at key moments.

Not every personality type resonates with the role of DE evaluator. Obsessive types with rigid personalities and high need for control, do not make good DE evaluators. If you aspire to be a developmental evaluator, you need to examine yourself: it is fit for you, can you thrive in that climate, and can you play a supportive role for people?

It is very important to understand that the DE evaluator is still an evaluator. Despite the role confusion around this, DE is guided by principles (see for example the Evaluation standards developed by the American Evaluation Society and the JSCEE). Also, somebody internally involved in DE is not in the position of assessing the merits and worth of a programme. It also depends on the scale and ambition:

Patton explained that the more claims are made by a programme, more rigour and external validation is needed.

3.2.2 Implications for roles: build trust, share values and not beliefs

A DE evaluator cannot play the role of independent external evaluator because he or she is involved in programming. DE creates outputs and material for formative and summative evaluation and will bring about a different kind of credibility.

Evaluators should be more assertive about having shared values. For example, my brother died of Aids, so I care, you better be good at the job against HIV and Aids. I explicitly say: "I am not pretending to be value free, I share their values, but I don't share their beliefs, I am not bought into their assumptions." And that is very clear in our relationship that is based on trust and the ability to ask the tough questions to test their beliefs. This requires trust between social innovators and evaluators, and talk through to each other. They know my values, I am able to offer tougher feedback, because they know we share.

Much of evaluation is methods driven or accountability driven. DE is relationships driven, between the people, learning to apply and interpret systematic reality testing feedback, around the issue of what is being developed. Are problems being solved? Can we connect the dots? What results are achieved? Thinking deeply about things, gathering and interpreting data is a challenging process.

Patton is worried about DE being stuck in a box labelled 'process-orientated'. DE is not only a process methodology! It is about process and content.

3.2.3 Monitor the early steps and have conversations

It is critical to monitor the early steps. Similar to forestry, if saplings die after one year, you know you will not have a forest. How well people take these early steps, where there is no logical model but some signals of the existence of a basic management capacity. Can you get the first steps done; work with the community, community collaboration? Are they able to organise the meeting, take notes, some procedures of conflict mitigation? These are early indicators of people's ability to manage some kind of developmental work.

DE is based on conversations with people. If you work with people, try to find out metaphors for issues on which they work, ask people about important occurrences and how they will have come to understand such occurrences. Often metaphors represent a view on development. Simple changes in language define or point to entirely different development paradigms. For example, in New Zealand Maoris were talked 'about'. There is a fundamental difference between developing something "To Maori, With Maori, or As Maori". Dr. Patton stressed that doing things *as* who we are is Developmental Evaluation. (Patton, 2011, p. 276)

References and resources

Key references:

- Gamble, J.A.A., *A Developmental Evaluation Primer*, The J.W. McConnel Family Foundation, Canada, 2008
http://tamarackcommunity.ca/downloads/vc/Developmental_Evaluation_Primer.pdf
- Patton, M.Q. *Developmental Evaluation: Applying complexity concepts to enhance innovation and use*, the Guildford Press, 2011, Chapter 1 <http://www.guilford.com/excerpts/patton.pdf>
- Patton, M.Q. *Webinar on Developmental Evaluation*, UNICEF webinar
<http://mymande.org/index.php?q=content/developmental-evaluation>

Additional references:

- American Evaluation Association, *Guiding principles for evaluators*, 2004
<http://www.eval.org/Publications/GuidingPrinciples.asp>
- Gleick, J. *Chaos: making a new science*. New York: Penguin books 1987
- Joint commission on Standards for Educational Evaluation, *Program evaluation standards, synthesis*
<http://www.eval.org/EvaluationDocuments/progeval.html>
- Kusters, C., with S. van Vugt, S.Wigboldus, B. Williams and J. Woodhill, *Making evaluations matter: a practical guide for evaluators* 2011 <http://www.cdi.wur.nl/UK/resources/publications>
- Mintzberg, H. *Tracking Strategies: Toward a General Theory*, Oxford University Press, 2007.
- Patton, M. Q. "Developmental Evaluation," *Evaluation Practice* 15(3): 31, 1996.
- Patton, M. Q., "Developmental Evaluation." P. 116 in S. Mathison (ed.) *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage, 2005.
- Patton, M.Q. "Evaluation for the Way We Work", *Nonprofit Quarterly*, Spring, 2006.
<http://store.nonprofitquarterly.org/13evforwaywe.html>
- Patton, M.Q. *Utilization-Focused Evaluation*, 4th ed., Sage, 2008.
- Patton, M.Q., *Essentials of Utilization-Focused Evaluation*, Sage, 2012.
<http://www.sagepub.com/books/Book233973>
- Rogers, E. *Diffusion of innovations*, Glencoe: Free press, 1962
- Rosling, H. *The joy of stats*. (documentary). Rosling also makes inspiring presentations. For example
http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html
- Schulz, K. *Being wrong: adventures in the margins of error*, 2010 www.beingwrongbook.com
- Scriven, M. *Evaluation thesaurus* (4th edition) 1991 see also <http://michaelscriven.info/>
- Snowdon, D. J. and M.E. Boone, *A Leader's Framework for Decision Making*, Harvard Business Review, November 2007
- Taleb, N.N. *The black swan: the impact of the highly improbable*. New York: Random House 2007

- Westhorp, G., Prins, E., Kusters, C.S.L., Hultink, M., Guijt, I., Brouwers, J.H.A.M. *Realist evaluation: an overview*. Report from an Expert Seminar with Dr. Gill Westhorp, CDI, 2011
http://www.cdi.wur.nl/UK/resources/Publications/publications_2011/
- Westley, F., Zimmerman, B. and Patton, M.Q. *Getting to Maybe: How the World Is Changed*, Random House Canada, 2006

Appendix 1

Biography

Michael Quinn Patton

Patton is an independent organizational development and evaluation consultant. He is former **President of the American Evaluation Association**. He is the only recipient of both the *Alva and Gunnar Myrdal Award* from the Evaluation Research Society for "outstanding contributions to evaluation use and practice" and the *Paul F. Lazarsfeld Award* for lifetime contributions to evaluation theory from the American Evaluation Association. The Society for Applied Sociology honored him with the 2001 *Lester F. Ward Award* for Outstanding Contributions to Applied Sociology. He was the Gwen Iding Brogden Distinguished Lecturer at the 2008 National Conference on Systems of Care Research for Children's Mental Health.

He is the author of five evaluation books including a 4th edition of *Utilization-Focused Evaluation* (2008) and 3rd edition of *Qualitative Research and Evaluation Methods* (2002). These books have been used in over 500 universities worldwide. He is also author of *Creative Evaluation* (1987); *Practical Evaluation* (1982); and *Culture and Evaluation* (1985). He has co-authored a book on the dynamics of social innovation with two Canadians drawing on complexity theory and systems thinking: *Getting to Maybe: How the World is Changed* (Random House, 2006). His latest book is *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use* (Guilford Press, 2011).

After receiving his doctorate in Organizational Sociology from the University of Wisconsin, he spent 18 years on the faculty of the University of Minnesota (1973-1991), including five years as Director of the Minnesota Center for Social Research and ten years with the Minnesota Extension Service. He received the University's Morse-Amoco Award for outstanding teaching. He also served as a faculty member with the Union Institute & University for 16 years.

He was the keynote presenter for the launching of the Latin American Network in Peru in 2004, the African Evaluation Society in Nairobi, Kenya in 1999 and at the European Evaluation Society in Switzerland in 2000. He has twice keynoted the American, Canadian, and Australasian Evaluation Society conferences, as well as national evaluation conferences for the United Kingdom, New Zealand, Italy, Denmark, Japan, and Brazil. He is a regular trainer for the World Bank in Ottawa, the American Evaluation Association's professional development courses, and The Evaluators' Institute.

He has worked with organizations and programs at the international, national, state, and local levels, and with philanthropic, not-for-profit, private sector, and government programs. He has worked with peoples from many different cultures and perspectives. He is a generalist working across the full range of efforts at improving human effectiveness and results, including programs in leadership development, education, human services, the environment, public health, employment, agricultural extension, arts, criminal justice, poverty programs, transportation, diversity, managing for results, performance indicators, effective governance, and futuring. He uses a variety of evaluation and research methods, with special focus on mixed methods designs and analysis.

This report summarises the discussions and presentations of the Expert Seminar 'Developmental Evaluation', which took place in Wageningen on March 22, 2012. The Expert Seminar was organised by the Wageningen UR Centre for Development Innovation in collaboration with Learning by Design and Context, international cooperation.

More information: www.cdi.wur.nl

