Futures xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Futures

journal homepage: www.elsevier.com/locate/futures

Meaning and power in the design and development of policy experiments

Sreeja Nair [a], Michael Howlett [a,b,*]

[a] Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore
[b] Department of Political Science, Simon Fraser University, Burnaby, BC, Canada

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Policy experiments
Uncertainty
Policy pilots
Policy implementation
Politics
Evaluation
Diffusion

ABSTRACT

Poorly designed or implemented policies impede a society's ability to adapt to changes in the policy environment. In order to avoid such situations, pilot projects and other forms of policy experiments can and are often used to test new approaches before their full-scale roll-out. Policy experimentation can provide meaning to policymaking by helping in framing or projecting the future, deriving alternate response strategies and monitoring any changes in the policy environment. At least in theory, the small scale and experimental nature of pilots can encourage policy innovations and reduce policy risks. The discussion in this paper examines three key challenges to policy experimentation all of which centre on questions of meaning in terms of understanding the future, and power in terms of the ability of governments to design and implement such actions. These are (1) the influence of politics and key stakeholders therein on the design and evaluation of experiments, (2) problems in the technical evaluation of policy experiments and (3) problems encountered in the diffusion of experiments and retaining the lessons drawn from them.

1. Uncertainty and the need for policy experiments

A critical challenge that often faces decision-makers and planners in designing policies is doing so under conditions of uncertainty, that is, with limited and sometimes no information about the likely future policy environment. In this sense, as Swanson et al. (2010) argued, much policymaking in the 21st century is akin to gardening in so far as it is "muddy, attentive and experiential, because we really do not know what growing conditions will prevail".

Outputs, outcomes and target group behaviour are only a few of the many aspects of policy-making which are uncertain. Day and Klein (1989) note while most government policies are crafted in response to events that are 'reasonably predictable', policy events can also be, (1) unpredictable, 'unforeseen' and 'unprojectable', (2) catastrophic and (3) events where interpretation of uncertainty signals is convoluted because of associated moral and social issues.

Coping with the challenges posed by uncertainty involves assessing degree or level in any given policy-making situation. Some of these uncertainties stem from a lack of knowledge of cause and effect relationships between policy interventions and outcomes. Failing to correctly identify the bounds and range of these uncertainties is a major cause of policy over and under-reaction (Maor, 2012, 2014) and over and under-design, and uncertainties must be correctly understood and diagnosed by policy-makers in specific circumstances if policy failures are to be avoided, both in the short and long-term.

* Corresponding author at: Department of Political Science, Simon Fraser University, Burnaby, BC, Canada. Tel.: +1 77888888.
E-mail address: howlett@sfu.ca (M. Howlett).

http://dx.doi.org/10.1016/j.futures.2015.02.008
0016-3287/

Faced with this kind of uncertainty governments often resort to experiments or pilot projects which can generate policy-relevant knowledge with respect to the future policy environment, identify appropriate policy responses and allow policies to effectively 'adapt' with the rate and level of projected future change in their environments (Nair & Vreugdenhil, in press; Swanson & Bhadwal, 2009).

But not all environments change as rapidly as others and not all uncertainties demand the same response, so policy experiments themselves must be carefully designed and calibrated (Nair & Howlett, in press). Policy experimentation can provide meaning to policymaking by helping in framing or projecting the future, deriving alternate response strategies and monitoring changes in the policy environment. At least in theory, the small scale and experimental nature of pilots can encourage policy innovations and reduce policy risks. The discussion in this article examines three key challenges to policy experimentation all of which centre on questions of meaning in terms of understanding the future, and power in terms of the ability of governments to design and implement such actions. These are (1) the influence of politics and key stakeholders therein on the design and evaluation of experiments, (2) problems in the technical evaluation of policy experiments and (3) problems encountered in the diffusion of experiments and retaining the lessons drawn from them.

2. Types of uncertainty and the benefits of policy experiments

The recent uncertainty classifications by Kwakkel, Walker, and Marchau (2010), Walker et al. (2003) and Walker, Marchau, and Swanson (2010) can be used to develop a set of propositions for policy-making dealing with a range of future scenarios and knowledge gaps related to them. Walker et al. (2010) usefully identify five levels of uncertainty which experiments can address. These include Level 1 shallow or parameter uncertainty where multiple alternative states representing the system with specific probabilities are present and Level 2 medium or fuzzy uncertainty where multiple alternatives exist within a scenario but these can be ranked based on the 'perceived likelihood' of their occurrence. Level III situations are where different scenarios exist but can still be ranked in terms of their likelihood. Level IV uncertainty represents a more complex situation in which multiple plausible alternative scenarios can be enumerated without being able to rank the alternatives in terms of their perceived likelihood. Finally, in the most complex Level V situations there is an inability to present or agree upon a full range of possible alternative scenarios and the "strong possibility of being surprised" must be acknowledged Walker et al. (2013).

Testing policy responses through active experimentation is a key tool for policy-makers attempting to deal with these different kinds of uncertainty through approaches ranging from improved knowledge of target group behaviour under different policy interventions, to improved projections of the likelihood of future states of events. Transposing the idea of "controlled experimentation" – a concept that is well-acknowledged in the social and natural sciences - into policy research has received much attention by policy scholars and practitioners alike attempting to deal with uncertainties in the policy environment (Anderson, 1975).

Policy experiments including pilots can play an important role in policy-relevant knowledge production which could include early evaluation of the impacts of new initiatives such as subsidies and incentive programmes for education or training of economically disadvantaged communities and utilisation of data generated via these pilots (Stromsdorfer, 1985). The development sector, for example, has often conducted experiments to evaluate alternative strategies and accordingly allocate resources to those that emerge as most feasible in promoting development goals (Rondinelli, 1993). Experiments have also served as a source of evidence for policy-making in many sectors including education, healthcare, environment, social welfare among others (Bennion, 2011).[1] The role of social learning fostered through experimentation has also been found to be crucial in overcoming 'lock-in' of unsustainable trajectories and enabling restructuring of current social–technical systems to transition towards sustainable pathways (Rotmans et al 2001).

"Promoting variation" (Swanson & Bhadwal, 2009) by crafting multiple policy alternatives can make the emergence of a successful solution more likely (Cummings, 2013). But enhanced experimentation and learning can be instrumental in "keeping pace with the dynamic drivers and expressions of risk" in a changing policy environment (ÓBrien & Sygna, 2013), only if the experiments themselves are designed to address the level of uncertainty involved in a specific case.[2]

[1] More recently, policy pilots have also gained much attention in the transition management literature (Rotmans, Kemp, & van Asselt, 2001; Van den Bosch & Rotmans, 2008) specifically those intended to generate long-term societal change. Many experiments that aim towards enabling sustainable transitions are largely technological. Social learning requires open and flexible networks while learning derived from technological experimentation largely happens in 'closed networks' (Bos & Brown, 2012). There is little empirical analysis on the effects of policy experiments however, in particular how their design influences their potential as "learning incubators" (McFadgen, 2013). It is usually just assumed that experimentation is an important instrument to support sustainability transitions by providing a 'space' for social learning to occur (Bettini, 2014).

[2] Acknowledgement of the limitations to "rational calculation, planning and forecasting" under conditions of high uncertainty have drawn attention to the role of experiential learning and experimentation over time for adapting to change (March & Olsen, 1975). The idea of adaptive policies has been discussed widely in the context of decisions for the long-term such as infrastructure planning and climate change (Buurman, Zhang, & Babovic, 2009; Gersonius, Ashley, Pathirana, & Zevenbergen, 2013; Giordano, 2012; Ranger & Garbett-Shiels, 2011; Swanson & Bhadwal, 2009; Walker et al., 2001). The process of adaptive policymaking can be (1) passive and operate on the available scientific information till new knowledge comes up or (2) active and consciously experiment with policy alternatives to identify better strategies as the new conditions emerge (Walter, 1992). Apart from technical or 'hard' knowledge, pilots may also produce 'soft' types of knowledge such as knowledge on management processes, actor preferences, policies and implementation designs. Learning is established through formal means (monitoring and evaluation), but (social) learning between stakeholders during the process is often at least as important for the future of the innovation or project (Nair & Vreugdenhil, in press).

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008

3. Designing and assessing policy experiments: issues of meaning and power

Policy experimentation can be considered as a form of 'predictive method' and a part of ex-ante evaluation deployed by various agencies, including the government, to pre-test different programmes and policies for their likely impacts; testing processes of implementation and stakeholder acceptability prior to launching these fully or on a large-scale. The underlying motivation in designing such policy experiments is that these will provide results that are largely indicative of what outputs, outcomes and challenges can be expected when programmes and policies based on results of these experiments are implemented fully (Nair & Vreugdenhil, in press).

Policy experimentation thus potentially provides meaning by helping in framing or projecting the future, deriving alternate response strategies and monitoring changes in the policy environment. One of the ways of doing this, for example, is using strategic foresight exercises that attempt to integrate multiple perspectives and methods for identifying current and emerging issues and trends and helping assess policy options for attaining a desired future (Habegger, 2010). Foresight can be instrumental for environmental planning by providing insights about a range of futures of social-ecological systems and critical thresholds, and thus aid in anticipatory planning to avoid adverse impacts (Bengston, Kubik, & Bishop, 2012) and informing policy by enhancing the knowledge base for thinking about and designing policies (Da Costa, Warnke, Cagnin, & Scapolo, 2008). Adaptive foresight is an emerging field of study that aims to shape individual and collective responses to future scenarios by enhancing the "rational basis of decision-making" (Eriksson & Weber, 2009). It combines elements of adaptive planning into conventional foresight methods and enables individuals and groups to develop response strategies by assessing the range of assumptions and limitations of different scenarios including pilot projects and experiments to test concepts and instrument combinations (Eriksson & Weber, 2009).

Policy pilots can form a common and important form of policy experimentation in conjunction with scenario forecasting and involve the introduction of major government policies or programmes at a "controlled small-scale" (Weiss, 1975) or phase-wise manner, "allowing them to be tested, evaluated and adjusted before being rolled out nationally" (Cabinet Office, 2003). Planning "well-designed pilots" alongside a fully-functioning policy can help test a policy's performance along with identification of emerging issues and make necessary policy adjustments (Swanson & Bhadwal, 2009). A pilot in this sense can form an important first step of regular policy monitoring and evaluation over the long-term (Cabinet Office, 2003) But this is possible only if several common barriers to understanding and designing policy experiments can be overcome.

4. Barriers and challenges to effective experimentation

Pilots can aid in policy appraisal (Turnpenny, Radaelli, Jordan, & Jacob, 2009) and provide useful insights for dealing with complex policy issues and high uncertainty (Vreugdenhil, Slinger, Thissen, & Ker Rault, 2010). The small-scale and experimental nature of pilots can encourage policy innovations (Cabinet Office, 2003),[3] and policymakers are thus often urged or consider pilot projects and other forms of policy experiments in order to test new policy and programme approaches (Martin & Sanderson, 1999; Vreugdenhil et al., 2010).

However there are several challenges posed to realizing the benefits of policy experimentation in practice which need to be overcome if success is to be achieved in this area. The factors leading to success of a pilot and the links between pilot diffusion and transition to policies are not well-established, theoretically or empirically. Furthermore, the study of the impact of pilots on policy development is often limited to 'learning from failure' (Vreugdenhil, Frantzeskaki, Taljaard, Ker Rault, & Slinger, 2009).

Some scholars have suggested that 'how pilots work' signifies a more useful investigation rather than 'whether it works', especially when experiments are considered as 'prototypes' for future policies (Chitty, 2000). In the context of urban water sustainability, for example, Farrelly and Brown (2011) examined eleven local-scale experiments in Australian cities and found sustainable transitions to urban water management required changes in underlying culture and beliefs along with structural reforms. The role of 'bridging organizations' was found to be critical if learning from local-scale experiments was to inform future policy and practice.

In another example, van der Heijden (2014) explored how the content and process of policy-design experiments affect their outcomes by studying 31 experiments for sustainable building design in Australia, the Netherlands, and the United States. Results of the study again showed a mismatch between recognition of these experiments as good practices and their actual translation into practice – a mismatch that was attributed to fault in the process of policy-design rather than content. A key factor that influenced the adoption of these experiments in practice related to the extent of financial risk involved. Additionally, the multiplicity of several parallel design experiments resulted in "fatigue" or "confusion" among participants of experiments and created competition between experiments when it was perceived, rightly or wrongly, that only some experiments would be allowed to scale-up. This can undermine learning in many ways.

[3] The idea of "design experiments" is receiving increasing attention and presents a useful way of assessing policy innovations and their impact as the innovations develop further. Evaluation of these design experiments however differs from the regular evaluations of randomised control trials, as the focus here is on the quality of design of an intervention and how it can be adjusted over time for achievement of favourable results through learning and feedback (Stoker & John, 2008).

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008

Three of the key barriers to effective experimentation found in the literature are summarised in Table 1 and set out in more detail below. While 'meaning'-related issues relate to how and why the future is interpreted; issues related to 'power' deal with how and why subsequent courses of action are chosen.

4.1. Challenge no. 1: the politics of policy experiments

Political factors can both facilitate and impede the design, development and evaluation of successful policy experiments. Policy experimentation for example, can raise issues of fairness in investing public resources only on certain fractions of the society covered by a pilot (Stoker, 2010). Thus in the late 1980s and early 1990s policy pilots came under scrutiny in many countries, especially developing ones, as these were often seen as being 'donor-driven', dependent on external aid and less focused on local priorities and engagement than on specific external donor priorities. Assessing the success or failure of the results of such efforts also proved similarly problematic since they involved the possibility of losing donor support (PHR, 2004).

It is also the case that policymakers may remain hesitant towards 'accepting uncertainty' and its risks of failure (Howlett, 2012) despite the (successful) results of pilot results. Majumdar and Mukand (2004), for example, have modelled the impact of electoral liabilities on a government's decision to learn through policy experimentation by focusing on the impact problem areas have on their reputation. They argue that new governments may initially tend to experiment with new policies but over time become more conscious of their "reputational stake" in policies attributable to their tenure in office and consequently exercise more restraint in experimentation that may imply failure of their earlier efforts. In a related attempt to model why long-term policy preferences of policymakers are often contrary to rational expectations Callander and Hummel (2014) argue that though political power is only held by policymakers temporarily, they can sometimes extend their influence for a long-time through the conduct of experiments. Policymakers can do so by influencing the "informational environment" of their successors enhancing the longevity for their preferred policy choices through "pre-emptive policy experimentation" (Callander & Hummel, 2014).

Much political disagreement is over beliefs or ideology rather than outcomes and these can be brought into the design and evaluation of experiments simply through their conduct. Additionally, if the policy change that is promoted and involved in experiments poses significant new costs, it may increase the 'stickiness' of existing policies, despite what was indicated through outcomes of the experiment (Callander, 2011).

Institutional arrangements can also interfere with experimentation. The existence of federal or multi-level governance arrangements can play a significant role here when aspects of the subjects under investigation cross-over jurisdictional lines of authority and responsibility. China's economic transformation, for example, has been marked by policy experiments wherein central government encourages local policymakers to explore "new ways of problem-solving" and provide insights for national policy formulation (Heilmann, 2008). Recent studies of urban housing policy changes in China, however, reveal that the strong control from the central government on the content and process of experimentation at the local level acts as a deterrent for novel solutions to emerge via active experimentation (Mei & Liu, 2013).

Table 1
'Meaning' and 'power'-related challenges for using experimentation as a tool for policymaking under uncertainty (compiled from various sources).

Challenges	Issues of 'Meaning'	Issues of 'Power'
Politics of policy experiments	- Where to experiment and why?	- Hesitation towards accepting uncertainty and risk of failure - Hesitation in changing from policy status quo, especially when additional costs are involved - Power struggle between central and local levels to control emergence of novel solutions via active experimentation - Experiments launched to mask/delay politically unpalatable policy reforms
Evaluation of policy experiments	- Difficulties in measuring and attributing causality and isolating policy outcomes that can be solely attributed to experiments - How and when to conduct the evaluation?	- Hastening evaluation to obtain 'evidential support' for pre-determined policy solutions - Bias towards 'prototyping' versus true experimentation to try what works - Some experiments have long gestational periods before results become apparent, going beyond administrative or career timelines - Discounting future policies against 'less-than optimal' policies choices today
Diffusion of policy experiments	- Scanty and fragmented evidence on process of diffusion of experiments and influencing factors - Design and results of experiments are often heavily bound by context-specificity; thus there are limitations to exact replicability of experimental outcomes when these are scaled-up	- Diffusion of experiments influenced by multiple stakeholders - Stakeholder attitude and support for scaling up can change over temporal and spatial scale

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008

In a worse case situation, this last factor may lead pilot projects to be offered up simply as an excuse for policymakers to delay large-scale policy reforms beyond their term in office and hence avoid action that might be politically unpalatable, at least in the short term (PHR, 2004). For example controversial pilots might not be adopted immediately but delayed until "the political mood is ripe for a more enduring course of action" (Jann & Wegrich, 2007). Hence, sometimes instead of enhancing evidence-based policy-making, pilot projects may be used as tools for conflict avoidance.

4.2. Challenge no. 2: the evaluation of policy experiments

By helping to evaluate the potential effectiveness of a policy innovation ex-ante, policy experimentation in theory allows policymakers to better understand the effects of a policy intervention (McFadgen, 2012). Positive results from a pilot evaluation, for example, at least theoretically can help build consensus for a project (Duflo, 2004).

But isolating the actual impacts resulting from a pilot from other factors can be very difficult in practice, especially when such evaluations are heavily politicised, such as when they address deep-rooted social and economic issues such as unemployment, homelessness or crime, among others.[4] Political pressure can hasten the process of evaluation of pilots in a premature bid to obtain "evidential support" for the implementation of certain already favoured decisions (Sanderson, 2002).

More prosaically, a common criticism of experiments is in terms of how precisely they measure and attribute causality. Attributing absolute causal inference in policy experimentation especially where complex systems are involved is a difficult task. While randomised experiments are closer to providing precision in causality attribution, it can be difficult if not impossible to accurately evaluate a proposed policy's effectiveness when only quasi-experiments or sometimes poorly designed or idiosyncratic interventions are used (McFadgen, 2012).

Pilot studies are often not reported widely and when they are, they usually do not go beyond justification of the research methods or tools. Calling for better documentation of pilots whether these are successes or failure, Van Teijlingen and Hundley (2001) argue "well-designed and well-conducted pilot studies can inform us about the best research process and occasionally about likely outcomes".

Policy experiments and pilots are also usually conducted over short time-periods, often involving "one-off evaluations" to measure success (Stoker & John, 2008). Evaluation of policy experiments and pilots also depends on what the 'real purpose' of the experiment was in the first place. For example there are cases where pilots are not representative of the context in which they are applied and are sustained only because there is significant political will to ensure that the 'pilot works'. This is then more a case of "prototyping" rather than experimentation i.e. there is more emphasis on evaluating how the pilot works rather than whether it works (Sanderson, 2002).

Moreover, the timing and time needed for appropriate piloting of policies and subsequently their evaluation also vary significantly, crossing in some cases beyond electoral or other administrative or career timelines. Some pilot initiatives such as those related to the education or health sector, for example, may have very long gestation periods before producing a measurable effect. In such cases, a "call for action tends to overwhelm the call for well-grounded prior evidence" (Cabinet Office, 2003).

And even when the intent of experimentation for long-term policies is to reveal critical information in terms of 'expected utility' of the policy action today as well as over the long-term, there are trade-offs that citizens as well as policymakers have to make in terms of discounting future streams of costs and benefits against present day ones. Sometimes these trade-offs could mean settling for less-than optimal policies which are 'good enough' today (Callander, 2011), regardless of the results of experiments which might show greater benefits in the future.

4.3. Challenge no. 3: 'Diffusion' of policy experiments

A third major potential barrier exists beyond design and evaluation issues when the desire exists to 'scale-up' the results of an experiment or pilot. Diffusion (i.e. expansion or continuation) of pilots can occur via replication of the pilot into other or similar pilots and 'scaling up' into policies or bigger pilots. In general empirical evidence about pilots and the process of their diffusion is lacking (Vreugdenhil et al., 2009).

Many factors can influence scaling dynamics, including the pilot design and the context. This includes factors such as the type of stakeholders involved, the availability of knowledge and resources; the choice of scale and the choice for pilot sites; the mode of governance that influences the nature of stakeholder engagement and learning; the level of innovativeness of the pilot and how it converges or diverges from the current policy context; flexibility to make changes to adapt to local conditions and finally the timing of the strategy for pilot diffusion (Vreugdenhil et al., 2009). The scaling up of policy experiments has been argued to be a "craft rather than science", that is, apart from being a technical process it is also highly driven by the interests, behaviour and attitudes of stakeholders (Spicer et al., 2014). For example, if there is top-down

[4] Many pilots have been studied and documented in isolation, making them a context-specific subject and generalisation of their learning effects has not been attempted. Hoffmann (2011) argues that rather than a piece-meal approach, it is more valuable to look at experiments collectively along with their combined impacts. He collates 58 independent initiatives for action against climate change ('climate governance experiments') – that include initiatives by cities, provinces and states, citizen groups, and corporations globally. He argues system-level characteristics are 'emergent' in nature and cannot be estimated by simply examining individual efforts. Finds the nevertheless finds the patterns arising from individual 'climate governance experiments' to constitute the "contours of the nascent experimental governance system".

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008

pressure for rapid scaling of the pilots, "implementation gets ahead of capacity, and local partners may become frustrated and disengaged" (PHR, 2004).

There can be several "channels of diffusion" or actors who can influence the diffusion process (see Fig. 1) (Vreugdenhil et al., 2009). Temporal aspects of the experiments are also critical as the champions of specific experiments may well have moved on before these have been completed and thus the political or administrative support they had at the outset may have disappeared by the time they are finished (PHR, 2004). In such cases, diffusion can face impediments if a clear strategy for diffusion management is entirely absent, is done poorly or where there is widespread opposition from some critical stakeholders (Vreugdenhil, 2010).

5. Conclusions and avenues for future research

Policy experiments are potentially a very powerful tool to help policymakers deal with uncertain futures through pre-testing concepts and instruments. Policy experiments including pilot projects have been deployed by policymakers and practitioners in many sectors to help test policies before these are launched on a full scale.

To some extent, most policies can be considered as experiments, with the aim of "promoting continual learning and adaptation in response to experience over time" (Swanson et al., 2010) and through trial and error, especially when there is imperfect information (Callander, 2011). In this sense, governments have a critical role to play as "laboratories" and policymakers must continually monitor and learn from experimentation (Moynihan, Shipan, & Volden, 2012). Experiments have also been useful as a source of evidence for policymaking. However there are challenges in realisation of the benefits of policy experimentation in practice, both in relation to 'meaning' in terms of understanding the future, and 'power' in terms of undertaking related policy action.

As the discussion above has noted, successful design, implementation and scaling up of experiments is not an automatic process. Firstly, the politics surrounding experimentation and presence of diverse stakeholder perspectives pose continual challenges to effective experimentation. Temporal aspects such as when to experiment and when to evaluate, and spatial aspects such as where to experiment are all vulnerable to political influence and interference. Furthermore, the treatment and acceptance of uncertainty as well as selection of a course of policy action between that which is desirable and that which is politically preferred, conducive or acceptable are often choices that are politically motivated, irrespective of the evidence provided by policy experiments.

In order to realize the benefits of policy experimentation for designing policies despite uncertainty about the future environment, strategies to overcome the challenges noted above will need to be considered. Careful thought and deliberation must accompany the construction and design of policy experiments just as it must full-scale policies. While issues of 'meaning' can be better addressed by enhancing the knowledge base for informed policymaking, issues of 'power'

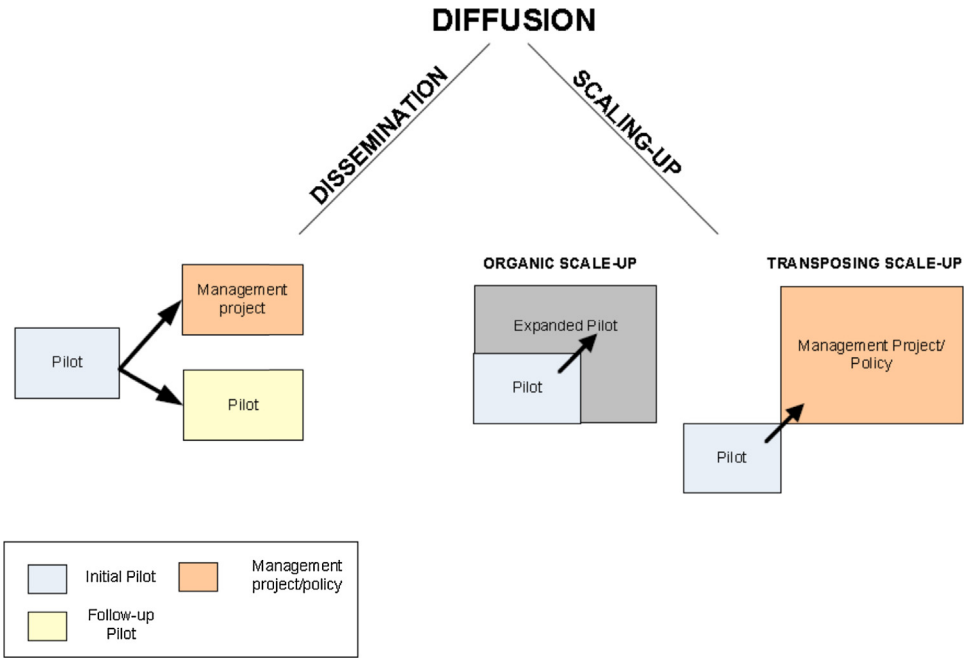

Fig. 1. Pilot diffusion through replication and scaling-up (Vreugdenhil et al., 2009).

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008

need higher levels of engagement and dialogue between stakeholders to agree on preferred and suitable courses of policy action for the future, based on the insights gained from well-crafted policy experiments.

Some corrective practices and design principles have been set out to help deal with the three barriers cited above. Moynihan et al. (2012), for example, caution against centralising policies that have a complex and 'emergent' nature over time and argue allowing experimentation and learning to occur in a decentralised manner is the best way for sustenance of such policies over time. Polycentric governance arrangements can enable innovation and experimentation, deploy a larger diversity of response and address policy problems at multiple scales compared to centralised structures (Cummings et al., 2013). The increased use of experimentation as a tool can also help make governance more collaborative rather than top-down in nature (John, 2013). Secondly, the answers to questions such as, how to evaluate, evaluation by whom and for whom which have plagued evaluations of full policies also hold relevance for evaluation of policy experiments or pilots. Lastly, diffusion and scaling up of experiments is critical for realisation of its objectives. Scaling up depends on the acceptability of experiments to potential beneficiaries and the political support that these experiments receive. Given the often inevitable context-specificity of many policy experiments, a key challenge in scaling up is to tease out those design characteristics of policy experiments that are context specific and those that are universal and "assure the universal elements are applied, but leaving room for local adaptation; and to evaluate, learn and change the approach as scaling up proceeds" (Hartmann & Linn, 2007).

References

Anderson, A. B. (1975). Policy experiments: Selected analytical issues. *Sociological Methods & Research, 4*, 13.
Bengston, D. N., Kubik, G. H., & Bishop, P. C. (2012). Strengthening environmental foresight: Potential contributions of futures research. *Ecology and Society, 17*(2), 10.
Bennion, E. (2011). Experiments. In J. Ishiyama & M. Breuning (Eds.), *21st century political science: A reference handbook* (pp. 523–532). Thousand Oaks, CA: SAGE Publications, Inc.
Bettini, Y. (2014). Understanding institutional capacity for urban water transitions. *Technological Forecasting and Social Change*. http://dx.doi.org/10.1016/j.techfore.2014.06.002
Bos, J. J., & Brown, R. R. (2012). Governance experimentation and factors of success in socio-technical transitions in the urban water sector. *Technological Forecasting & Social Change, 79*, 1340–1353.
Buurman, J., Zhang, S., & Babovic, V. (2009). Reducing risk through real options in systems design: The case of architecting a maritime domain protection system. *Risk Analysis, 29*, 366–379.
Cabinet Office (2003). Trying it out: The role of 'pilots' in policy-making. In *Report of a review of government pilots*. Strategy Unit, Government of the United Kingdom.
Callander, S. (2011). Searching for good policies. *American Political Science Review, 105*(November (4)).
Callander, S., & Hummel, P. (2014). Preemptive policy experimentation. *Econometrica, 82*(4), 1509–1528.
Chitty, C. (2000). Why pilot? Paper presented to the Royal Statistical Society Conference. *The evaluation of social and economic policies*. RSS.
Cumming, G. S., Olsson, P., Chapin III, F. S., & Holling, C. S. (2013). Resilience, experimentation, and scale mismatches in social-ecological landscapes. *Landscape Ecology, 28*(6), 1139–1150.
Da Costa, O., Warnke, P., Cagnin, C., & Scapolo, F. (2008). The impact of foresight on policy-making: Insights from the FORLEARN mutual learning process. *Technology Analysis & Strategic Management, 20*(3), 369–387.
Day, P., & Klein, R. (1989). Interpreting the unexpected. The case of AIDS policymaking in Britain. *Journal of Public Policy, 9*, 337–353.
Duflo, E. (2004). Scaling up and evaluation. *Annual World Bank Conference on Development Economics*. The International Bank for Reconstruction and Development/The World Bank.
Eriksson, A. E., & Weber, M. K. (2009). Adaptive foresight: Navigating the complex landscape of policy strategies. *Technological Forecasting & Social Change, 75*, 462–482.
Farrelly, M., & Brown, R. (2011). Rethinking urban water management: Experimentation as a way forward? *Global Environmental Change, 21*, 721–732.
Gersonius, B., Ashley, R., Pathirana, A., & Zevenbergen, C. (2013). Climate change uncertainty: Building flexibility into water and flood risk infrastructure. *Climatic Change, 116*, 411–423.
Giordano, T. (2012). Adaptive planning for climate resilient long-lived infrastructures. *Utilities Policy, 23*, 80–89.
Habegger, B. (2010). Strategic foresight in public policy: Reviewing the experiences of the UK, Singapore, and the Netherlands. *Futures, 42*, 49–58.
Hartmann, A., & Linn, J. F. (2007). Scaling up: A path to effective development. In *2020 focus brief on the world's poor and hungry people*. Washington, DC: IFPRI.
Heilmann, S. (2008). From local experiments to national policy: The origins of China's distinctive policy process. *The China Journal,* (59), 1–30.
Hoffmann, M. J. (2011). *Climate governance at the crossroads: Experimenting with a global response after Kyoto*. Oxford University Press.
Howlett, M. (2012). The lessons of failure: Learning and blame avoidance in public policy-making. *International Political Science Review, 33*(5), 539–555. http://dx.doi.org/10.1177/0192512112453603
Jann, W., & Wegrich, K. (2007). Theories of the policy cycle. In F. Fischer, G. J. Miller, & M. S. Sidney (Eds.), *Handbook of public policy analysis: Theory, politics, and methods*.
John (2013). Experimentation, behaviour change and public policy. *The Political Quarterly, 84*(2).
Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2010). Classifying and communicating uncertainties in model-based policy analysis. *International Journal of Technology, Policy and Management, 10*(4), 299–315.
Majumdar, S., & Mukand, S. W. (2004). Policy gambles. *The American Economic Review, 94*(4), 1207–1222.
Maor, M. (2012). Policy overreaction. *Journal of Public Policy, 32*(3), 231–259. http://dx.doi.org/10.1017/S0143814X1200013X
Maor, M. (2014). *Risk and policy underreaction* (working paper) Jerusalem: Department of Political Science, Hebrew University Available online at: www.moshemaor.net
March, J. G., & Olsen, J. P. (1975). The uncertainty of the past: Organizational learning under ambiguity. *European Journal Of Political Research, 3*, 147–171.
Martin, S., & Sanderson, I. (1999). Evaluating public policy experiments: Measuring outcomes, monitoring progress or managing pilots? *Evaluation, 5*(3), 245–258.
McFadgen, B. (2012). *What is the value of twisting the lion's tail? Evaluating the use of policy experiments in adaptation governance and how they can facilitate learning* The Netherlands: Institute for Environmental Studies.
McFadgen, B. (2013). Learning from policy experiments in adaptation governance. In *Paper presented at the international conference on public policy, Grenoble, France*.
Mei, C., & Liu, Z. (2013). Experiment-based policy making or conscious policy design? The case of urban housing reform in China. *Policy Sciences, 47*(3), 321–337. http://dx.doi.org/10.1007/s11077-013-9185-y
Moynihan, D. P., Shipan, C. R., & Volden, C. (2012). Policy diffusion: Seven lessons for scholars and practitioners. *Public Administration Review, 72*(6), 788–796.
Nair, S., & Howlett, M. (in press). Dealing with failure over the long-term: Policy synchronization and adaptive policy design. *Public Policy and Administration* (in press).

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008

Nair, S., & Vreugdenhil, H. (in press). Pilots as predictive methods. In *Encyclopaedia of public administration and public policy*. Taylor and Francis (in press).
ÓBrien, K., & Sygna, L. (2013). Responding to climate change: The three spheres of transformation. *Proceedings of Transformation in a Changing Climate*. University of Oslo ISBN: 978-82-570-2000-1.
PHR (Partners for Health Reformplus) (2004). *The role of pilot programs: Approaches to health systems strengthening*. Bethesda, MD: PHRplus Abt Associates Inc.
Ranger, N., & Garbett-Shiels, S.-L. (2011). *How can decision-makers in developing countries incorporate uncertainty about future climate risks into existing planning and policymaking processes?* London: Centre for Climate Change Economics and Policy, Grantham Research Institute.
Rondinelli, D. (1993). *Development projects as policy experiments: An adaptive approach to development administration* (2nd ed.). New York: Routledge.
Rotmans, J., Kemp, R., & van Asselt, M. B. A. (2001). More evolution than revolution: Transition management in public policy. *Foresight, 3*(1), 15–31.
Sanderson, I. (2002). Evaluation, policy learning and evidence-based policymaking. *Public Administration, 80*(1), 1–22.
Spicer, N., Bhattacharya, D., Dimka, R., Fanta, F., Mangham-Jefferies, L., Schellenberg, J., et al. (2014). Scaling-up is a craft not a science': Catalysing scale-up of health innovations in Ethiopia, India and Nigeria. *Social Science & Medicine, 121*, 30–38.
Stoker, G. (2010). Translating experiments into policy. *The Annals of the American Academy of Political and Social Science, 628*, 47.
Stoker, G., & John, P. (2008). Design experiments: Engaging policy makers in the search for evidence about what works. *Political Studies, 57*(2), 356–373.
Stromsdorfer, E. W. (1985). Social science analysis and the formulation of public policy: Illustrations of what the president knows and how he comes to know it. In J. A. Hausman & D. A. Wise (Eds.), *Social Experimentation* (pp. 257–282). University of Chicago Press.
Swanson, D., & Bhadwal, S. (Eds.). (2009). *Creating adaptive policies: A guide for policymaking in an uncertain world*. New Delhi/Ottawa: Sage/IDRC.
Swanson, D., Barg, S., Tyler, S., Venema, H., Tomar, S., Bhadwal, S., et al. (2010). Seven tools for creating adaptive policies. *Technological Forecasting & Social Change, 77*(6), 924–939.
Turnpenny, J., Radaelli, C. M., Jordan, A., & Jacob, K. (2009). The policy & politics of policy appraisal: Emerging trends & new directions. *Journal of European Public Policy, 16*(4), 640–653.
Van den Bosch, S., & Rotmans, J. (2008). *Deepening, broadening and scaling up: Towards a conceptual framework for transition experiments KSI report*.
van der Heijden, J. (2014). Experimentation in policy design: Insights from the building sector. *Policy Sciences, 47*(3), 249–266.
Van Teijlingen, E. R., & Hundley, V. (2001). *The importance of pilot studies, social research update, issue 35*. University of Surrey.
Vreugdenhil, H. (2010). *Pilot projects in water management: Practicing change and changing practice* (Ph.D. thesis) The Netherlands: Delft University.
Vreugdenhil, H., Frantzeskaki, N., Taljaard, S., Ker Rault, P., & Slinger, J. (2009). The next step in policy transitions: Diffusion of pilot projects. *13th annual conference of the international research society for public management*. Copenhagen Business School.
Vreugdenhil, H., Slinger, J., Thissen, W., & Ker Rault, P. (2010). Pilot projects in water management. *Ecology and Society, 15*(3), 13.
Walker, W. E., Harremoes, P., Rotmans, J., van der Sluijs, J., van Asselt, M. B. A., Janssen, P., et al. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment, 4*(1), 5–17.
Walker, W., Marchau, V., & Swanson, D. (2010). Addressing deep uncertainties using adaptive policies. *Technological Forecasting & Social Change, 77*(6).. (Special Section 2).
Walker, W. E., Marchau, V. A. W. J., & ; Kwakkel, J. H. (2013). Uncertainty in the Framework of Policy Analysis. In Public Policy Analysis: New Developments. In W. A. H. Thissen & W. E. Walker (Eds.), Berlin, Germany: Springer.
Walter, C. J. (1992). Perspectives on adaptive policy design in fisheries management. In K. Jain & L. W. Botsford (Eds.), *Applied population biology* (pp. 249–262). The Netherlands: Kluwer Academic Publishers.
Weiss, C. (1975). Evaluation research in the political context. In E. L. Struening & M. Guttentag (Eds.), *Handbook of evaluation research* (pp. 13–26). CA, USA: Sage Publications.

Please cite this article in press as: Nair, S., & Howlett, M. Meaning and power in the design and development of policy experiments. *Futures* (2015), http://dx.doi.org/10.1016/j.futures.2015.02.008