

IVM Institute for Environmental Studies

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

Deliverable 2.1.1

Belinda McFadgen



This report is released by: Dave Huitema
Associate Professor



This report was commissioned by: Knowledge for Climate

IVM

Institute for Environmental Studies
VU University Amsterdam
De Boelelaan 1087
1081 HV AMSTERDAM
The Netherlands
T +31-20-598 9555
F +31-20-598 9553
E info@ivm.vu.nl

Knowledge for Climate Programmabureau
Daltonlaan 400
3584 BK Utrecht
T +31 317 48 6540
M +31 6 2120 2447
E info@kennisvoorklimaat.nl

Copyright © 2012, Institute for Environmental Studies

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the copyright holder

Contents

1	Introduction	4
2	Policy Experiments	6
2.1	What is a policy experiment?	6
2.2	What is not a policy experiment?	10
2.3	Concerns with the use of policy experiments	10
3	Learning	14
3.1	What is learning?	14
3.2	Adopted typology of learning	15
3.3	Evaluation learning within the policy domain	17
4	Relationship between Experiments and Learning	18
4.1	Learning outcomes form within the experiment	18
4.2	Learning outcomes within the policy domain	18
5	Conclusion	20
	References	21

1 Introduction

The social and ecological impacts of climate change are becoming more apparent and society is being urged to adapt more extensively than current efforts (IPCC Fourth Assessment Report, 2007). Adaptation to climate change is reactive but it has an anticipatory dimension because it needs to be based on an assessment of future conditions (Adger et al, 2005). Therefore, society must restructure its existing governance mechanisms in order to meet the challenge of successful adaptation, and one way to do this is to develop new governance arrangements (Termeer et al, 2011). Governance in this instance is defined as “the interactions between public and/or private entities ultimately aiming at the realisation of collective goals” (Jordan et al., 2010) and adaptation governance focuses on building effective arrangements between those entities. In realising collective goals adaptation governance calls on actors from all strata in society- governments, businesses, and civil society actors, and is open to using both regulatory and non-regulatory steering mechanisms (ibid.).

Theoretically, policy experiments make connections across boundaries and recognise new linkages of people and ideas (Huiteima et al, 2009; Termeer et al, 2011) so they are expected to play a role in realising adaptation governance by connecting state and non-state actors, different policy sectors and different institutional levels. This research project aims to explore both conceptually and empirically the practical implications of using policy experiments to connect actors and generate evidence for policy making from the perspective of learning, which will add to the growing body of scientific knowledge about adaptation governance.

However, the recommended use of policy experiments is not confined to adaptation governance. A review of the literature finds the concept of experimenting with policy alive and thriving in the natural resource management (e.g. Holling, 1978; Cooke et al, 2004; Walters and Holling, 1990), transitions management (e.g. Raven et al, 2008; Berkhout et al, 2010; Smith, 2007), and adaptive governance (e.g. Folke et al, 2005; Olsson et al, 2006) literatures, and a more divided opinion on their use in the policy sciences literature (e.g. Fischer, 1995; Vedung, 1997; Campbell, 1998; Greenberg et al, 2003; Stoker, 2010). This scope is initially rather daunting to capture; however, once “tamed” it provides a rich and diverse theoretical base from which to draw inspiration for this project. For instance, an experiment can help decrease uncertainty and help cope with system complexity (Folke et al 2005), create protected space for innovations to emerge (Raven et al, 2008), support knowledge acquisition and learning (Voss and Bornemann, 2011) and test policy ideas against experience (Sanderson, 2009).

Coming at this study from a learning perspective is logical because experiments and learning are often discussed in-tandem; for example, adaptive management experiments allow for “learning-by-doing” (Lee, 1999) and social policy experiments provide the evidential basis for learning how to improve action (Sanderson, 2009). The literature is very positive when it comes to the discussion about policy experiments and learning: experimentation is considered a “key mechanism” and can quicken the pace of learning (Armitage et al 2008); experiments are expected to create room for systematic learning (Maarleveld and Dangbegnon, 1999); experiments function as a boundary object where participants can learn with and from each other (Huiteima et al 2009); an Experimenting Society is a “learning society” (Campbell 1998).

In summary, it is apparent from the literature that policy experimentation and learning are concepts that logically pair together. This makes it all the more surprising that

there have not been any large empirical studies done on their relationship. Studies that analyse on a case-by-case basis exist; for instance, Armitage et al (2008) analyse an active experiment in Vietnam for its learning outcomes, Lee (1999) writes about the impact of experimentation on learning-by-doing in adaptive management, Raven et al (2008) assess the impact of local experiments on field-level learning; but no studies have been found that systematically study a broad range of policy experiments from a design perspective and how this impacts learning.

I aim to fill this knowledge gap. In the following section I highlight how experiments are designed in different ways for a variety of purposes, so in order to study experiments comparatively I focus on their institutional design; i.e. who participates, what positions are available, what information is created and shared, who makes the decisions, etc. It is hypothesised that an experiment’s design features influence the learning that occurs both within the experiment and within the wider policy domain.

2 Policy Experiments

2.1 What is a policy experiment?

Despite an extensive literature review (e.g. Folke et al, 2005, Olsson et al, 2006, Raven et al, 2008, Berkhout et al, 2010, Smith, 2007, Holling, 1978, Cooke et al, 2004, Walters and Holling, 1990, Peters, 1998, Campbell, 1998, Stoker, 2010, Vedung, 1997, Fischer, 1995, Pahl-Wostl, 2006, Ostrom 1999) I did not find any dominant or consistent definition of a policy experiment. Rather, the concept of experimenting with policy takes different forms; ranging from field experiments that follow strict experimental methods (i.e. with randomly chosen subjects and control groups) to an open approach where any new policy is treated as an hypothesis. This section starts with a discussion on the range of policy experiments found in the literature and highlights the similarities and differences between the conceptualisations, before switching the focus onto the political, bureaucratic and social dynamics of experiments.

The ordinal diagram below (Figure 2.1) aims to illustrate the diverse applications of a policy experiment, as described in the literature. On the horizontal line, experiments vary according to their commitment to measuring the causal impact of a policy. The vertical line displays whether the experiment tests a completed policy or develops a policy idea/innovation. These simple delineations capture the essence of policy experiments, as explained in more detail below.

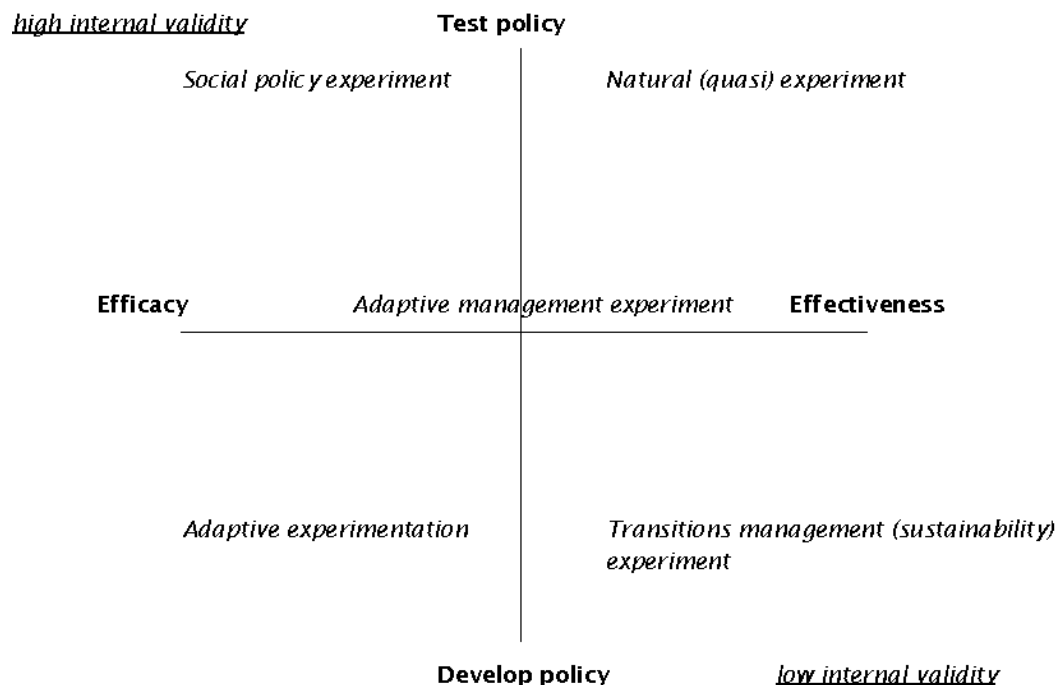


Figure 2.1 Ordinal diagram to illustrate types of policy experiments.

2.1.1 From efficacy to effectiveness

On the horizontal axis experiments can be placed along a spectrum of how precise they measure causality. Policy experiments that value efficacy are those that, despite being conducted in a field setting, have a design similar to that of a laboratory experiment. They are randomised, with both control and treatment groups. With this design a policy experiment tests an isolated variable (i.e. an expected change in social policy) to determine whether the policy will have the expected impact. This is the approach Donald Campbell conceptualised for his “Experimenting Society” and it is still used today to measure the economic impacts of social policy (Greenberg et al, 2003; Vedung, 1997; Fischer, 1995). Confidence in the causal effect of a policy intervention is the most common design criticism of policy experimentation, for despite the use of randomisation and control groups being an example of scientific rigor, scholars point out that causal inference will never be strong enough to predict with any certainty what the true effects of a policy intervention will be (Vedung 1997, Fischer 1995, Higgins 1981).

However, causal inference is difficult to establish in even the most controlled of laboratory experiments (Shadish et al, 2002) so in a policy experiment conducted in the field, absolute confidence that the intervention caused the observations is impossible in complex systems like society or ecosystems (Parsons, 1996). Arguably it remains logical that obtaining any evidence on the effects of a proposed policy is better than relying on hunches (Stoker 2010, Sanderson 2002) so less rigorous “quasi” experiments that test effectiveness are of value.

These experiments sit towards the other end of the horizontal spectrum. A quasi-experiment (or “non-random field trial”) aims to evaluate a proposed policy’s effectiveness, which is a less rigorous test than measuring its efficacy (Petticrew et al, 2005). Instead of using randomisation to improve causality, a baseline can be identified at the beginning of an experiment that allows for knowledge claims to be made at the end (for instance like that described by Martin and Sanderson where evaluations are based on a before/after comparison of the experiments’ progress (Martin and Sanderson, 1999)). In this way reliable results can still be rendered via controlling mechanisms and strict monitoring and feedback cycles (Lee, 1999). At the far right end sit experiments that do not focus on establishing causality (i.e. those in transitions management) although monitoring and evaluation for learning purposes is required against interim objectives (Voss and Bornemann 2011). Evidence from these experiments can be considered indicative rather than conclusive (Pettigrew et al 2005).

Obviously, evidence as to whether the policy works is less reliable from non-random field experiments than from randomised experiments but as Campbell notes: although a lack of randomisation undermines the validity of an experiment, for practical, political, and/or ethical reasons it may be required (Campbell, 1998).

2.1.2 The nature of the intervention: either testing policy or developing policy

This delineation essentially refers to the flexibility that allows experimental designers to amend the experiment as it progresses. When policy experimentation is conducted as impact assessment or evaluation, the policy is essentially being tested for its effects. It is complete and there is no room to amend or change the experiment as it progresses- amending it can even be seen as compromising it; i.e. from “whether it works” to “how to make it work” (Sanderson 2002). However, the reiterative monitoring and feeding of new knowledge back into the experiment by amending the hypotheses

is a fundamental part of adaptive management experiments (Lee, 1999; Armitage et al, 2008; Voss and Bornemann, 2011). The experimental design is tweaked as it goes (i.e. “learning by doing”) and the aim is to craft a successful policy or technology that solves the problem, can meet political or other predetermined objectives (Cook et al, 2004) and can be up-scaled. These experiments are arguably better prepared for change because the flexibility makes them more adaptable.

Moving south along the vertical spectrum, some policy experiments are designed to develop a new policy, technology or governance strategy innovation *in situ* to learn about the effects as they emerge (e.g. Raven et al 2008). Focus is less on learning the impact of policy change and more on providing a space where institutional structures are relaxed and ideas and learning from the experiment can develop (Berkhout et al, 2010). These types of arrangements are also referred to as “pilot projects”, which can be conducted in the spirit of experimentation (Vrugdenhil, 2010). Anyone can initiate this sort of experiment but the link to policy may then be more tenuous; as demonstrated by Bai et al (2010) who found that if an experiment is used to develop a technological innovation with a range of stakeholders, its success at influencing policy may be contingent on government participants being involved. Policy experiments can be conducted to test governance arrangements as well (Voss and Bornemann 2011). In these “deliberative” experiments (Metze, 2010) stakeholders take part in sharing problems and have a role in deciding actions to be taken by all participants, backed up by soft law approaches (Hajer 2003:187). Notably, in this approach, the state is a participant but should not (theoretically at least) dominate (Hajer 2003).

An interesting pattern between policy experiments that are open to development and those that are not is the type of participants involved. Policies that are tested for their causality involve experts, generally scientists and policy makers (i.e. Greenberg et al 2003) and there is very little public input into the design and administration. Moreover, there is an absence of discussion in the literature on the role of the public, beyond the observations that a social policy experiment is initiated by the government and conducted on citizens as the subjects of experimentation- not as experimenters themselves (Greenberg 2003, Fuller, 1998). The participants involved in an experiment may be public officials at several layers of government (Stoker 2010) but the degree that non-state actors are involved in the social policy experiment extends only to public “engagement” (Martin and Sanderson 1999). Instead, the public merely reacts to the research or adapts to it (Fuller, 1998).

In contrast, other experiments are inclusive and take more of a collaborative approach. In order to plan for the unknown, adaptive management attempts to bridge the gap between “disciplines, data, techniques, knowledge, institutions, and people” (Holling, 1978) and emphasises the integration of various knowledge sources (Voss and Bornemann (2011). The more scientific adaptive management experiments limit participation to policy makers and scientists (e.g. Holling 1978) but the field has evolved to the incorporation of other societal actors as well (e.g. in the development of the adaptive co-management approach). In experiments that develop novel ideas and solutions the focus is on involving people that contribute something; whether it is knowledge, capabilities, or resources (Berkhout et al 2010).

Smith et al (2004) point out that the concept of governance incorporates both state and non-state actors in the formulation of public policy, and the fields of adaptive management and transitions management see non-state actors as agents that contribute to the development of policy with their knowledge and resources, or should be involved for the sake of legitimacy. On the other hand, social policy experiments are still considered a government activity, run by bureaucrats with the citizen as a

subject, and may be viewed as outdated, a “Type 1” government solution. However, each experimental type has learning as its main focus and different designs may render different learning outcomes. Identifying and analysing these learning outcomes is the main drive of this research project.

In sum, the concept of a policy experiment can be divided into two “ideal types”: those with a focus on assessing the impact of a policy and those that focus on developing technical and institutional policy innovations. Figure 2.2 below illustrates the differences:

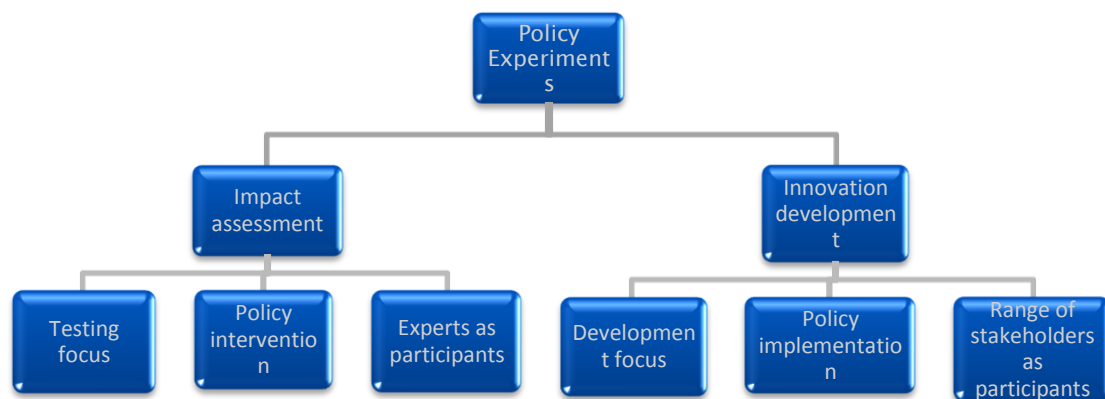


Figure 2.2 Breakdown of experiment types.

However, despite these differences, a definition of a “policy experiment” can have the following characteristics:

1. It attempts to test or develop an innovation, whether an innovation in technology, policy, or even governance. Testing ranges from explicit findings of cause and effect to establishing a baseline for monitoring effects;
2. It provides a “protected space” within which to do so by temporarily changing the institutional context. Ethically this may not be fair on citizens who lose out if there is a change to the status quo, but it may be vital if experiments are to make an impact;
3. It requires the involvement of participants and varies according to who participates, to what degree, and for what purpose;
4. What differentiates it from other types of experiments is that it has a connection with government policy (whether directly as a policy intervention or indirectly as a tool to develop innovations that impact policy) and seeks to influence it; and finally
5. It has intentions to upscale, so it relies on the perceived success of the experiment (although a strict policy experiment will not aim to be adopted at full scale and be willing to accept failure of the innovation as a learning experience).

2.2 What is not a policy experiment?

In its most relaxed application, an experimental approach to policy making sees all policies as ongoing experiments if they are monitored and evaluated regularly with feedback loops for learning (Huitema et al 2009). An example is from Pahl-Wostl (2006) where she suggests that facilitated stakeholder processes are an extension of an experiment from the adaptive management field. In the same vein, some scholars equate policy changes with experimentation; for instance, Ostrom argues that intended changes to institutional arrangements are effectively policy experiments (Ostrom 2005). Sanderson introduces the idea of “intelligent policy making” in the spirit of pragmatist John Dewey, who wrote about policies as working hypotheses subject to experimentation (Sanderson, 2009). Due to the immense respect I have for these scholars it is tempting to include these approaches in the definition of an experiment. However, monitored changes to existing policy are arguably still within the *status quo* and I cannot consider them an explicit, intentionally designed “policy experiment”. Expanding the definition to fit any policy would also mean an experiment could go on indefinitely and this counters the requirement for an experiment to emerge in an artificial institutional space.

A second delineation I want to make is between policy experiments and pilot projects. Quite often in the literature the terms policy experiment and pilot project/programme are used interchangeably. Sanderson (2002) describes pilot projects as having an experimental/ quasi-experimental design, but suggests they are more susceptible to negative influence. For instance, he points out that some pilots encourage ongoing improvements and sharing of best practices among test pilots, which create difficulties for impact evaluation. A focus on best practice moves the spotlight from whether the policy creates the impact expected (whether it works) to how best to make it work; in other words, the purpose changes from being about testing to being about “exemplifying”. Vreugdenhil et al (2010) conclude much the same with their distinction of a pilot project’s three roles - research, management, and political-entrepreneurial. The third role is used for advocacy purposes, to “convince other actors of one’s own point of view or to lobby for specific solutions to envisaged future problems” (Vreugdenhil et al 2010).

In conclusion, a policy experiment is effectively a temporary institutional space that is created to test or develop a policy, governance arrangement, or technological innovation that has a connection to policy. Importantly, a policy experiment is meant to generate effects. It creates social effects because it interrupts the *status quo* both legally and institutionally and its inclusion of the public either as stakeholder or subject may be ethically questionable. It generates bureaucratic effects because they are mostly initiated, attended, and funded by the state (Greenberg et al 2003) although shifts in governance are changing this. It has political effects because it wants to influence policy and the evidence produced has to compete with political interests. The competition with political interests features high in the next section.

2.3 Concerns with the use of policy experiments

On the face of it, policy experimentation is a reliable way to generate evidence for policy making. However, there are many criticisms about methodological issues and scepticism about the relationship between policy experiments and politics.

2.3.1 Methodological design

The problems with establishing a policy experiment’s causal validity were referred to in an earlier section, but there are also criticisms made of a policy experiment’s “generalisability” (where the results of one experiment can, with a degree of certainty, be replicated over and over again). Justifying the up-scaling or “roll out” of an experiment is not feasible from its results, they argue, because no natural or social system is ever the same in place, population, or time period (e.g. Greenberg et al 2003). Furthermore, sites are rarely randomly chosen and changes in social attitudes, governments, and business cycles render the experiment inescapably contextual. The limited temporal and spatial scale to which a policy experiment is applied also affects up-scaling, as it may not pick up on effects of the intervention at a community level (Greenberg et al 2003, Vedung 1997). The same argument is made for natural resource management policies that manage the ecosystem at different spatial scales (Walters and Holling 1990) where small, localised experiments can easily miss the effects of large scale biophysical processes.

These criticisms are valid; however, an accumulation of evidence can help to override the problem of whether results can hold for other populations, settings, or measurement variables, and combining the results of experiments on the same topic may maintain generalisability (Stoker 2010). Groups like the Campbell Collaboration (<http://www.campbellcollaboration.org>) assist by performing systematic reviews and meta-analyses (where possible) on policy experiments and disseminating the results. In other cases, the fact that an experiment delivers different results in one context compared to another is taken as a given and encouraged (Raven et al 2008).

A slightly different criticism of experimental design is made by McLain and Lee (1996) who reviewed the use of experiments in three case studies in Northern America. They concluded that the experiments there relied excessively on the use of rational planning models and failed to value non-scientific forms of knowledge. They argued that some experiments pay “inadequate attention to policy processes that promote the development of shared understandings among diverse stakeholders”. This criticism provides support for ensuring that experimental design facilitates not only instrumental (or cognitive) learning by experts, but also more reflexive learning effects, such as normative and relational learning between participants.

2.3.2 Use of evidence in policy making

Some criticisms of experiments do not concern the method or design *per se*, but are of a more political nature. There are numerous themes: political misuse (Sanderson, 2002), time frustrations (Vedung, 1997), influence on resources (Sanderson, 2002), ethical concerns (Fischer, 1995), and the different intentions of policy makers and scientists (Lee, 1999). These political issues occur both within the policy experiment itself and in the space where the experiment interacts with the wider policy domain.

The results of policy experimentation may be subject to disingenuous use by politics. As noted earlier, Sanderson (2002) provides a critical report on the UK’s use of pilot programmes, highlighting how political interests bend experimental results according to their needs. Within the policy domain, experimenting just becomes a show trial of a policy and not actually any kind of test to be learned from. Policy makers already “know” their principles work and so they do not require experimentation. The process becomes one of demonstration (Sanderson 2002). An experiment can even be used as a tactic to delay proper action (Halbert 1993), with policy makers requiring “more research” before making a final decision. Supporting this is Greenberg et al (2003),

who found that the experiments in their case studies were conducted to replace real action with symbolic action, although they point out this may backfire politically if the experiment shows the policy works. Within the experiment itself, experimental design that ignores politics can allow powerful political actors to exploit the experiment’s structure to the detriment of weaker participants (Voss and Bornemann, 2011).

2.3.3 Time

Time is a factor that can inhibit the use of policy experiments to produce evidence for policy making. Political considerations constrain the length of time a policy experiment can operate (Sanderson, 2002) and the need for urgent action in combination with politics makes planned experimentation stiff and time consuming (Vedung, 1997). For instance, if an issue has media attention or is high up on a party’s policy agenda there is a sense of urgency to show something is being done and the time it takes to conduct an experiment may be too long compared with the time a policy maker has for finding a swift solution to a policy issue. Conversely, even if the experiment is conducted over a short time, a policy issue itself may be so broad that its deep seated effects take a long time to show themselves and the experiment is not indicative of the true nature of the situation. Policies that aim to change behaviours or carry out institutional reform may take a while to show any measurable effects, thus the experiment may be lengthened and risk political changes (change in government, issue focus) that could threaten the experiment’s relevance.

It should be noted that time can also be useful to an experiment, for if timed correctly it can be effective in keeping a policy idea alive or by coinciding with a window of opportunity and creating momentum for a policy to emerge (Greenberg et al, 2003; Huitema and Meijerink, 2009).

2.3.4 Ethics

Experiments are often criticised because of the ethical dynamics they generate (e.g. in Fischer, 1995). Conducting out policy experiments on the general population can draw serious criticism mainly due to the ethics of treating people differently for experimental purposes, i.e. giving only some eligible people a benefit instead of everyone, or having people think they are receiving a treatment when they are not (Greenberg et al, 2003). In order to avoid the development of adverse reactions in subjects when they learn they are receiving different treatment to others, Campbell suggests that subjects should not be aware they are participating in an experiment at all (this would also diminish the subjects’ reflexivity on the experiment, otherwise known as the Hawthorne Effect). However, he concludes this sort of experimental design undermines moral and democratic values and the threat to validity must be absorbed into the experiment (Campbell 1998).

2.3.5 Novelty

The novelty of policy experimentation can cause people to act differently than normal, both within the experiment and within the wider policy domain. An experiment is essentially a temporary change in legal and institutional context and test subjects may react to the fact they are part of an experiment, with more dedication and a higher level of commitment than when the policy is launched for real (Sanderson, 2002). Within the policy domain, if there is a high level of political interest in making the experiment a success an experiment may draw more resources than those made available for the actual policy launch (Sanderson, 2002; Bille, 2010).

2.3.6 Science policy interactions

There is tension between scientists and policy makers at the science-policy interface when it comes to applying experimental results to policy making in the wider domain. It is said that scientists and policy makers are motivated by different goals (Stoker, 2010) e.g. experimentation has a quest for truth but for politics truth is opaque and reliance is heavier on values (Lee 1999). Scientists that conduct experiments to inform policy making can see their results unused if they deliver evidence of things a policy maker does not want to hear (Sanderson 2002). Policy makers are also aware that once a policy is created it is difficult to terminate (Peters, 1998). Halbert (1993) departs from a Popperian perspective and points out that deliberately setting out to learn can threaten the incumbent social structure because learning can only occur via the identification of error. She argues that failure is politically hazardous and learning from failure is neither politically nor scientifically acceptable, which is a problem if failure needs to be an acceptable part of the learning process.

In relation to the tensions caused between scientists and policy makers within the experiment, the precise goals and objectives of an experiment can be at loggerheads with a policymaker’s need for diffuse aims and objectives for political purposes. Science requires that an experiment have goals that are identified precisely; but in politics, where success relies on negotiation between divergent interests, it is best if goals are left implicit or vague (Higgins 1980).

2.3.7 Summary

With such a broad sweep of criticisms, it is a wonder that there is any support for policy experimentation at all. However, despite the misgivings, the call for producing policy evidence via experimentation is growing (e.g. article “Is Our Adults Learning?” from the New York Times newspaper, April 2012) and research is needed about how we can design experiments so they make best use of the learning processes they facilitate. Enhancing learning in experiments may even address some of the concerns referred to above- for instance, this project focuses on how learning can enhance the production and use of evidence in the policy decision making process, but learning may also reduce problems with time and the science-policy interface.

Firstly, however, the concept of learning needs to be defined for the purposes of this project, which is the focus of the next section.

3 Learning

Learning is arguably the most sought-after verb in the social sciences. The management of our social-ecological systems is complex and uncertain and learning is a concept that helps us address those challenges. It allows us to realise the challenge of “learning our way out of complex environmental problems”. The governance literature is very clear on how vital learning is for managing social-ecological systems. It has been suggested that modern environmental problems have a “wicked” nature about them, making them highly complex and essentially unsolvable by scientific means (Rittel and Webber 1973). Many studies refer to this issue and suggest different approaches. Voss and Bornemann (2011) for instance argue that the best way to navigate through uncertainty and complexity is by gathering knowledge “to explore the system’s true structure” and learning to adapt.

This project takes the view that there are two spaces created by policy experiments that foster learning. One is within the experiment where social learning takes place between participants. This learning between the participants is inhibited or enhanced by the experiment’s institutional design. The second space that learning takes place is beyond the experiment in the policy domain. This chapter aims to identify what learning means and how it originates, and how it is measured and evaluated in regards to these two spaces.

3.1 What is learning?

Learning is arguably one of the most sought-after verbs in the social sciences. It can be defined as being essentially about change – the process by which knowledge, skills, and attitudes are acquired (Muro and Jeffrey, 2012). The management of our social-ecological systems is complex and uncertain and learning is a concept that helps us address those challenges.

The governance literature is very clear on how vital learning is for managing social-ecological systems. It has been suggested that modern environmental problems have a “wicked” nature, making them highly complex and essentially unsolvable by scientific means (Rittel and Webber 1973). Many studies refer to this issue and suggest different approaches.

Voss and Bornemann (2011) for instance argue that the best way to navigate through uncertainty and complexity is by gathering knowledge “to explore the system’s true structure” and learning to adapt.

In this project I take the view that there are two dimensions created by policy experiments that foster learning. One is within the experiment where mutual learning takes place between participants. It is hypothesised that this learning between the participants is influenced by the experiment’s institutional design. The wider dimension is within the policy domain, where the evidence of an experiment aims to influence the policy decision making process, and its ability to do so is measured as policy learning. This chapter aims to identify how learning is measured and evaluated in regards to these two spaces.

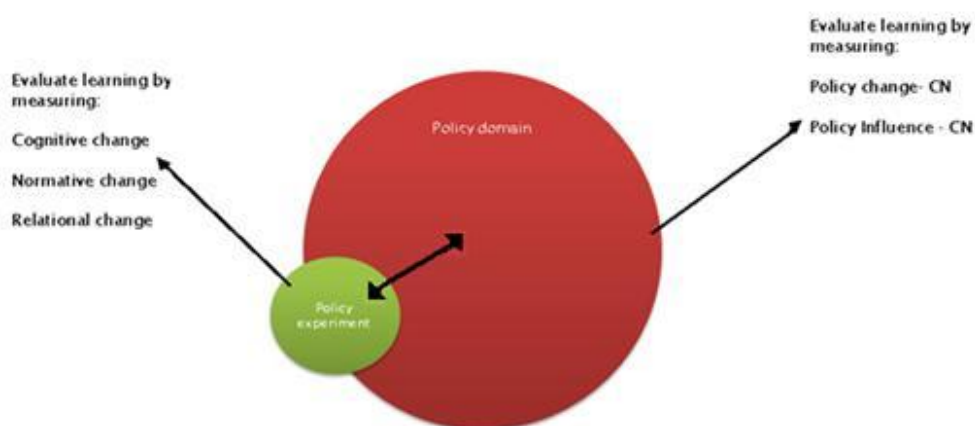


Figure 3.1 Learning Arenas.

3.2 Adopted typology of learning

The literature distinguishes between two fundamental sorts of learning: *cognitive (or instrumental)* learning, which is described as the acquisition of knowledge, content management, and problem-solving task-oriented actions (Pahl-Wostl 2006); and *relational (or communicative)* learning, which is concerned with social processes, how individuals relate to and build trust with one another, and whether they understand and appreciate one another’s frames (see Muro and Jeffrey, 2012; Haug et al 2010, Huitema et al 2010, Pahl-Wostl 2006, Armitage et al 2008).

A review of the literature reveals that multiple loop learning is the most popular evaluative concept for social and organisational learning (for example, see Maarleveld and Dangbegnon, 1999; Armitage et al 2008, Thomas and Allen (2006)). It originates from Argyris and Schön (1978), who conducted research into learning in organisations. They drew distinctions between the learning that occurs when errors are fixed and the learning which emerges when the underlying values and goals of the issue are reflected upon (Argyris and Schön, 1978). Armitage et al (2008) have defined multi loop learning in adaptive co-management terms, with single loop learning as the sharing of stakeholder interests and the use of them to build alternative strategies. Double loop learning involves a change of management strategies based on reflection on values underlying incumbent routines.

However, an alternative to a multiple loop learning analysis is the measurement of three separate types of learning. The attraction to this method of conceptualising learning is that the types are not layered upon each other in order of importance. Instead, *cognitive, normative, and relational* learning types are distinguished because, depending on the context, learning new knowledge and ideas can be just as important as reflecting on goals and values. In effect, the two are not mutually inclusive (Haug et al 2010).

This learning typology originates from Haug et al (2010) where in order to measure the learning effects of policy games in climate policy appraisal, three learning types: cognitive, normative, and relational learning, were conceptualised. The policy learning literature focuses on the knowledge-based dimensions of learning, and the public participation/impact assessment literature analyses social learning and group

dynamics between subjects, so the typology succeeds in knitting the learning theories together. The identification and isolation of these three learning effects provide a useful typology that can be used to recognise and analyse learning from experiments, as discussed below.

3.2.1 Cognitive learning

A cognitive learning process results in “the acquisition of new, or the improved structuring of existing, knowledge” (Haug et al 2010). It refers to the learning of factual knowledge that is generated by carrying out the experiment and also the knowledge participants share with one another. Therefore, not only it is important that an experiment generates high levels of information that the participants learn, the design of the experiment must incorporate mechanisms to ensure high levels of cognitive learning between the participants.

When experiments are designed for learning in particular, for instance “learning by doing” experiments in adaptive management (Lee, 1999), cognitive learning is what they have in mind. An experiment aims to test something and generate knowledge; however, just because an experiment generates knowledge does not necessarily mean it produces high levels of cognitive learning. For instance, cognitive learning relies on a clear presentation of results and if the person communicating the results is not comprehensible then low levels of cognitive learning may occur (particularly on technical issues that not all participants are familiar with).

Evidence that an experiment generated cognitive learning in the wider policy environment may be measured in evidence of new information or ideas that emanated from experiments being applied to policy decisions (following Busenberg 2001).

Cognitive learning is a process that results in the cognitive change of a participant. What is learned is new information that can correct errors in what we already know and enhance our understanding of issues (Mostert et al 2007). Cognitive learning does not result in a change in underlying norms, values, or belief systems (Huitema et al, 2010). If learning new information leads someone to question their view on a topic, this learning is assessed as “normative learning.

3.2.2 Normative learning

Normative learning links implicitly to the other two learning effects. Haug et al (2010) define it as “changes in the viewpoints, norms and values of participants”. A visit to one’s views, values, and norms could be in light of new knowledge- a cognitive influence; or due to deliberation or persuasion from another participant- a relational influence (Haug et al 2010). Normative learning can be measured as a convergence of views of participants (Huitema et al, 2010; Muro and Jeffrey, 2012), which may be achieved through “ensuring strong deliberation processes and clear open communication channels” (Voss and Bornemann 2011).

Fischer (1995) states that a policy experiment is useful for measuring whether an experiment reaches its policy goals, but not whether the goals in themselves are appropriate (i.e. an experiment cannot provide a venue for normative learning). This is relevant if the experiment is solely testing causality (see also Weiss, 1997), but for other experiment types it is possible that the design of the experiment may facilitate goal reflection and change, particularly if goal reflection is key aspect of the experiment (Armitage et al 2008) and the experiment is seen as an evolving practice (Bai et al 2010).

3.2.3 Relational learning

Relational learning refers to the collaborative effects of learning; like an improved ability to cooperate, increase in trust, and a better understanding of the mindsets and frames of other participants (Haug et al, 2010; Huitema et al, 2009). It follows from Webler’s class of learning as “moral development” which includes being able to take on other participants’ perspectives and learning how to cooperate with others (Webler et al 1995).

An experiment that fosters the relational learning process may produce evidence of participants sharing and understanding each others’ frames of particular issues; for instance how each participant views the policy issue, the reasons for the experiment, how they define the experiment in terms of success, etc. Recognition of different frames can improve trust relations and create room for better cooperation (Mostert et al 2007). If “reframing” occurs, whereby a participant changes their view on an issue after engaging with other participants, then this is arguably evidence of a normative learning process, but any evidence of increased trust and cooperation and discussion on frames is due to relational learning. Trust is in itself a complex and multi-faceted phenomenon that can be assessed as commitment, respect, or honesty (Muro and Jeffrey, 2012).

The table below outlines the framework for analysis of the three learning types.

Table 3.1 The framework for analysis of the three learning types.

Learning effects	Criteria	Description
Cognitive- change in knowledge	Increase in knowledge about interventions;	Refers to increase in knowledge generated from intervention;
	Sharing of knowledge between participants.	Refers to exchange of knowledge between participants.
Normative- change in perspectives	Change in goals;	Refers to the shift in focus;
	Change of views and beliefs	Refers to convergence of views and perspective among participants.
Relational- change in attitude	Increase in trust;	Refers to level of commitment of others;
	Increase in co operations;	Refers to creation of working groups to get tasks done;
	Increased understanding of others’ mindset.	Refers to the understanding of others’ frames.

3.3 Evaluating learning within the policy domain

Busenberg (2001) defines policy learning as *the process in which individuals apply new information and ideas in policy decisions*, but scholars believe policy learning may only be observed when policy change has taken place (Bennett and Howlett (1992)). However, evaluating learning from an experiment by equating it with policy change may be too high a benchmark and we will miss the other impacts experiments have on policy. For instance, Greenberg et al (2003) found in their case studies that although no experiment was pivotal to a decision to adopt or not adopt a policy, results were used in a variety of ways.

4 Relationship between Experiments and Learning

As noted above, there are two arenas where policy experiments can affect learning outcomes: within the experiment itself and within the wider policy domain where the experiment engages. The following section summarises the potential effects experimental design can have on learning within these two arenas and the lens through which I plan to analyse the relationship between learning and design.

4.1 Learning outcomes from within the experiment

Analysing a policy experiment’s function as a boundary object highlights its use as a platform where various stakeholders become participants, brought together to generate knowledge and share ideas. In policy experiments that focus on impact assessment, the participants’ role is largely to generate evidence of the policy’s impact. In more collaborative experiments, participants can be expected to share knowledge, align their goals and expectations, and build trust that enables better decision making and policy design. Whether learning outcomes improve if an experiment has a collaborative design, is a fundamental question of this research.

The importance of learning is discussed in the participation, organizational learning, and natural resource management literatures. The essential message is that by engaging with and learning from each other, participants with different perspectives may adapt, which results in shared or complimentary perspectives that enable the development of a common understanding and shared action (Muro and Jeffrey 2008). Others are more circumspect about learning and the merging of perspectives, and prefer to link learning with open venues for negotiation and structured conflict (e.g. van den Hove, 2004; Cuppen, 2009).

Participants cannot be forced to learn, but they can be encouraged to learn through the creation of learning situations within social spaces (Rist et al, 2006). This project hopes to identify design choices that experimental designers can make so their experiment is a “learning situation”. This is beneficial because the function of a policy experiment as a boundary object that encourages learning may make the experiment more relevant to the policy process and the scientific outcomes “more responsive and transparent” (Huitema et al 2009). Moreover, a collaborative design could alleviate some concerns discussed above: it may increase the level of commitment amongst participants after the experiment ends; a flexible design may address the time issue by allowing the experiment to keep up with shifting political priorities; and/or it may address power issues that surface between participants.

Therefore, one of the fundamental questions of this research is: do experiments that are designed in a more collaborative manner, with co-management “principles” (power sharing, knowledge use, participatory (Huitema et al, 2009)) result in higher learning outcomes than other experimental designs?

4.2 Learning outcomes within the policy domain

All policy experiments- whether functioning as policy evaluation, impact assessment, or policy development- provide knowledge and evidence of their effects, which in turn are used to inform policy decision-making (Sanderson, 2009). Therefore, when the literature focuses on experiments and learning, it does so in the context of how an experiment can *influence* the policy process (cf. Sanderson 2002).

In regards to design, since the learning outcomes in this arena emanate from the experiment’s intervention itself, an interesting perspective to study policy learning from experiments is at the *science-policy interface* (SPI). A policy experiment is essentially a platform where bureaucrats and researchers meet and its design could have an impact on how knowledge created by the policy intervention moves successfully or unsuccessfully into the policy making realm.

This is an interesting perspective because an experiment’s design features may influence its successful functioning as a SPI platform, in turn influencing whether an experiment increases policy learning. It does this by improving the movement of science from the experiment up to the policy domain through the relationship of scientists and policy makers. In order to test this assumption, I propose to use the typology of factors from Cash et al (2003) that assess the quality of information generated by a SPI. *Salience* refers to the relevance of the experiment to decision makers; for instance, whether the experiment has clear objectives, methods for review, a clear connection between science and policy. *Credibility* refers to how sufficient the scientific evidence and arguments are; for instance, whether participants review the outputs, whether there is continuous evaluation, or if there is a clarity of roles. Finally, *legitimacy* refers to how respectful the experimental process has been to the participants’ divergent beliefs and values and fair in its treatment of the differences (Cash et al, 2003); for example, are all relevant parties involved, is knowledge shared, are a variety of perspectives considered.

Therefore, a second fundamental research question is: will policy learning be more common in policy experiments that are effective SPIs?

5 Conclusion

This state of the art paper describes the two concepts *policy experiment* and *learning* and reflects on how the two relate. Policy experiments vary in many ways; for instance in the participants involved, the information they generate, and the innovations they test. Each field tends to discuss experimental design within its own discipline, which is understandable because experimental design is contingent on its use. I was surprised at the scope of what is defined as a policy experiment, but also heartened to see there were strong similarities. There is space in the science for an overview of policy experiments because disciplines do not often overlap; for instance, natural resource management and transitions management experiments do not tend to make or consider the existent political and design criticisms (a criticism made by Voss and Bornemann (2011)). I hope to fill this gap by analysing experiments in their institutional form, so findings can be applied across the board regardless of discipline.

The literature is supportive of policy experiments but also very critical of their use, particularly in their design and their relationship with politics. Learning is a heavily theorised concept in the literature but essentially two types of learning are identified: the more cognitive elements and the relational elements. This paper assessed learning from experiments in regard to three types- cognitive, relational, and normative, and highlighted how these learning effects might emerge within a policy experiment. It is clear that policy experiments and learning are concepts that are intimately related and research into how experimental design can influence and improve learning would add greatly to our knowledge about this relationship.

References

- Adger, W.N., N. W. Arnell, E.L. Tompkins. 2005. Successful adaptation to climate change across scales. *Global Environmental Change* 15(2): 77-86.
- Argyris, C. 2003. A Life Full of Learning. *Organization Studies* 24:1178.
- Armitage, D., M. Marschke, and R. Plummer. 2008. Adaptive co-management and the paradox of learning. *Global Environmental Change* 18:86-98
- Bai, X., B. Roberts, and J. Chen. 2010. Urban Sustainability Experiments in Asia: Triggers, Actors, Linkages, Barriers and Pathways. *Environmental Science and Policy* 13(4): 312-325.
- Berkhout, F., G. Verbong, A.J. Wiecek, R. Raven, L. Lebel, and X. Bai. 2010 Sustainability Experiments in Asia: Innovations Shaping Alternative Development Pathways? *Environmental Science & Policy* 13(4): 261-271.
- Bille, R. 2010. Action without change? On the use and usefulness of pilot experiments in environmental management. *SAPIENS, Veolia Environment Institut* 3(1).
- Brookes, D. 2012. Is Our Adults Learning? *New York Times Newspaper*. 24 April 2012.
- Busenberg, G.J. 2001. Learning in Organizations and Public Policy. *Journal of Public Policy*. 21:173-189.
- Campbell, D.T. 1998. The Experimenting Society, p35. In Dunn, W. (Ed.), *The Experimenting Society: Essays in Honor of Donald T. Campbell*. Policy Studies Review Annual volume 11. Transaction Publishers, New Brunswick, New Jersey.
- Cook, W., D. Casagrande, D. Hope, P. Groffman, S. Collins. 2004. Learning to roll with the punches: adaptive experimentation in human dominated systems. *Frontiers in Ecology and Environment* 2(9), 467-474.
- Farrelly, M., R. Brown. 2011. Rethinking urban water management: Experimentation as a way forward? *Global Environmental Change* 21:721-731
- Finger, M. and P. Verlaan. 1995. Learning our way out: A conceptual framework for social-environmental learning. *World Development*. 23(3):503-513.
- Fischer, F. 1995. *Evaluating Public Policy*. Nelson Hall, Chicago, Illinois, USA.
- Folke, C., T. Hahn, P. Olsson, and J. Norberg. 2005. Adaptive governance of social-ecological systems. *Annual Review of Environment and Resources* 30(1):441-473.
- Freire, P. 1993. *Pedagogy of the Oppressed*. New rev. 20th Anniversary ed. New York.
- Fuller, S. 1998. Making Science into an Experimenting Society, p69. In Dunn, W. (Ed.), *The Experimenting Society: Essays in Honor of Donald T. Campbell*. Policy Studies Review Annual volume 11. Transaction Publishers, New Brunswick, New Jersey.
- Greenberg, D., D. Linksz, and M. Mandell. 2003. *Social experimentation and public policy making*. The Urban Institute Press, Washington, D.C., USA.
- Hajer, M. 2003. Policy without polity. Policy analysis and the institutional void. *Policy Sciences* 36:175-195.
- Halbert, C. 1993. How adaptive is adaptive management? Implementing adaptive management in Washington State and British Columbia. *Reviews in Fisheries Science* 1(3)261-283.
- Hall, P. 1993. Policy paradigms, social learning and the state: the case of economic policymaking in Britain. *Comparative Politics* 25:275-296.

- Haug, C., D. Huitema, and I. Wenzler. 2011. Learning through games? Evaluating the learning effect of a policy exercise on European climate policy. *Technological Forecasting and Social Change* 78(6): 968-981.
- Higgins, J. 1980. The Unfulfilled Promise of Policy Research. *Social Policy and Administration* 14(3).
- Holling, C. S. 1978. *Adaptive environmental assessment and management*. John Wiley and Sons, London, UK.
- Huitema and Meijerink (2009) *Water policy entrepreneurs. A research companion to water transitions around the globe*, Edward Elgar Publishing UK
- Huitema, D., E. Mostert, W. Egas, S. Moellenkamp, C. Pahl-Wostl, and R. Yalcin. 2009. Adaptive water governance: assessing the institutional prescriptions of adaptive (co-) management from a governance perspective and defining a research agenda. *Ecology and Society* 14(1): 26. [online] URL: <http://www.ecologyandsociety.org/vol14/iss1/art26/>.
- Huitema, D., C. Cornelisse, and B. Ottow. 2010. Is the jury still out? Toward greater insight in policy learning in participatory decision processes—the case of Dutch citizens' juries on water management in the Rhine Basin. *Ecology and Society* 15(1):16.[online]URL: <http://www.ecologyandsociety.org/vol15/iss1/art16/>
- Jha-Thakur, U., P. Gazzola, D. Peel, T.B. Fischer, S. Kidd. 2009. Effectiveness of strategic environmental assessment - the significance of learning. *Impact Assessment and Project Appraisal*. 27(2): 133-144.
- Jordan, A., D. Huitema, H. Van Asselt, T. Rayner, and F. Berkhout. 2010. *Climate Change Policy in the European Union: Confronting the Dilemmas of Mitigation and Adaptation*. Cambridge University Press, UK.
- Kolb, D. A. 1984. *Experiential learning: experience as the source of learning and development*. Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Lee, K. N. 1999. Appraising adaptive management. *Conservation Ecology* 3:3-16.
- Maarleveld, M., and C. Dangbegnon. 1999. Managing natural resources: A social learning perspective. *Agriculture and Human Values* 16(3): 267-80.
- McLain, R. J., and R. G. Lee. 1996. Adaptive management: promises and pitfalls. *Environmental Management*. 20(4):337-448.
- Metze, T. 2010. *Innovation Limited*. Eburon Academic Publishers, Delft, NL.
- Mostert, E., C. Pahl-Wostl, Y. Rees, B. Searle, D. Tàbara, and J. Tippett. 2007. Social learning in European river-basin management: barriers and fostering mechanisms from 10 river basins. *Ecology and Society* 12(1): 19. [online] URL: <http://www.ecologyandsociety.org/vol12/iss1/art19/>.
- Muro, M., and P. Jeffrey. 2008. A critical review of the theory and application of social learning in participatory natural resource management. *Journal of Environmental Planning and Management* 51:325-344.
- Muro, M. 2008. *The role of social learning in participatory planning & management of water resources*. PhD Thesis: <http://dspace.lib.cranfield.ac.uk/handle/1826/3513>.
- Olsson, P., L. H. Gunderson, S. R. Carpenter, P. Ryan, L. Lebel, C. Folke, and C. S. Holling. 2006. Shooting the rapids: navigating transitions to adaptive governance of social-ecological systems. *Ecology and Society* 11(1): 18. [online] URL:<http://www.ecologyandsociety.org/vol11/iss1/art18/>
- Ostrom, E. 1999. Coping with the tragedy of the commons. *Annual Review Political Science* 2:493-535.

- Ostrom, E. 2005. *Understanding Institutional Diversity*. Princeton University, New Haven, Connecticut, USA.
- Pahl-Wostl, C. 2002. Towards sustainability in the water. The importance of human actors and processes of social learning. *Aquatic Sciences* 64:394-411.
- Pahl-Wostl, C. 2006. The importance of social learning in restoring the multifunctionality of rivers and floodplains. *Ecology and Society* 11(1): 10. [online] URL: <http://www.ecologyandsociety.org/vol11/iss1/art10/>.
- Parry, M.L., O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E Hanson. ed., *Climate Change 2007: Impacts, Adaptation and Vulnerability, Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, ISBN 978-0-521-88010-7.
- Parsons, W. 1996. *Public Policy: An Introduction to the Theory and Practice of Policy Analysis*. Edward Elgar Publishing Ltd.
- Peters, G. B. 1998. The Experimenting Society and Policy Design, p125. In Dunn, W. (Ed.), *The Experimenting Society: Essays in Honor of Donald T. Campbell*. Policy Studies Review Annual volume 11. Transaction Publishers, New Brunswick, New Jersey.
- Raven, R.P.J.M., E. Heiskanen, R. Lovio, M. Hodson and B. Brohmann. 2008. The Contribution of Local Experiments and Negotiation Processes to Field-Level Learning in Emerging (Niche) Technologies: Meta-Analysis of 27 New Energy Projects in Europe. *Bulletin of Science Technology Society* 28: 464.
- Reed, M. S., A. C. Evely, G. Cundill, I. Fazey, J. Glass, A. Laing, J. Newig, B. Parrish, C. Prell, C. Raymond, and L. C. Stringer. 2010. What is social learning? *Ecology and Society* 15(4): r1.
- Rittel, H., and M. Webber. 1973. Dilemmas in a general theory of planning. *Policy Sciences* 4:155-169.
- Sabatier, P. A. 1998. The advocacy coalition framework: revisions and relevance for Europe. *Journal of European Public Policy* 5:98-130.
- Sanderson, I. 2002. Evaluation, policy learning and evidence-based policy making. *Public Administration* 80(1): 1 - 22.
- Shadish, W.R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalised causal inference*. Houghton Mifflin Company, Boston, USA.
- Smith, A., A. Stirling, and F. Berkhout. 2005. The governance of sustainable socio-technical transitions. *Research Policy* 34 : 1491-1510.
- Smith, A. 2007. Translating Sustainabilities between Green Niches and Socio-Technical Regimes. *Technology Analysis & Strategic Management* 19(4) 427-450.
- Stoker, G. 2010. Exploring the promise of experimentation in political science: micro-foundational insights and policy relevance. *Political Studies* 58 300-319.
- Stoker, G. 2010. Translating experiments into policy. *The ANNALS of the American Academy of Political and Social Science* 628:47.
- Termeer, C., A. Dewulf, H. van Rijswick, A. Van Buuren, D. Huitema, S. Meijerink, T. Rayner, and M. Wiering. 2011. The regional governance of climate adaptation: A framework for developing legitimate, effective, and resilient governance arrangements. *Climate Law* 2: 159-179.
- Thomas, K. and S. Allen. 2006. The learning organisation: a meta-analysis of themes in literature. *The Learning Organization* 13(2)123-139.
- Vedung, E. 1997. *Public policy and program evaluation*. Transaction Publishers, New Brunswick, USA.

- Voß, J., and B. Bornemann. 2011. The politics of reflexive governance: challenges for designing adaptive management and transition management. *Ecology and Society* 16(2): 9. [online] URL: <http://www.ecologyandsociety.org/vol16/iss2/art9>.
- Vreugdenhil, H., J. Slinger, W. Thissen, and P. Ker Rault. 2010. Pilot projects in water management. *Ecology and Society* 15(3): 13. [online] URL: <http://www.ecologyandsociety.org/vol15/iss3/art13/>
- Walters, C.J., C.S. Holling. 1990. Large-Scale Management Experiments and Learning by Doing. *Ecology* 71(6): 2060-2068.
- Webler, T., H. Kastenholz, and O. Renn. 1995. Public participation in impact assessment: a social learning perspective. *Environmental Impact Assessment Review* 15:443-463.