

WORKING PAPER

Accepted for publication in: Policy Sciences

Experimentation in policy-design: Insights from the building sector

Dr. Jeroen van der Heijden

Regulatory Institutions Network (RegNet), Australian National University
Amsterdam Law School, University of Amsterdam

Contact:

j.j.vanderheijden@anu.edu.au

www.jeroenvanderheijden.net

Abstract

The current article questions how experimentation in policy-design plays out in practice. In particular, it is interested in understanding how the content and process of policy-design experiments affect their outcomes. The article does so by building on an original study into 31 real-world examples of experimentation in policy-design in the building sector in Australia, the Netherlands and the United States. All examples aim to improve the environmental sustainability of the building sector. The article finds that these 31 examples have attracted moderate to substantial numbers of participants (policy outcome HO.i), but have not achieved substantial numbers of buildings built or retrofitted with high levels of sustainability (policy outcome HO.ii). By carefully unpacking these policy-designs into a number of key characteristics it finds that this mismatch between the two outcomes may partly be explained by flawed policy-design processes. The article concludes with the main lessons learnt, and provides some suggestions on how to improve experimentation in policy-design.

Keywords

Policy-design, policymaking, policy evaluation, experimentation, policy learning

Experimentation in policy-design: Insights from the building sector

Abstract

The current article questions how experimentation in policy-design plays out in practice. In particular, it is interested in understanding how the content and process of policy-design experiments affect their outcomes. The article does so by building on an original study into 31 real-world examples of experimentation in policy-design in the building sector in Australia, the Netherlands and the United States. All examples aim to improve the environmental sustainability of the building sector. The article finds that these 31 examples have attracted moderate to substantial numbers of participants (policy outcome HO.i), but have not achieved substantial numbers of buildings built or retrofitted with high levels of sustainability (policy outcome HO.ii). By carefully unpacking these policy-designs into a number of key characteristics it finds that this mismatch between the two outcomes may partly be explained by flawed policy-design processes. The article concludes with the main lessons learnt, and provides some suggestions on how to improve experimentation in policy-design.

Keywords

Policy-design, policymaking, policy evaluation, experimentation, policy learning

1 Introduction

policies and proposal for social action [should] be treated as working hypotheses, not as programs to be rigidly adhered to and executed. They will be experimental in the sense that they will be entertained subject to constant and well-equipped observation of the consequences they entail when acted upon, and subject to ready and flexible revision in the light of observed consequences – John Dewey (1991 [1927]: 202-203)

an experimental approach to social reform, an approach in which we try out new programs designed to cure specific social problems, in which we learn whether or not these programs are effective, and in which we retain, imitate, modify, or discard them on the basis of apparent effectiveness on the multiple imperfect criteria available – Donald Campbell (1969: 409)

As the above quotes by Dewey and Campbell highlight, experimentation has long been advocated as an optimal approach to policy-design. Both authors would likely have been happy knowing that their visions appear to have become reality. There is now a large literature that provides a wide range of examples of this experimentation in policy-design (for overviews of the literature and a range of examples, see Greenberg & Schroder, 2003; Hoffmann, 2011; Van der Heijden, 2012). Yet, the current literature on experimentation in policy-design differs slightly from the visions of the past. Dewey and Campbell had in mind experimenting with the *content* of policy programmes (that is, testing, piloting, or demonstrating a particular policy-design, to use some terms from contemporary empirical policy-design literature; e.g., Markusson, Atsushi, & Stephens, 2011; Vreugdenhil, Taljaard, & Slinger, 2012). They do not appear interested in the *process* of policy-design – i.e., that an experimental policy-design has come about is a given to them.

Currently, however, scholars have become more and more interested in experimentation with both the content *and* process of policy-design. Like Dewey and Campbell they are interested in experimental policy-designs that differ from, for instance, traditional direct governmental regulation or subsidies (Jordan, Wurzel, & Zito, 2010). Yet, they are also interested in how different actor constellations are involved in the process of developing such policy-designs (Hoffmann, 2011). A typical example of this contemporary literature builds around what is termed ‘experimentalist governance’ (e.g., De Burca, 2010; Overdevest & Zeitlin, 2012; Sabel & Zeitlin, 2011), which is defined as ‘a recursive process of provisional goal setting and revision based on learning from comparison of alternative approaches to advancing these goals in different contexts’ (Overdevest & Zeitlin, 2012: 4).

Scholars, past and current, have expressed high hopes of what it is that such experimentation in policy-design may achieve. For instance, experimentalist governance is advocated as being more effective, more accountable, more democratic, and better suited to address complex societal problems than any of the traditional approaches to policy-design (De Búrca & Scott, 2006; Sabel & Zeitlin, 2012). However, the literature, past and current, remains rather silent as to the actual outcomes of such experimentation (cf., Greenberg & Schroder, 2003; Jordan, Wurzel, & Zito, 2013). There is an oft repeated reason for this silence: true experimentation in policy-design (with content, process, or both) is often impossible. In an ideal laboratory-like setting the experimental policy-design should then be randomly assigned to a test context and not to a control context, should affect people or processes in a representative setting, and should be carried out long enough to achieve meaningful results (Greenberg & Schroder, 2003). Practical constraints, ethical constraints, or both often stand in the way of doing so, which severely affects the generalizability of findings from policy-design experimentation (cf., Campbell, 1969). Yet, given the wide range of policy-design experiments, either documented or currently undertaken, it should at least be possible to now arrive at *moderatum* generalizations: ‘claims to basic patterns, or tendencies, so that other studies are likely to find something similar but not identical’ (Payne & Williams, 2005, 306).

This, then, is the aim of this article: it seeks to add knowledge on contemporary experimentation in policy-design (experimentation with content *and* process), and develop a number of moderate insights into such experimentation that might be confirmed or refuted through further evidence from other studies. In particular, this paper seeks to understand what design characteristics add to achieving successful policy outcomes from experimentation in policy-design. It does so through a study of a stratified sample of 31 recent real-world examples of experimentation in policy-design in the building sector in Australia, the Netherlands and the United States (US). In what follows, in section 2 first the research design is discussed. Then, in section 3, the research findings are presented. Finally, section 4 concludes by discussing the main lessons learnt.

2 Policy experimentation in the Australian, Dutch and US building sectors

The construction, maintenance and use of buildings (together: the ‘building sector’) is widely considered key in addressing climate change (Evans, Joas, Sundback, & Thobald, 2005; Yudelson & Meyer, 2013). The sector accounts for roughly 35% of global carbon emissions (IPCC, 2007). The good news is that the technology and knowledge are available to significantly reduce these emissions by 50% over the next 20 years (Newman, Beatley, & Boyer, 2009). Furthermore, it is expected that this can be done in a cost-effective manner, making the building sector the only sector in the world where high levels of environmental performance (e.g., significant reductions of carbon

emissions) will not require long-term financial sacrifices (IPCC, 2007). Unfortunately, in recent decades it has not proven easy to realize the unique potential the building sector holds. The sector faces a number of regulatory barriers that hamper traditional direct intervention by governments, and market barriers that hamper voluntary approaches to increasing the sector's environmental performance (for extensive discussions of these barriers, see Abaire, 2008; Bastianoni, Galli, Niccolucci, & Pulselli, 2006; Brown, 2001; Van der Heijden, 2013; Zimmerman & Martin, 2001).

Seeking to overcome such barriers in the sector, governments, businesses and civil society groups around the globe are now collaborating in the development and implementation of policy-design experiments that seek to realize the potential the building sector holds (for a range of examples, see Hickson, 2009). A typical example is the *Greener Building Collaboration* in Australia (fictional name), which brings together the city council of one of Australia's major cities, a national bank, a major fund manager, and property owners in the city's central business district. In particular, it addresses a specific barrier in the industry: property owners can often not find mortgages for retrofitting and upgrading their buildings. Banks are risk-averse in supplying mortgages as they consider that the cost of the upgrade is not (yet) represented in an increase in a building's market value (cf. Pivo, 2010). Yet, it has been evidenced that the cost of the upgrade will be paid back by lower amenity costs, and expected higher rent rates (Eichholtz, Kok, & Quigley, 2010; Pivo, 2010). The *Collaboration* allows the council to enter into agreements with property owners and finance providers as a way of funding works to improve the environmental performance of those property owners' buildings. Under the *Collaboration* the council and a property owner enter into a public agreement on the future performance of the property owner's building(s). The agreement states the future environmental performance that is to be achieved by the property owner, and stipulates a time frame for achieving this result. The council then lends funds from the bank and provides for the property owner to make upgrades to their building(s). The property owner repays these funds through a local council charge on the land (i.e., through an increased property tax), which is expected to be covered by the reduced energy and maintenance costs of their building(s). Under the *Collaboration* the council gains as it realizes a reduction in the city's carbon emissions, property owners gain as they are provided with low-cost funding to retrofit their buildings and in doing so provide much sought-after 'green' office space, and finance providers gain as they can provide mortgages for a low-risk client (the council).

2.1 Research design

Based on the known availability of examples of experimentation with policy-designs in the building sector (further 'cases'), such as the *Greener Building Collaboration*, and the relevance of the sector in addressing climate change, the sector was chosen as the sectoral case environment for the study reported on in this article. Australia, the Netherlands and the US were selected as the country case environments because these countries are particularly active in the type of policy-design experimentation addressed in this article (Hoffmann, 2011; Jordan, Wurzel, & Zito, 2005), and because of their similarities in case context. That is, the study follows a most-similar systems design, but is interested in relevant differences between the individual cases that may help to gain a better understanding of experimentation in policy-design (Goertz & Mahon, 2012).

A pool of potential cases to study was identified based on an extensive Internet search using keywords such as 'sustainable development AND [country]', 'sustainable building AND [country]', 'green building AND [country]', 'sustainable construction AND [country]' and 'green construction AND [country]'. In addition, the author used social media (predominantly, sustainable and 'green'

building groups on LinkedIn) and his network of policymakers, administrators, architects, engineers, constructors, developers, investors and the like, in Australia, the Netherlands and the US, to gain additional information about potential cases identified in this Internet search. This network was further explored for additional potential cases to study.

From the pool of potential cases to study (roughly 30 in Australia, 20 in the Netherlands and 50 in the US), 12 Australian (including the above-discussed *Collaboration*), 9 Dutch and 11 US cases were selected for further analysis when they met a number of criteria (i.e., a stratified sample). First, they are an example of experimentation with policy-design in terms of the content of the case – i.e., only policy-designs were selected that are different from traditional designs such as direct governmental construction regulation or subsidies (cf., Jordan, et al., 2013). Second, they are an example of experimentation with policy-design in terms of the process of the case – i.e., only policy-designs that were the outcome of a collaborative development process including actors from various backgrounds (i.e., government, industry representatives, citizen representatives) were selected (cf., Gunningham, 2009). Third, they have matured to at least two years of actual implementation – i.e., it was expected that some time is needed for the cases to achieve outcomes. Fourth, the policy-designs selected all move beyond requirements as laid down in building legislation and regulation. The latter criterion was chosen to be able to select policy-designs that may be expected to result in meaningful action.

2.2 Data collection and analysis

In order to understand the policy-design content, policy-design process, implementation process and outcomes of the cases, a series of in-depth face-to-face interviews was carried out (McCracken, 1988; Richards, 1996). Interviewees were selected for their expert knowledge on and experience with one or more of the cases studied (Longhurst, 2003). This sampling resulted in a pool of 99 interviewees (53 Australian, 27 Dutch, 19 US) from various backgrounds – i.e. policymakers, administrators, investors, developers, architects, engineers and property owners. In addition, in the Netherlands a half-day mini-symposium on this topic was organized by the author during the 2012 *International Green Building Week*, which attracted 32 participants.¹ Then, compared to the Australian and Dutch cases studied, the US cases are particularly well documented by both case organisations and third parties and well discussed in the social media. This eased the interview load for the US cases. Table 1 provides a brief overview of these interviewees.

Table 1 – Interviewees' background (mini-symposium participants not included)

Interviewee background	Government	Non-government
Policy maker	8 (4 Aus/4 Dutch/1 US)	
Administrator	26 (22 Aus/4 Dutch/3 US)	15 (12 Aus/3 Dutch/ 13 US)
Architect, engineer, advisor		11 (5 Aus/6 Dutch)
Contractor, developer		7 (3 Aus/4 Dutch)
Property owner		7 (4 Aus/3 Dutch)
Other		6 (3 Aus/3 Dutch /2 US)
Total	38 (26 Aus/8 Dutch/4 US)	61 (27 Aus/19 Dutch/15 US)

¹ The relatively small size of the Netherlands is promising in attracting participants from various backgrounds and from all over the country to a mini-symposium.

Australian interviews were predominantly carried out between September 2011 and February 2012; Dutch interviews were carried out between June and September 2012; and US interviews were carried out between May and June 2013. Interviews were based on a semi-structured questionnaire that provided a structure of checks and balances to assess the validity of the findings (cf. Silverman, 2001). The questionnaire built on a series of guiding variables that are further discussed in section 2.3. Interviews were recorded and transcribed into a report that was sent back to the interviewees for validation (Fielding & Fielding, 1986). Note that the interviewees were often aware of and involved in more than one case. It is expected that this (partly) helps to overcome a sampling bias of administrators (and participants) who are overly enthusiastic about their 'own' policy-design experiment (Sanderson, 2002). Finally, between October 2013 and February 2013 a series of follow-up interviews was carried out with a number of initial interviewees to discuss the development of the policy-design experiments under scrutiny.

The data were first explored using Qualitative Comparative Analysis (QCA) tools, techniques and software (Ragin, 1987; Ragin et al., 2006; Schneider & Wagemann, 2010) to gain an insight into potential patterns of how cases' characteristics may affect their outcomes (further 2.3). Contrary to most conventional techniques, QCA understands cases as configurations of variables and seeks the specific *configurations* of variables that are likely related to a certain outcome. QCA has quickly evolved as an accepted research practice for the type of study presented in this article (Goertz & Mahon, 2012; Rihoux, Rezsöházy, & Bol, 2011). Here fuzzy-set QCA (fsQCA) is applied to gain an understanding of the possible relation between the cases' characteristics and their outcomes, whilst allowing the relative differences between the cases to be brought into the analysis.² The data were further processed by means of a systematic coding scheme (Seale, Gobo, Gubrium, & Silverman, 2004) and qualitative data analysis software (Atlas.ti). By using this approach the data were systematically explored and insights were gained into the 'repetitiveness' and 'rarity' of the experiences shared by the 99 interviewees. Finally, a document study of existing information on these 31 cases and existing research on experimental in policy-design was carried out to cross-check the validity of the data and findings – i.e. triangulation.

2.3 Guiding variables

Before setting out to explore the 31 cases, one question related to the aim of this article remains: what *exactly* is it that an experiment in policy-design should achieve to assess it as a success? In other words, what outcomes may be expected from such experimentation? This seemingly easy question turns out to be difficult to answer, and the extant literature is not yet settled on this question. Some scholars argue that an experiment in policy-design is successful if it generates relevant lessons about how it operates in practice (May, 1992; Petts, 2007; R. Rose, 2001); for instance, about how building contractors understand and are willing to comply with the policy-design. Others argue that a policy-design can only be considered successful when it meets its intended outcomes (cf., Dunn, 2003; Marsh & McConnell, 2010), for instance, a certain number of buildings constructed. For yet others a policy-design is successful if it achieves 'soft outcomes', for

² The author applies QCA methodology, logic and tools mainly to gain an insight into potential relevant patterns in the data collected that are difficult, if not impossible, to trace otherwise. These patterns are then used as a starting point for a more descriptive analysis of the data collected. The author is critical of too strong a reliance on QCA tools as it, eventually, forces researchers to break down their rich qualitative data into quantifiable units, which may work well for some qualitative studies or criteria, but surely not for all (ANONYMOUS).

instance, spillover effects from participants to non-participants in the policy-design experiment, the setting of a focal point, or the normalization of certain (intended) behaviour (cf., Darnall & Sides, 2008; Lyon & Maxwell, 2007).

This article builds on established work in the field of environmental policy by Potoski and Prakash (2009) in evaluating policy-design experiments, and considers (HO.i) the number of participants in a case, and (HO.ii) the number of buildings constructed or retrofitted within a case as a proxy for its success. Both outcome variables are observable and countable, and can be related to the experiments' stated ambitions – making these variables suitable for the aim of the research (cf., Stone, 2002). For example, the above-mentioned *Collaboration* targets over a thousand buildings, but has thus far only resulted in 27 agreements between the city council and property owners for a total of 45 buildings to be retrofitted, and only a handful of buildings that had been or were being retrofitted at the time of this study. It has, however, attracted hundreds of participants since it was launched in 2009 (three years before the research was carried out). Consequently, the programme is evaluated as not having achieved meaningful results in terms of buildings built or retrofitted (HO.ii), but as having achieved meaningful results in terms of participants (HO.i).

Building on the extant policy-design and experimentalist governance literature a series of variables are distilled that are assumed to be related to these outcomes. They relate to the content and process of the policy-design experiment. *Content* characteristics that may positively add to achieving outcomes are: (C.i) a direct financial gain for those participating in the experiment (Heyes & Maxwell, 2004); (C.ii) a strong focus on the private gain for participating, for instance in terms of building close networks with policy makers, limiting access to the market for competitors, or tapping into a potential future client base (e.g., Croci, 2005) – please note, direct financial gain is often found to be the leading 'private interest' for participation, which is why it is considered separately in this article (cf., Van der Heijden, 2012); (C.iii) a strong focus of the policy-design on and explicit reference to the greater public interest (here, addressing climate change) participants serve (e.g., Rivera & de Leon, 2004); and (C.iv) the opportunity to show leadership and being recognized as a leader (e.g., Khanna & Anton, 2002). *Process* characteristics that may positively add to achieving outcomes are: (P.i) an engaged community of participants in terms of participants sharing their knowledge and insights during the design-process (e.g., Solomon, 2008); (P.ii) collaboration and deliberation, and a sharing of decision making powers during the design-process (e.g., Lobel, 2004); (P.iii) financial support for the experimentation process from government (e.g., Héritier & Eckert, 2008); and (P.iv) the opportunity to be able to steer the trajectory of future policy through direct interaction with government (e.g., Termeer, 2009). As with the outcome variables, these explanatory variables are, to a certain extent, measurable (albeit based on the qualitative interview data) and are therefore suitable for the aim of this research (cf., Stone, 2002).

3 Research findings

Table 2 provides a summary of the data by unpacking each case into the key variables discussed (section 2.3). A four-point scale is used to indicate whether a case meets a variable or not, and to give some insight into the relative variance between the cases studied. A '++' indicates a substantial positive score (e.g., there is a very clear focus on and explicit mentioning of the public interest in the policy-design documentation – C.ii; or, there is a very large number of buildings constructed or retrofitted within a case as compared to the case's ambitions – HO.ii); a '+' indicates a moderate positive score (e.g., there is some financial gain relative to other cases studied – C.i); a '-' indicates a moderate negative score (e.g., collaboration was flawed because not all relevant parties were

involved – P.ii; or, the case has only attracted a very small number of intended participants as compared to its ambitions – HO.i); finally, a ‘--’ indicates a substantial negative score on or full absence of the criterion in the case (e.g., there is no opportunity to show leadership or be recognized as a leader – C.iii).³

Table 2 – Summary of data

Cases*	Characteristics of the experimental policy-designs studied**									
	Content				Process				Outcomes	
	C.i	C.ii	C.iii	C.iv	P.i	P.ii	P.iii	P.iv	HO.i	HO.ii
1 (Aus)	++	++	-	++	++	+	+	-	++	++
2 (Aus)	++	++	-	++	++	++	++	+	++	+
3 (Aus)	+	-	++	--	++	+	-	--	++	+
4 (Aus)	+	-	--	-	--	--	--	++	+	-
6 (Aus)	+	+	-	+	+	+	++	-	+	--
9 (Aus)	+	--	+	+	+	++	-	++	++	-
11 (Aus)	+	-	++	+	-	--	--	--	+	-
12 (Aus)	++	--	+	++	-	-	++	--	+	-
13 (Aus)	++	--	+	++	-	-	++	--	+	-
15 (Aus)	+	--	++	--	++	++	+	+	++	-
17 (Aus)	++	+	+	++	++	+	++	--	+	-
18 (Aus)	-	+	-	-	++	+	-	++	+	+
19 (NL)	++	+	--	--	--	--	--	--	+	+
20 (NL)	--	+	+	++	+	+	++	++	+	-
21 (NL)	++	++	--	-	++	+	-	+	--	--
22 (NL)	+	+	-	+	-	+	+	++	+	-
23 (NL)	++	--	+	+	-	+	+	-	+	+
24 (NL)	++	--	+	++	-	-	++	--	+	-
25 (NL)	+	+	+	+	+	+	+	++	+	+
26 (NL)	+	+	-	++	++	++	++	+	++	-
27 (NL)	++	++	-	++	++	++	+	-	++	+
42 (USA)	++	+	-	--	+	-	++	-	+	+
44 (USA)	+	+	-	+	+	+	++	-	-	--
45 (USA)	+	+	-	+	++	-	++	-	+	+
47 (USA)	+	-	++	+	++	+	++	-	++	-
48 (USA)	++	++	-	+	+	-	+	--	+	+
49 (USA)	+	+	+	++	++	+	++	++	+	-
50 (USA)	+	++	++	-	+	-	+	--	+	+
52 (USA)	++	++	+	++	-	--	++	-	-	-
53 (USA)	++	++	--	++	++	++	-	-	++	+
55 (USA)	++	++	-	++	++	+	+	-	++	++

* Please note that this article reports on a study that sits in a larger study, which addresses more than the 31 cases discussed here. This explains why the case numbers surpass the number 31.

** Abbreviations as per section 2.3.

In the following sections the data are first explored using fsQCA techniques and logic (Ragin, 2000; Ragin, et al., 2006) to trace patterns in the data. Through a descriptive analysis the data and these patterns are then further explored (Seale, et al., 2004; Silverman, 2001).

³ For the fsQCA analysis these scores were translated as: ‘++’ equals ‘1’; ‘+’ equals ‘0.66’; ‘-’ equals ‘0.33’; and ‘--’ equals ‘0’ (cf. Ragin et al. 2006).

3.1 QCA analysis

The various policy experiments analysed show better outcomes in terms of the participants they have attracted (HO.i) than in terms of the number of buildings built or retrofitted they have realized (HO.ii). This confirms earlier literature on the subject (Potoski & Prakash, 2009). At question now is what clusters of characteristics are likely related to these outcomes; and whether the cluster(s) of characteristics related to outcome HO.i is(are) similar to or different from the cluster(s) of characteristics related to outcome HO.ii. That is, these clusters may help to gain insight into how cases that do achieve moderate to substantial numbers of participants may be adjusted to be more likely to achieve moderate to substantial numbers of buildings built or retrofitted.

In looking at the data collected through the lens of fsQCA, only two patterns of characteristics stand out that are likely related to the achievement of a moderate to substantial number of participants. These are:

- (I) a financial gain for those participating in the experiment, combined with a strong focus on the private gain for participating (*C.i + C.ii*)
- (II) a financial gain for those participating in the experiment, combined with a strong focus on the greater public interest (here, addressing climate change) participants serve, the opportunity to show leadership and being recognized as a leader, and financial support for the experimentation process (*C.i + C.iii + C.iv + P.iii*)

Then, in looking at data collected through the lens of fsQCA only one pattern of characteristics stands out that is likely related to the achievement of a moderate to substantial number of buildings built or retrofitted:

- (I) a financial gain for those participating in the experiment, combined with a strong focus on the private gain for participating (*C.i + C.ii*)

This seems to imply that within the set of cases studied participants are attracted to participate in policy-design experiments where the financial risks are low: those that resemble free-market situations (I), and those that promise visibility in terms of showcasing leadership with little cost involved for participants (II.) Yet, only cases with a clear focus on participants' financial interest are most likely to result in successful outcomes in terms of substantial numbers of buildings built or retrofitted with high levels of environmental performance (I).

Although an interesting finding, this may be considered bad news for policy-design experimentation as such. After all, this finding seems to imply that predominantly a small part of the *content* of the policy-design experiments (*C.i* and *C.ii*) matters to achieve moderate to substantial outcomes (at least HO.i and HO.ii), whereas the *process* of the policy-design experiments hardly plays a role. At first glance, this seems to go against a large part of the discussed (normative) literature that considers exactly this process to be relevant for experimentation with policy-designs. Further, given that the clusters of characteristics related to outcome HO.i overlap with the cluster of characteristics related to outcome HO.ii, the QCA analysis tells us little about how to improve the policy-designs of these experiments so as to result in more buildings built or retrofitted with high levels of environmental performance.

But the QCA analysis is only a part of the story. Why exactly have so many cases not achieved substantial outcomes in terms of buildings built or retrofitted? And what exactly is the role

of the process of the policy-design experiments studied? To answer these questions a descriptive analysis of the 99 interviews is provided in what follows.

3.2 Attracting participants

In terms of attracting participants, the cases show relative success – related to their stated ambitions. A third of the cases studied have attracted a substantial number of participants (32%, n=10), and more than half of the cases studied have attracted a moderate number of participants (58%, n=18). Interviewee accounts support the outcomes of the QCA analysis, and give some better understanding of why (prospective) participants are especially attracted to policy-design experiments that promise high financial gains, or at least low financial risks, and those that serve their private interest. Further, these interview accounts also provide insights into the difficulty of attracting participants in the 31 cases studied, and possible ways of attracting higher numbers of participants to this type of experiment in policy-design.

Financial gain

In the large majority of the cases (94%, n=29) participation will result in financial gain for the non-state parties involved, with many of the cases analysed having a strong focus on overcoming financial barriers for developers, constructors, owners and users of buildings in reducing carbon emissions. Interviewees often stressed that this financial gain was key to getting participants on board. As a representative of a Dutch builders' umbrella organization explained: 'What our members wish to know is whether or not it is financially viable to go sustainable, and within what boundaries successful outcomes of experiments are likely to be expected' (int69).⁴ Notably, this focus on financial gain achieved some serious criticism from interviewees. 'They [the participants of these cases] are not thinking about the environment, they can make money out of [participating]!' one of the interviewees colourfully expressed (int46). These interviewees were concerned that without the (promise) of a financial gain (prospective) participants would not be interested in developing, buying or occupying buildings with higher levels of sustainability. They further questioned whether a financial incentive truly is a suitable approach to achieve long-term change in the sector.

That having been said, often the financial gain of participating is minimal compared to the huge costs of developing and maintaining buildings, so some interviewees explained: 'the amount of the grant [case#12] is relatively small if you take into account that a development could range somewhere around 80 to 85 million dollars. Is one or one and a half million [the grant provided] an influence as to whether you go ahead with the deal, or not? However, the grant did serve to stimulate a focus on sustainability measures and a learning environment in which the developers and planners were thinking how to go further, how to push boundaries' (int27). This may indicate that the financial gain not only has monetary but also symbolic value for the participants (cf., Cialdini, 2009).

⁴ In line with the custom of qualitative social science research interviewees provided me with their insights in confidence. As such the identities of the interviewees (nor the cases studied) cannot be provided. To give the reader insight in the variance of the interviews voice is given to them by referring to individual interviewees with a number (e.g. 'int50'). Please note that some interviewees are referred to with numbers higher than 99 (the number of interviews used for this article). Please note that this research essay reports on a study that sits in a larger study, which addresses over 50 cases based on over 200 interviews – brief descriptions of *all* the cases in this larger study can be obtained from www.REMOVED FOR REVIEW.info.

Private interest

About two-thirds of the cases studied (67%, n=21) have a clear focus on the private interests of participants when participating in the process of policy-design experimentation. As one of the interviewees explained: 'If you talk to anyone you find that they know right from wrong. You find that they want to do the right thing. With any reasonable opportunity they run for it [sustainable development]. They are happy to do so. They feel better about themselves. But it does not stretch into self-sacrifice very much at all. Self-preservation kicks in over ambition fairly early' (int29). Mostly the self-interest this interviewee is referring to comes in the form of financial gains when participation gives participants a competitive advantage over non-participants, such as in case#1, case#2, case#27 and case#55 (all examples of benchmarking tools). In other cases participation brings non-state actors closer to state actors, which may give them some influence in broader policymaking processes. For instance, an interviewee in case#18 (a particular type of covenant) explained: '[Participation in an experiment] is a bit for the profile of the [participating] organizations. It can also be a point to coordinate activities with governments. Often the steering committee [of an experiment] is formed with various government agencies. It provides a bit of a funnel for interaction' (int51). Also, obtaining information, mostly about how to save costs, was considered a strong motivator for participation. Case#3, case#20, case#45 and case#47, for example, are fully designed about sharing such information among state actors and non-state actors in general, and among participating non-state actors in particular. This all confirms existing literature in this field (cf., Van der Heijden, 2012).

What hampers the attracting of participants?

Not only do interview accounts give an insight into why the majority of experiments have been successful in attracting participants, they also give some insights into why others have not been. Interestingly, interviewees were critical of, in particular, the design *process* of a range of experiments studied.

First, in about a third of the cases (35%, n=11) interviewees were critical of the process of collaboration and deliberation between state and non-state stakeholders in developing the policy-design experiment. Such collaborations are considered relevant in the current policy and governance literature for successful outcomes (e.g. Lobel, 2004; Solomon, 2008). It is expected to result in an understanding about what it is the governed (non-state) actors are willing to change from the current status quo, and what they are (reasonably) able to change. Through a shared process of developing a policy-design it may then be expected that more (prospective) participants are willing to join the experiment as it builds from their experiences (i.e., bottom up) and they feel ownership of the experiment. Interviewees were, however, critical of the practical limitations of collaboration, especially in situations where many actors are involved. It is not difficult to understand that the larger a group of (prospective) participants, the more difficult it is to reach consensus among them. As an administrator of case#17 (which aims for a wide uptake of solar panels in the Dutch rental sector) made clear: 'It is difficult to develop something on a large scale and get everyone on board. There are so many different organizational foci, so many decision-making processes, so many people involved' (int93). In this particular case the majority of the initial participants in the collaborative development process decided not to participate in the implementation of the experiment as they did not recognize their input in the final policy-design experiment.

Second, interviewees noted a potential 'participation fatigue' and 'participation confusion' in the sector as a result of the wide variety of possible experiments to join and the ongoing

introduction of new experiments at local, regional, national and international level. The set of cases studied in this article is but a selected subset of a much larger set of policy-design experiments in the building sector. Although interviewees did not consider competition between the various experiments, they noted the absence of collaboration between or coherence among these.⁵ 'The sustainable movement is very fragmented. Everyone moves in the same direction, but hardly anyone is working together. Some do, but most are in their own little tower,' a senior representative of a private firm in the US noted (int183). Here interviewees considered that more collaboration among the various experiments can take away some of the confusion and fatigue in the sector. An illustrative example here is case#3, an experimental tenant community started in one of Australia's major cities. When implementing the policy-design its administrators had no intention of scaling the experiment beyond the city limits. However, other Australian cities were interested in the experiment as well. Instead of developing their own experiments they followed the 'brand' as introduced by this initial city and introduced local adjustments to it. The policy-design is now implemented in most of Australia's major cities and as a national framework has attracted over 425 tenancies throughout Australia. In particular, the clear 'brand' of the experiment was considered key in attracting participants by the administrators of this experiment in various Australian cities.

How to attract more participants?

When asked directly how to attract more participants interviewees generally referred to the process of the policy-design experiments as being of importance. Following on from the critique of the collaborative processes, they believed that more interaction with potential participants may result in experiments that are indeed sought after by potential participants. In addition, the process of developing policy-designs may in certain cases have been rushed, leaving insufficient time to come to designs that were acceptable for a (potentially) large number of participants, so explained interviewees. For instance, it is telling that in only a few cases (23%, n=7) were various alternatives considered before choosing the final policy-design. In all other cases no alternatives were considered. Furthermore, in over half of the cases studied (61%, n=19) interviewees even referred to a process that resembles prototyping more than actual experimentation (cf., Sanderson, 2002). In these cases the experiments have been designed with the aim implementing them irrespective of the outcomes of the process of experimenting. Interviewees criticized this lack of collaboration and true experimentation during these policy-design processes.

Further, interviewees noted that the various policy-design experiments often attract the same participants. They referred to a small group of truly concerned developers, contractors, architects, engineers, property owners and tenants who are often willing to participate in these type of experiments, whilst the majority of the sector is not. A quote from an administrator of case #53 (which particularly aims to convert that majority) is typical of the tenor among interviewees: 'You cannot push businesses into it. We learnt that however much work we put into outreach you still only reach those who already slightly get it. You don't get the ones who don't have some basic understanding or are already a bit concerned about sustainability. That is the hard lesson we have learnt' (int194). Some interviewees thought that it may be more fruitful to work through the networks of participants to reach out to a larger community of potential participants than stick to more classical approaches such as keeping a website and attending trade fairs. They thought that

⁵ Even more, in many of the cases studied interviewees were not aware of the other cases included in this study – even those that were implemented in, for instance, the same city.

targeting non-participants through the network of participants may create a 'ripple effect' throughout the sector.

3.3 Achieving buildings with high levels of environmental performance

Only a very small number of cases (7%, n=2) has achieved a substantial number of buildings built or retrofitted, and just a bit over a third of the cases has achieved a moderate number of buildings (39%, n=12) – related to the ambitions of the cases. Interviewee accounts largely confirm the outcomes of the QCA analysis, but also shed some more light on why so few cases have, thus far, achieved a substantial number of buildings built or retrofitted.

Financial gain and private interest

The most successful policy-design experiments in terms of buildings built or retrofitted are building benchmarking tools (case#1, case#2, case#27 and case#55). The aim of this type of policy-design is to rank buildings against each other, and communicate a building's relative performance through a label. This further allows building developers or property owners to market their buildings as performing better than those of their competitors who not participate in benchmarking. Such benchmarking or labelling has been found to be successful in other sectors as well (Cashore, Auld, & Newsom, 2004; Cooper & Symes, 2009), and in the buildings sector there is some evidence that buildings in such benchmarking arrangements yield better returns than those that are not (Eichholtz, et al., 2010). In short, benchmarking very clearly serves the self-interest of participants – and especially building developers and property owners.

The other types of experimental policy-designs (e.g., novel forms of financing, competitive grants, tenant communities) proved to be less successful in achieving substantial numbers of buildings built or retrofitted. Here interviewees in all countries pointed to a significant barrier to improving environmental performance in the building sector: a perception of higher upfront costs for buildings with high levels of environmental performance combined with a (perceived) weak economy, often addressed as the global financial crisis (GFC). In these cases it is less clear for participants that building or retrofitting buildings will indeed result in financial gain – this is exactly one of the things the experiments seek to evidence, so interviewees explained. Nevertheless, interviewees expressed high hopes of what their experiments may achieve in the future. Generally, interviewees did not consider a change in economic circumstances necessary to increase the number of buildings built, but an increase of business cases evidences that these types of buildings ultimately are cheaper to operate and maintain than conventional buildings. As the programme manager of a US experimental policy-design of tripartite financing mentioned (case#42, along the lines of the above-mentioned *Collaboration*): 'There probably is a group dynamic. When more projects are getting [built], when there are more success stories, when there is demonstration to the owner in terms of increasing their cash flow, or improving building value, then it will start to become more of a robust market' (int179).

What hampers the achievement of buildings built or retrofitted?

As with participation, interviewees were critical of the process of the policy-design experiments. They considered that without true collaboration, programme administrators are unable to understand what it is that (potential) participants are willing and able to do to build or retrofit buildings with high levels of environmental performance. This indicates that the process of policy-design is more relevant than the QCA analysis initially indicated – which is unsurprising because the

QCA analysis presented looks at the potential causes of success only, and does not consider potential causes of non-success.

In addition, interviewees in a small number of cases (16%, n=5) referred to a mismatch between the policy-design experiment and its legal or regulatory context. In short, although these cases have attracted a moderate to substantial number of interested participants, they cannot yet build or retrofit their buildings as current legislation does not allow the solution the experimental policy-design provides. A typical example is case#6, the above-mentioned *Greener Building Collaboration*. Although the experiment was started in 2009 and had attracted a moderate number of participants by 2011, it was not before 2012 that a change in the city act finally allowed the city to borrow money and give this to commercial property owners for retrofits. For interviewees in other Australian cities this was exactly the reason to treat the case with some care before starting related experiments: 'We are legislatively bound and [copying case#6] would require legislative changes that could take up to eighteen months. Two questions need to be answered first: is [case#6] successful in [city X], and if so, is the effort to implement the programme [here] worth the outcome?' (int27).

How to achieve more buildings with high levels of environmental performance?

When asked directly how the outcomes of these policy-design experiments may be improved, interviewees predominantly pointed towards what they considered the shortcoming of the experiments analysed: participation in all these is voluntary. There is, of course, a clear logic to this. It will normally be politically untenable to mandate participation in an experiment when the (possible) outcomes of the experiment are unclear or uncertain. Yet, as interviewees argued, many experiments have been in force for more than five, and sometimes ten years (e.g., case#55); and many (77%, n=24) have resulted in knowledge that evidences the possibilities of constructing or retrofitting high-performing buildings in a cost-effective way. Making a move from voluntary experiments to mandatory action was considered the most effective way to scale up and speed up the implementation of lessons learnt in these experiments. One interviewee highlighted: 'The speed in which we react is out of sync with the problems we face. Although a lot of [experiments] make sense, they are not fast enough in addressing problems. Regulation is needed' (int33).

Subsequently, interviewees in various cases referred to the importance of peer pressure for ensuring that participants act in line with the goal of the policy experiments. Administrators of policy-design experiments, especially from Australia and the US, referred to the positive effects of rewarding leaders within the experiment. As they further explained, rewarding mechanisms can be relatively low cost and are easily built in to a design – for instance, by showcasing leading buildings on their websites, by running yearly award ceremonies, or by seeking media coverage for leaders in the experiment. Such activities, as interviewees explained, add to the creation of strong communities where acting in line with the ambitions of the policy-design is the norm.

4 Discussion and conclusion

Building on the existing experimentation in policy-design literature, this article has assessed 31 examples of contemporary policy-design experiments in the building sectors of Australia, the Netherlands and the US. These examples are all characterized by experimentation with the content and process of policy-design. Through a QCA analysis the article first sought to understand what clusters of characteristics of these policy-design experiments are likely related to successful outcomes in terms of attracting participants, and in terms of buildings built or retrofitted with high levels of environmental performance. Then, through a descriptive analysis of 99 interviews with key

actors in these experiments and secondary data, the article sought to understand why so many of these 31 policy-design experiments fail to achieve substantial numbers of buildings built or retrofitted with high levels of environmental performance, even though they attracted moderate or substantial numbers of participants.

The QCA analysis confirms much of the existing (qualitative and quantitative) research in this area. Experimentation in environmental policy is likely to be successful if participation comes at low financial risk, and preferably with financial gain (cf., Baron & Diermeier, 2007; Croci, 2005). Yet, interview accounts highlight that even a small monetary gain may be enough to attract participants. Future research may wish to address the relevance of such symbolic monetary gains in policy-design experimentation, and how symbolic gains can best be framed to attract participants. Future research may further wish to address whether objective gains (i.e., evidence based), anticipated or expected gains (i.e., as promised by the policy-design developers), or symbolic gains matter most in attracting participants. Then, the QCA analysis also points to the ability of participants to show leadership and be recognized as a leader as a key criterion for attracting participants to an experiment. Yet, interview accounts are less clear on this aspect of the experiments studied. It may be that those participating in these types of experiments, and especially the administrators of these experiments, are so used to working with leaders in the sector that they consider leadership the norm and not the exception in policy practice (cf., J. Rose, 2011). This issue may be further explored in future research.

The QCA analysis further indicated that in achieving policy outcomes (here defined as attracting participants and achieving buildings with high levels of environmental performance) the content of the policy-design experiments mattered more in the 31 cases studied than the process of experimentation. This is of interest as it partly contradicts the current understanding in the literature of the role of collaboration and participation in the development of policy-designs (e.g., Lobel, 2004; Solomon, 2008). That having been said, the descriptive analysis of the interviews helps to better explain the findings from the QCA analysis: interviewees were rather critical of the characteristics of the process of the experiments. Collaboration was often found to be weak; only in a few cases were content alternatives considered before entering into a testing of the policy-design; and in a number of cases studied the process of experimenting came closer to prototyping than an actual open experimentation process. In light of these critiques, then, it is not surprising that the experiments perform poorly. Future research may continue the line of research started in this article by carefully studying policy-design experiments with less flawed processes, aiming to understand the exact role and impact of the policy-design process on the experiments' outcomes.

The descriptive analysis further taught two major lessons related to the ongoing proliferation of policy-design experiments. On the one hand, this proliferation may result in 'participation fatigue' or 'participation confusion' when (prospective) participants are either asked too often to participate in experiments, or when they simply cannot decide on what experiment to join as these all promise different outcomes and require different input. This is a relevant lesson for those wishing to start a new experiment in policy-design. Instead of adding a fully novel experiment to the pool of existing experiments it may be wiser to join an existing one (local, regional, national, or international). The current literature on experimentation in policy-designs indicates that all over the world a very wide range of policy-designs is being experimented with, making it likely that an existing experiment can be joined. By teaming up with an existing experiment resources can be shared in terms of marketing the experiment and attracting participants.

On the other hand, the proliferation of experiments may result in competition for participants between the experiments. Although interviewees did not directly refer to such

competition, some interviewees were aware of how their experiment performs compared to (certain) others and acted to this. 'We keep the [performance] data to ourselves for now. We are being compared with other [experiments]. [Experiments] that run much longer. We will show our data when there are clear success stories to show,' one of the interviewees stated (int187). There is, of course, a clear logic for such behaviour. Successful experiments are more likely to attract more participants and government funds. However, such (perverse) behaviour seems to go exactly against the goal of experimentation: drawing lessons on whether and how novel policy-designs achieve their intended goals. This is the value of experimentation. It is not about evidencing that a policy-design is able to achieve (some) success stories. Future research may wish to critically assess the reported outcomes of policy-design experiments, aiming to understand how often and under what circumstances administrators refrain from presenting the full story of the experimentation.

Finally, and most strikingly, the majority of the experiments studied (87%, n=27) aimed to evidence something obvious: that buildings with high levels of environmental performance are not more costly, and are sometimes cheaper to build, operate, maintain and use than traditional buildings with low levels of environmental performance. Yet, these opportunities and possibilities have been known for some decades now and are very well documented (e.g, Beddoes & Booth, 2012; Register, 1987). Why then repeat the same experiment, albeit in different forms, over and over and over again? This may point to the feeling that, at least in the building sector, there is something inherently wrong with policy-design experimentation: the experiments are not truly novel policy-designs, but different approaches to treating the same symptom. The symptom here is that existing technology and knowledge are not used by actors in the building sector. The core of the problem, according to many of the interviewees, is, however, not unwillingness to use this technology and knowledge, but unfamiliarity with this technology and knowledge, and unfamiliarity with what it may achieve in terms of reduced carbon emissions and financial savings. Thus, instead of keeping a narrow symptom focus, those involved in policy-design experimentation may wish to focus on the core of the problem and develop policy-designs that address this. Cases #3 and #47, for instance, are illustrative examples where participants are educated about possibilities, and learn from each other's solutions to improving the environmental performance of their buildings. Sharing lessons and re-educating actors in the building sector based on already widely available knowledge may be a more fruitful approach to achieving a necessary scaling up of sustainable performance in this sector (cf., Biermann, 2008) than reinventing the wheel over and over again.

To conclude, this research has shed some new light on contemporary policy-design experimentation in the building sector, and has raised serious concerns about such experimentation. These lessons may be taken to other sectors as well, and may be further explored in future research. After all, the characteristics of the building sector give promise to positive outcomes of experimental policy-designs that seek to improve the environmental performance of the sector. With other sectors sharing far fewer of these positive characteristics (cf., IPCC, 2007) it may be a question of whether the rather poor performance of the cases studied in the building sector echoes in other sectors as well, or is even magnified due to less positive characteristics – or, unexpectedly, is not. That having been said, asking whether experimentation in policy-design generally works or not is likely to be the wrong question. The study presented highlights that experimentation in policy-design sometimes works, and sometimes does not. Rather, it is more important to learn about the specific conditions under which experimentation achieves its outcomes and how. By carefully unpacking 31 real-world examples of policy-design experimentation this article has taken a small step toward trying to better understand these conditions.

References

- Abaire, J. (2008). Green Buildings: What It Means To Be "Green" and the Evolution of Green Building Laws. *The Urban Lawyer*, 40(3), 623-632.
- Baron, D. P., & Diermeier, D. (2007). Strategic Activism and Nonmarket Strategy. *Journal of Economics and Management Strategy*, 16, 599-634.
- Bastianoni, S., Galli, A., Niccolucci, V., & Pulselli, R. M. (2006). The ecological footprint of building construction. *Ecology and the Environment*, 93, 345-356.
- Beddoes, D., & Booth, C. (2012). Insights and Perceptions of Sustainable Design and Construction. In C. Booth, F. Hammond, J. Lamond & D. Proverbs (Eds.), *Solutions for Climate Change Challenges in the Built Environment* (pp. 127-140). Oxford: Wiley-Blackwell.
- Biermann, F. (2008). Earth Systems Governance. In O. Young, L. King & H. Schroeder (Eds.), *Institutions and Environmental Change* (pp. 277-301). Cambridge, MA: MIT Press.
- Brown, M. A. (2001). Market failures and barriers as a basis for clean energy policies. *Energy Policy*, 29(14), 1197-1207.
- Campbell, D. (1969). Reforms as experiments. *American Psychologist*, 24(4), 409-429.
- Cashore, B., Auld, G., & Newsom, D. (2004). *Governing Through Markets: Forest Certification and the Emergence of Non-State Authority*. New Haven: Yale University Press.
- Cialdini, R. (2009). *Influence*. Boston: Pearson.
- Cooper, I., & Symes, M. (2009). *Sustainable Urban Development. Volume 4. Changing Professional Practice*. London: Routledge.
- Croci, E. (2005). *The Handbook of Environmental Voluntary Agreements*. Dordrecht: Springer.
- Darnall, N., & Sides, S. (2008). Assessing the Performance of Voluntary Environmental Programs: Does Certification Matter? *Policy Studies Journal*, 36(1), 95-117.
- De Burca, G. (2010). New Governance and Experimentalism: An Introduction. *Wisconsin Law Review*, 227, 227-238.
- De Búrca, G., & Scott, J. (2006). *New governance and constitutionalism in Europe and the US*. Oxford: Hart.
- Dewey, J. (1991 [1927]). *The Public and Its Problems*. Ohio: Swallow Press.
- Dunn, W. N. (2003). *Public policy analysis an introduction*. Harlow: Prentice Hall.
- Eichholtz, P., Kok, N., & Quigley, J. (2010). Doing Well by Doing Good? Green Office Buildings. *American Economic Review*, 100, 2492-2509.
- Evans, B., Joas, M., Sundback, S., & Thobald, K. (2005). *Governing Sustainable Cities*. London: Earthscan.
- Fielding, N. G., & Fielding, J. L. (1986). *Linking Data*. London: Sage.
- Goertz, G., & Mahon, J. (2012). *A Tale of Two Cultures*. Princeton: Princeton University Press.
- Greenberg, D., & Schroder, M. (2003). *The Digest of Social Experiments* (3rd ed.). Washington: Urban Institute Press.
- Gunningham, N. (2009). The New Collaborative Governance: The Localization of Regulation. *Journal of Law and Society*, 36(1), 145-166.
- Héritier, A., & Eckert, S. (2008). New Modes of Governance in the Shadow of Hierarchy. *Journal of Public Policy*, 28(1), 113-138.
- Heyes, A. G., & Maxwell, J. W. (2004). Private vs. Public Regulation: Political Economy of the International Environment. *Journal of Environmental Economics and Management*, 48(2), 978-996.
- Hickson, K. (2009). *The ABC of Carbon*. Toowong: ABC Carbon.
- Hoffmann, M. (2011). *Climate Governance at the Crossroads*. Oxford: Oxford University Press.
- IPCC. (2007). *Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
- Jordan, A., Wurzel, R., & Zito, A. (2005). The Rise of New Policy Instruments in Comparative Perspective: Has Governance Eclipsed Government? *Political Studies*, 53(3), 477-496.

- Jordan, A., Wurzel, R., & Zito, A. (2010). 'New' Instruments of Environmental Governance. *Environmental Politics*, 12(1), 1-24.
- Jordan, A., Wurzel, R., & Zito, A. (2013). Still the century of 'new' environmental policy instruments? *Environmental Politics*, 22(1), 155-173.
- Khanna, M., & Anton, W. R. Q. (2002). Corporate Environmental Management: Regulatory and Market-Based Incentives. *Land Economics*, 78(4), 539-558.
- Lobel, O. (2004). The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought. *Minnesota Law Review*, 89(2), 263-293.
- Longhurst, R. (2003). Semi-structured Interviews and Focus Groups. In N. J. Clifford & G. Valentine (Eds.), *Key methods in geography* (pp. 117-132). London: Sage Publishers.
- Lyon, T. P., & Maxwell, J. W. (2007). Environmental Public Voluntary Programs Reconsidered. *The Policy Studies Journal*, 35(4), 723-750.
- Markusson, N., Atsushi, I., & Stephens, J. (2011). The social and political complexities of learning in carbon capture and storage demonstration projects. *Global Environmental Change*, 21(2), 293-302.
- Marsh, D., & McConnell, A. (2010). Towards a framework for establishing policy success. *Public Administration*, 88(2), 564-583.
- May, P. J. (1992). Policy Learning or Failure. *Journal of Public Policy*, 12(4), 331-354.
- McCracken, G. (1988). *The long interview* (Vol. 13). London: Sage.
- Newman, P., Beatley, T., & Boyer, H. (2009). *Resilient Cities*. Washington: Island Press.
- Overdevest, C., & Zeitlin, J. (2012). Assembling an experimentalist regime: Transnational governance interactions in the forest sector. *Regulation & Governance*, 6(x), 1-29.
- Payne, G., & Williams, M. (2005). Generalization in Qualitative Research. *Sociology*, 39(2), 295-314.
- Petts, J. (2007). Learning about learning: lessons from public engagement and deliberation on urban river restoration. *The Geographical Journal*, 173(4), 300-311.
- Pivo, G. (2010). Owner-Tenant Engagement in Sustainable Property Investing. *The Journal of Sustainable Real Estate*, 2(1), 183-199.
- Potoski, M., & Prakash, A. (2009). *Voluntary Programs: a club theory perspective*. Cambridge: MIT Press.
- Ragin, C. C. (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Ragin, C. C. (2000). *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.
- Ragin, C. C., Robinson, C., Schaefer, D., Anderson, S., Williams, E., & Giesel, H. (2006). *User's guide to Fuzzy-Set / Qualitative Comparative Analysis*. Tuson: University of Arizona.
- Register, R. (1987). *Ecocity Berkeley: Building Cities for a Healthy Future*. Berkeley: North Atlantic Books.
- Richards, D. (1996). Elite Interviewing: Approaches and Pitfalls. *Politics*, 16(3), 199-204.
- Rihoux, B., Rezsöházy, I., & Bol, D. (2011). Qualitative Comparative Analysis in Public Policy Analysis. *German Policy Studies*, 7(3), 9-82.
- Rivera, J., & de Leon, P. (2004). Is Greener Whiter? Voluntary Environmental Performance of Western Ski Areas. *The Policy Studies Journal*, 32(3), 417-437.
- Rose, J. (2011). Diverse Perspectives on the Groupthink Theory. *Emerging Leadership Journeys*, 4(1), 37- 57.
- Rose, R. (2001). *Ten steps in learning lessons from abroad*. Hull: Future Governance Programme.
- Sabel, C., & Zeitlin, J. (2011). Experimentalist Governance. In D. Levi-Faur (Ed.), *The Oxford Handbook of Governance* (pp. 169-185). Oxford: Oxford University Press.
- Sabel, C., & Zeitlin, J. (2012). Experimentalism in the EU. *Regulation & Governance*, 6(3), 410-426.
- Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, 80(1), 1-22.
- Schneider, C., & Wagemann, C. (2010). Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets. *Comparative Sociology*, 9(3), 397-418.

- Seale, C., Gobo, G., Gubrium, J. F., & Silverman, D. (2004). *Qualitative Research Practice*. London: Sage.
- Silverman, D. (2001). *Interpreting Qualitative Data. 2nd edition*. London: Sage.
- Solomon, J. (2008). Law and Governance in the 21st Century Regulatory State. *Texas Law Review*, 86, 819-856.
- Stone, D. A. (2002). *Policy Paradox. The art of political decision making* (Revised Edition ed.). New York: Norton.
- Termeer, C. (2009). Barriers to New Modes of Horizontal Governance. *Public Management Review*, 11(3), 299-316.
- Van der Heijden, J. (2012). Voluntary Environmental Governance Arrangements. *Environmental Policies*, 21(3), 486-509.
- Van der Heijden, J. (2013). Greening the building sector: Roles for building surveyors. *Journal of Building Survey, Appraisal & Valuation*, 2(1), 24-32.
- Vreugdenhil, H., Taljaard, S., & Slinger, J. (2012). Pilot Projects and Their Diffusion: a Case Study of Integrated Coastal Management in South Africa. *International Journal Of Sustainable Development*, 15(1/2), 148-172.
- Yudelson, J., & Meyer, U. (2013). *The World's Greenest Buildings*. Abingdon: Routledge.
- Zimmerman, A., & Martin, M. (2001). Post-occupancy evaluation: benefits and barriers. *Building Research & Information*, 29(2), 168-174.