



Researching Flood Risk Governance in Europe: background theories

Corinne Larrue, Dries Hegger, Jean-Baptiste Trémorin (Eds)

Date: November 2013
Report Number: D2.2.2
Revision number: 6

Milestone Number: MS1: Assessment Framework for Flood Risk Governance
Due date for deliverable: 15 November 2013
Actual submission date: 14 November 2013

STAR-FLOOD receives funding from the EU 7th Framework programme (FP7/2007-2013) under grant agreement 308364

Document Dissemination Level
PU Public

Co-ordinator: Utrecht University
Project Contract No: 308364
Project website: www.starflood.eu

Document information

Work Package	2
Consortium Body	WP Team
Year	2013
Document type	Draft Deliverable
Status	Draft for Approval
Date	15 November 2013
Editors	Corinne Larrue, Dries Hegger, Jean-Baptiste Trémorin (Eds)

Document History

Date	No.	Prepared by	Organisation	Approved by	Notes
2 July 2013	1	Editors	Université François Rabelais de Tours	-	Discussed at WP2 meeting in Paris
4 September 2013	2	Editors	Université François Rabelais de Tours	-	Sent out for feedback in writing
20 September 2013	3	Editors	Université François Rabelais de Tours	-	Discussed at plenary consortium meeting
24 October 2013	4	Editors	Université François Rabelais de Tours	-	Draft for approval by project coordinator
15 November 2013	5	Editors	Université François Rabelais de Tours	Utrecht University	Version to be submitted to EC

Acknowledgement

The work described in this publication was supported by the European Community's Seventh Framework Programme through the grant to the budget of the Integrated Project STAR-FLOOD, Contract 308364.

Disclaimer

This document reflects only the authors' views and not those of the European Community. This work may rely on data from sources external to the STAR-FLOOD project Consortium. Members of the Consortium do not accept liability for loss or damage suffered by any third party as a result of errors or inaccuracies in such data. The information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and neither the European Community nor any member of the STAR-FLOOD Consortium is liable for any use that may be made of the information.

© STAR-FLOOD Consortium. This report should be referred to as follows:

Larrue, C., Trémorin, J.B., Hegger, D.L.T. (2013), Researching Flood Risk Policies in Europe: a framework and methodology for assessing Flood Risk Governance (report no D2.2.2), STAR-FLOOD Consortium, Utrecht, The Netherlands, ISBN: 978-94-91933-00-4.

Keywords

Appropriateness; Assessment Framework; Country Analysis; Embedded Case Study Analysis; Effectiveness; Efficiency; Explaining Policy Change; Flood Risk Governance Arrangements; Legitimacy; Policy Analysis; Policy Evaluation; Resilience.

Editor information

Prof. Corinne Larrue is Professor in Planning and Environmental Public Policies at the University of Tours and the University Paris XII.

Dr. Dries Hegger is post doc researcher in Environmental Governance at the Copernicus Institute of Sustainable Development, Faculty of Geosciences at Utrecht University.

Jean-Baptiste Trémorin is a research assistant at the University of Tours.

Author information

Dr. Meghan Alexander is a postdoctoral research fellow at the Flood Hazard Research Centre of Middlesex University.

Dr. Marloes Bakker is post doc researcher in Environmental Governance at the Copernicus Institute of Sustainable Development, Faculty of Geosciences at Utrecht University.

Dr. Anoeska Buijze is a postdoctoral research fellow in Administrative Law at the Faculty of Law, Economics and Governance (REBO) of Utrecht University.

Dr. Ann Crabbé is senior researcher and member of the research group Society and Environment of the University of Antwerp.

Dr. Carel Dieperink is assistant professor in Environmental Governance at the Copernicus Institute of Sustainable Development, Faculty of Geosciences at Utrecht University.

Willemijn van Doorn-Hoekveld MA is a PhD candidate in Environmental Law at the Research Centre Water and Sustainability of the Faculty of Law, Economics and Governance (REBO) of Utrecht University.

Prof. Peter Driessen is professor in Environmental Governance at the Copernicus Institute of Sustainable Development, Faculty of Geosciences at Utrecht University. He is also scientific Director of the Dutch National Research Programme Knowledge for Climate.

Dr. Marie Fournier is associated professor in Geography in University of Mulhouse and research fellow associated to the research centre CITERES of the Université François Rabelais de Tours.

Dr. Mathilde Gralepois is assistant professor in planning and local governance at the University of Tours.

Prof. Colin Green is professor of Water Economics at the Flood Hazard Research Centre of Middlesex University.

Marlous van Herten is project assistant for the STAR-FLOOD project at the Environmental Governance Section, Copernicus Institute of Sustainable Development, Faculty of Geosciences of Utrecht University.

Dr. Duncan Liefferink is assistant professor in the Department of Political Sciences of the Environment at Radboud University Nijmegen.

Dr Corinne Manson is Associated Professor in Public Law at the University of Tours.

Dr. Sally Priest is a senior research fellow at the Flood Hazard Research Centre of Middlesex University.

Prof. Marleen van Rijswijk is professor in European and Dutch Water Law at the Faculty of Law, Economics and Governance (REBO) of Utrecht University.

Cathy Suykens LL.M is a research fellow at the Institute for Environmental and Energy Law at the KU Leuven.

Dr. Mark Wiering is associate Professor in the Department of Political Sciences of the Environment at Radboud University Nijmegen.

Preface

This report is a deliverable of the EU 7th Framework Project STAR-FLOOD (www.starflood.eu). STAR-FLOOD focuses on flood risk governance. The project investigates strategies for dealing with flood risks in 18 vulnerable urban regions in six European countries: England and Scotland in the UK, Belgium, France, The Netherlands, Poland and Sweden. The project assesses the institutional embedding of these strategies from a combined public administration and legal perspective, with the aim to make European regions more resilient to flood risks.

Together with the report entitled "Researching Flood Risk Governance in Europe: a framework and methodology for assessing Flood Risk Governance" this report constitutes Deliverable D2.2, the main deliverable of Work Package 2. Whereas Work Package 1 provided an extended problem analysis related to Flood Risk Governance in Europe, Work Package 2 focuses on how Flood Risk Governance in Europe can be researched.

The report entitled "Researching Flood Risk Governance in Europe: a framework and methodology for assessing Flood Risk Governance" lays down an assessment framework and methodology for the empirical analyses that will be carried out within the STAR-FLOOD project. It focuses on providing guidance for researchers on how to do the empirical analyses. The assessment framework is, however, based on a thorough review of social scientific, legal and economic literature relevant for Flood Risk Governance. A more detailed elaboration of this literature is given in the current report. The two reports together focus on how stability and dynamics in Flood Risk Governance, at the country level, can be analysed, explained and evaluated.

The two reports of Deliverable D2.2 provide an extended elaboration on how the STAR-FLOOD researchers should go about the empirical research in order to maximise the scientific quality of the work as well as its societal relevance. The reports should be seen as working documents. Throughout the empirical research, further conceptual refinement as well as further operationalisation of STAR-FLOOD's key concepts will take place. It is also expected that the consortium's insights in how country-specific results can be compared and subsequently translated into good practices will progress in the course of the empirical work.

We trust that the current report is of interest for a broad readership. Besides the young researchers within the STAR-FLOOD consortium, the content of this report may also inspire other researchers and professionals with an interest in social scientific and legal research into Flood Risk Management, Disaster Risk Reduction or climate change adaptation.

Yours sincerely,

Prof. Corinne Larrue
Leader of WP2

Prof. Peter Driessen
STAR-FLOOD Project Coordinator

Executive Summary

This report, together with the report "Researching Flood Risk Governance in Europe: a framework and methodology for assessing Flood Risk Governance" forms the main deliverable for the Second Work Package of the EU 7th Framework Project STAR-FLOOD. Whereas the framework and methodology report mainly serves to provide guidance for researchers, the current report provides further theoretical background thereby grounding the framework in the literature. It also provides space for some further elaboration on issues that could only be touched upon in the guidance report. Both reports are expected to be equally important, but probably in different stages of the empirical research.

The report is structured according to the three main analytical steps that have been distinguished within STAR-FLOOD, being those of analysing, explaining and evaluating Flood Risk Governance in Europe. To each of these analytical steps one chapter is devoted.

Chapter 2 presents an extract of two texts on the Policy Arrangements Approach, the theoretical backbone of STAR-FLOOD's analytical framework. The first text is an extract of a book chapter by Duncan Liefferink entitled 'The dynamics of policy arrangements: turning round the tetrahedron'. The second text is an extract from: Wiering M, Arts, B 2006, Discursive shifts in Dutch river management: 'deep' institutional change or adaptation strategy? In: *Hydrobiologia*, vol. 565, pp. 327-338. Both texts together provide an overview of the four dimensions of the Policy Arrangements Approach – Actors, Discourses, Rules, Resources, examples of how they can be operationalised, also with regard to water management issues, and some reflection on how the four dimensions can be brought together into a characterisation of Policy Arrangements as a whole.

Chapter 3 provides an overview of background theories related to STAR-FLOOD's explanatory framework. The chapter first discusses the main features of some prominent explanatory theories from the policy sciences literature, including the Multiple Streams Framework, Punctuated Equilibrium Theory, and the Advocacy Coalitions Framework as well as publications closely related to or critically reflecting upon the frameworks. The chapter then introduces the Multi-Level Perspective from Transitions Theory and discusses some syntheses between all aforementioned explanatory frameworks. Subsequently, some attention is devoted to a specific issue of relevance for efforts at explaining policy change, that is, the relationship between social structures and the purposeful actions of agents involved in Flood Risk Governance. The chapter goes on by discussing some of the literature on change agency and policy entrepreneurship. The last part of the chapter is devoted to reflections on the relationship between explanations by policy analysts and legal scholars; the relevance of natural law and positive law studies for explaining stability and dynamics in Flood Risk Governance; and specific explanations for stability and dynamics *in* legal frameworks.

Chapter 4, finally, provides further background to the evaluation framework. The chapter starts off with a general reflection on evaluating governance, including an elaboration of previous experiences with evaluating governance. Subsequently, the literature on each one of STAR-FLOOD's evaluation criteria is discussed in some detail: resilience, appropriateness, legitimacy, efficiency and effectiveness. The chapter then reflects on how the sub-criteria of legitimacy, efficiency and effectiveness could be combined into the meta-criterion of appropriateness. The chapter concludes with a discussion on operational challenges for evaluation and a summary of the evaluation framework as a whole.

Contents

1. Introduction	13
<i>Corinne Larrue and Marie Fournier</i>	
1.1 The STAR-FLOOD project	13
1.2 Position of this report	13
1.2.1 STAR-FLOOD: where are we after Work Package 1?	13
1.2.2 Aim and scope of Work Package 2	13
1.2.3 Process followed in Work Package 2	14
1.3 Headlines of the conceptual and methodological approach	15
1.4 Outline of the report and guide for the reader	15
2 The Policy Arrangements Approach: introduction and basic ingredients for its operationalisation	17
<i>Mark Wiering and Duncan Liefferink</i>	
2.1 Introduction	17
2.2 The concept of policy arrangements	17
2.2.1 Introduction	17
2.2.2 Analysing policy arrangements: a matter of perspective	19
2.2.3 Actors/coalitions	19
2.2.4 Resources/power	22
2.2.5 Rules	23
2.2.6 Discourses	24
2.2.7 Summing up	25
2.3 Discursive shifts in Dutch river management: ‘deep’ institutional change or adaptation strategy?	26
2.3.1 Abstract	26
2.3.2 Discourse	26
2.3.3 Rules	27
2.3.4 Actors and coalitions	27
2.3.5 Resources	27
2.3.6 Discursive shifts in river management	28
2.3.7 Changing rules	30
2.3.8 Policy actors	32
2.3.9 Shifts in power	33
3 Theories on explaining policy and legal change	37
<i>Dries Hegger, Peter Driessen, Marlous van Herten, Anoeska Buijze, Jean-Baptiste Trémorin, Willemijn Van Doorn-Hoekveld, Marleen van Rijswijk and Carel Dieperink</i>	
3.1 Introduction	37
3.2 The Multiple Streams Framework (MSF)	38
3.3 Punctuated Equilibrium Theory (PET)	39
3.4 Advocacy Coalitions Framework (ACF)	40
3.5 The Multi-Level Perspective from transitions theory	41
3.5.1 Introduction	41
3.5.2 Explaining stability and change from a Multi-Level Perspective	41
3.6 Some syntheses between the policy analytical frameworks for explaining policy change	43
3.7 The relationship between structure and agency	44
3.8 Change agency literature	45
3.9 Relationship between explanations by policy analysts and legal scholars	47

3.10 Natural law, positive law and their relevance for explain stability and dynamics in flood risk governance.....	48
3.11 Stability and dynamics in the legal framework.....	48
3.11.1 Introduction.....	48
3.11.2 Stability in legal frameworks.....	49
3.11.3 Instability of the legal framework.....	49
3.11.4 Dynamics in legal frameworks.....	50
3.11.5 Conclusion.....	50
4 Theoretical background to the evaluation framework.....	51
<i>Sally Priest, Meghan Alexander, Colin Green, Anoeska Buijze, Willemijn Van Doorn-Hoekveld</i>	
4.1 Introduction.....	51
4.2 The purpose and logic of evaluation.....	51
4.2.1 Introduction.....	51
4.2.2 The purpose of evaluation.....	51
4.2.3 A logic of evaluation.....	52
4.3 Evaluating governance.....	54
4.3.1 The object of evaluation.....	54
4.3.2 Previous examples of evaluating governance.....	56
4.3.3 Methodological reflections for evaluating governance in STAR-FLOOD: the challenge of using indicators.....	64
4.3.4 Examining criteria for evaluation: What are we trying to evaluate and why?.....	67
4.4 Introducing evaluation criteria.....	68
4.4.1 A theoretical overview of resilience.....	68
4.4.2 Appropriateness.....	73
4.4.3 A theoretical overview of Effectiveness.....	74
4.4.4 Challenges of evaluating effectiveness.....	79
4.4.5 A theoretical overview of Efficiency.....	80
4.4.6 A theoretical overview of Legitimacy.....	83
4.4.7 Potential indicators of legitimacy.....	93
4.5 Operationalising the concept of Appropriateness: Combining the sub-criteria of legitimacy, effectiveness and efficiency.....	95
4.6 Operational challenges for evaluation.....	98
4.7 Evaluation framework: Summary.....	101
References.....	103

List of figures

2.1: The tetrahedron, symbolising the interconnectedness of the four dimensions of a policy arrangement	18
2.2: Example of a ‘map’ of actors and their relative positions in a policy arrangement	20
2.3: The tetrahedron: analytical perspectives	26
3.1 Multiple levels as a nested hierarchy (2002)	43
4.1: Clarifying the object of evaluation: Distinguishing governance arrangements, strategies and measures (where “X” indicates a number of possibilities).	55
4.2 A Framework for Programme Evaluation (CDC 1999)	56
4.3 Adaptive cycle and key ecosystem functions (r , K , Ω , α) (from Holling 2001)	70
4.4 Panarchical connections (from Holling 2001)	71
4.5: Using radial diagrams to visualise the “shapes” of different FRGAs according to the selected evaluation criteria	97
4.6: Clarifying the object of evaluation and applicable criteria	100

List of tables

2.1: Operationalization of the Policy Arrangement Approach (PAA)	28
3.2: Strategies of policy entrepreneurs found in the Dutch water management domain	46
4.1 Seven ideal logical requirements for evaluation criteria	52
4.2: Distinguishing stages for which evaluation criteria can be assessed	55
4.3 Criteria to evaluate Abrams et al.’s (2003) principles of good governance	57
4.4 Lockwood et al.’s eight principles for governance in natural resource management	60
4.5 Hierarchical framework for the assessment of governance quality (Cadman 2012: 21)	61
4.6 Evaluation criteria for social and institutional systems (after Ostrom 2006: 10)	61
4.7 Hypotheses developed by Newig and Fritsch (2009) to investigate the complex relationships between participatory, multi-level and scale-adapted governance and the effectiveness of environmental policy.	63
4.8 Example of a framework for organizing agreed governance principles, criteria and indicators (Abrams et al. 2003; 55)	66
4.9 Ecosystem functions organised into the different stages described in the “adaptive cycle” (Holling 2001; Holling et al. 2002)	69
4.10: Key principles of economic efficiency (Winch 1971)	80
4.11 Potential indicators and sub-indicators for evaluating public participatory processes	93
4.12 Main questions covered by the six intensity dimensions	95
4.13: Dominant GPPA depending on the stage of administrative decision making	98

List of boxes

1.1: Research questions of the STAR-FLOOD project	14
3.1 Examples of questions that could be posed inspired by the MSF	39
3.2 Examples of questions that could be posed inspired by PET	40
3.3 Examples of questions that could be posed using the Advocacy Coalitions Framework	41
3.4 Examples of questions that could be posed inspired by the MLP	43

3.5 Examples of questions that could be posed inspired by literature on change agency	46
4.1: A definition of evaluation	51
4.2: Characteristics of good indicators (Abrams et al. 2003: 38)	65
4.3 Definition of resilience	72
4.4 Some higher level reflections on effectiveness and comparison	78
4.5 Efficiency can be defined as	82

1. Introduction

Corinne Larrue and Marie Fournier

1.1 The STAR-FLOOD project

This report is a deliverable of the EU 7th Framework Project STAR-FLOOD (see www.starflood.eu for an outline of the project). STAR-FLOOD focuses on Flood Risk Governance. The project investigates strategies for dealing with flood risks in 18 vulnerable urban regions in six European countries: England and Scotland in the UK, Belgium, France, The Netherlands, Poland and Sweden. The project is assessing the institutional embedding of these strategies from a combined public administration and legal perspective, with the aim to make European regions more resilient to flood risks.

1.2 Position of this report

1.2.1 STAR-FLOOD: where are we after Work Package 1?

Deliverables for Work Package 1 proposed an extended problem analysis related to Flood Risk Governance in Europe. The different reports discussed the actual flood risks in the STAR-FLOOD consortium countries (report no D1.1.1, Green et al. 2013), the governance challenges related to Flood Risk Management (report no D1.1.2, Dieperink et al. 2013), European flood regulation (report no D1.1.3, Bakker et al. 2013) and the similarities and differences between the STAR-FLOOD consortium countries (report no D1.1.4, Hegger et al. 2013). In so doing, they gave a further specification of the scope of the STAR-FLOOD project and raised some preliminary conclusions, expectations and assumptions to be challenged in the subsequent Work Packages of the project.

More specifically, Work Package 1 clarified the main objective of the STAR-FLOOD project: the analysis of Flood Risk Management Strategies (FRMSs) and the Flood Risk Governance Arrangements (FRGAs) within which they are embedded. FRGAs can be defined as “the constellation resulting from a dynamic interplay between actors and actor coalitions involved in all policy domains relevant for Flood Risk Management – including water management, spatial planning and disaster management; their dominant discourses; formal and informal rules of the game; and the power and resource base of the actors involved” (Report D1.1.4: Hegger et al. 2013). This definition stresses that FRGAs have an actor dimension, a rule dimension, a power and resource dimension and a discursive dimension. By focusing on FRGAs, the STAR FLOOD consortium wants to get a better insight into the societal aspects of FRMSs and the way they are institutionally embedded in a broad sense. The concept allows insights from policy scientists and legal scholars to be combined and enables researchers to assess Flood Risk Governance from a combined perspective.

1.2.2 Aim and scope of Work Package 2

This report, together with the report "Researching Flood Risk Governance in Europe: a framework and methodology for assessing Flood Risk Governance" forms the main deliverable of the second Work Package of STAR-FLOOD. Whereas the first Work Package provided an extended problem analysis related to Flood Risk Governance in Europe, the focus of the second Work Package is on how Flood Risk Governance in Europe should be researched. Our assessment framework is based on a thorough review of social scientific, legal and economic literature relevant for Flood Risk Governance. Whereas the framework and methodology report mainly aims to provide guidance to researchers, the current report provides a more detailed elaboration of the literature on which the framework is based. The two reports together focus on how stability and dynamics in Flood Risk Governance, at the country level, can be analysed, explained and evaluated.

Box 1.1 lists the research questions that are being addressed in the STAR-FLOOD project. The assessment framework and methodology laid down in the guidance report are meant to enable sub-questions 5-9 to be answered in a consistent way in Work Package 3. Question 1-3 have been addressed in Work Package 1. Question 4 has been addressed in Work Package 1 and will be further addressed in Work Package 3. Question 10 will be addressed in Work Package 5.

Box 1.1: Research questions of the STAR-FLOOD project

i) Sub-goal 'identifying':

1. What are the main trends in and challenges for Flood Risk Governance in Europe?
2. What are the key elements of FRGs discussed in literature?
3. What kind of FRGs are characterised as 'good practice' in scientific and policy literature?
4. Which FRGs are developed and applied in different urban agglomerations in the selected countries?

ii) Sub-goal 'analysing':

5. What are the *historical dynamics* (or the absence thereof) of FRGs in the selected EU member states?

iii) Sub-goal 'explaining':

6. Which factors *explain* the FRGs and their dynamics and what is the relative importance of each factor?

iv) Sub-goal 'evaluating':

7. What are the main building blocks to specify the meta-criteria of appropriateness and resilience into an assessment framework for FRGs, what kind of indicators could be derived from these building blocks and how can these indicators be measured?
8. What are the strengths, weaknesses, opportunities and threats of FRGs in the selected EU member states in terms of their appropriateness (legitimacy, efficiency and effectiveness) and resilience?
9. What are the main similarities and differences between the selected EU Member States in terms of development and performance of FRGs? What is the scientific and societal importance of these similarities and differences?

v) Main goal 'designing':

10. Which *design principles* can be derived from the analysis, explanation and evaluation of our cases?

1.2.3 Process followed in Work Package 2

This report, together with the "assessment framework and methodology report" is the result of a six-month process of contributions, discussions, and exchanges of views between all the members of the STAR FLOOD consortium. As such, it benefits from the contribution of all the researchers involved in the project (both senior and junior researchers, from different academic backgrounds and countries).

For several months, the construction of this approach has been an on-going process with important milestones. Apart from the everyday exchanges and discussions which took place between the different members of the consortium, several events have been organised: teleconferences (30/05/2013), meetings of young researchers (First and Second Academic Master Classes on the STAR-FLOOD project: 6 and 7th of June 2013, 2nd and 3rd of October 2013), Management Team Meeting (2nd of July 2013) and a plenary consortium meeting (3rd and 4th of October 2013). These events have been important moments for discussion, revision and clarification of both documents.

1.3 Headlines of the conceptual and methodological approach

STAR-FLOOD's conceptual and methodological approach should enable the empirical research to contribute to the identification of good practices for Flood Risk Governance in Europe as well as guidelines regarding their applicability in different contexts. To reach this goal, the assessment framework includes the following main elements:

- An **analytical framework**, specifying how stability and dynamics in Flood Risk Governance Arrangements can be analysed using the four dimensions of the Policy Arrangements Approach (actors; discourses; rules; resources);
- An **explanatory framework**, indicating which factors might explain the emergence, dominance and dynamics of current FRGAs;
- An **evaluation framework**, elaborating on how STAR-FLOOD's evaluation criteria of appropriateness and resilience can be translated into measurable indicators.

Put in other words, the three steps of analysing, explaining and evaluating should enable us to assess what types of Flood Risk Governance Arrangements are in place, why and to what effect? The main object of analysis is the National Flood Policies and Regulations domain (NFPR) in each of the six STAR-FLOOD consortium countries. This domain concerns all flood-relevant policies at the national level.

The framework and methodology report provides an elaboration on how the steps listed above will be researched, how the work in Work Package 3 paves the way for the work in the subsequent Work Packages (WP4 and WP5) and the role of STAR-FLOOD's glossary of key terms in the project. The current report's role is, as said before, to provide theoretical grounding of the framework.

1.4 Outline of the report and guide for the reader

The three subsequent chapters each correspond with one of the three subsequent steps in the assessment framework: analysing, explaining and evaluating Flood Risk Governance.

Chapter 2 presents an extract of two texts on the Policy Arrangements Approach, the theoretical backbone of STAR-FLOOD's analytical framework. The first text is an extract of a book chapter by Duncan Liefferink entitled 'The dynamics of policy arrangements: turning round the tetrahedron'. The second text is an extract from: Wiering M, Arts, B 2006, Discursive shifts in Dutch river management: 'deep' institutional change or adaptation strategy? In: *Hydrobiologia*, vol. 565, pp. 327-338. Both texts together provide an overview of the four dimensions of the Policy Arrangements Approach – Actors, Discourses, Rules, Resources, examples of how they can be operationalised, also with regard to water management issues, and some reflection on how the four dimensions can be brought together into a characterisation of Policy Arrangements as a whole.

Chapter 3 provides an overview of background theories related to STAR-FLOOD's explanatory framework. The chapter first discusses the main features of some prominent explanatory theories from the policy sciences literature, including the Multiple Streams Framework, Punctuated Equilibrium Theory, and the Advocacy Coalitions Framework as well as publications closely related to or critically reflecting upon the frameworks. The chapter then introduces the Multi-Level Perspective from Transitions Theory and discusses some syntheses between all aforementioned explanatory frameworks. Subsequently, some attention is devoted to a specific issue of relevance for efforts at explaining policy change, that is, the relationship between social structures and the purposeful actions of agents involved in Flood Risk Governance. The chapter goes on by discussing some of the literature on change agency and policy entrepreneurship. The last part of the chapter is devoted to reflections on the relationship between explanations by policy analysts and legal

scholars; the relevance of natural law and positive law studies for explaining stability and dynamics in Flood Risk Governance; and specific explanations for stability and dynamics *in* legal frameworks.

Chapter 4, finally, provides further background to the evaluation framework. The chapter starts off with a general reflection on evaluating governance, including an elaboration of previous experiences with evaluating governance. Subsequently, the literature on each one of STAR-FLOOD's evaluation criteria is discussed in some detail: resilience, appropriateness, legitimacy, efficiency and effectiveness. The chapter then reflects on how the sub-criteria of legitimacy, efficiency and effectiveness could be combined into the meta-criterion of appropriateness. The chapter concludes with a discussion on operational challenges for evaluation and a summary of the evaluation framework as a whole.

2 The Policy Arrangements Approach: introduction and basic ingredients for its operationalisation

Mark Wiering and Duncan Liefferink

2.1 Introduction

This chapter presents two texts introducing the policy arrangements approach (PAA) as well as a few basic ingredients to operationalise the PAA for the purpose of the STAR-FLOOD -project. The first text is an extract from a book chapter. It sets out the basic logic of the PAA, elaborating the four dimensions (actors/coalitions, resources/power, rules of the game and discourses) and the dynamic relationship between them. It also shows how a policy arrangement can be analysed starting from any of the four dimensions. The argument is illustrated throughout the text with the example of the policy arrangement around organic farming in the Netherlands.

The second text is an extract from an article applying the PAA to the case of river management in the Netherlands – a case which comes close to the substance matter of the STAR-FLOOD -project. It applies various elements of the first text in practice and elaborates on specific aspects which may be of particular relevance to the STAR-FLOOD -project. Notably, it subdivides the discourse dimension of the PAA into three different ‘layers’: ontological, normative and strategic.

Both texts have been published. The first text is an extract of a book chapter by Duncan Liefferink entitled ‘The dynamics of policy arrangements: turning round the tetrahedron’ in: Arts, B, and Leroy, P (eds.), *Institutional dynamics in environmental governance*, Springer, Dordrecht, pp. 45-68. The second text is an extract from: Wiering M, Arts, B 2006, Discursive shifts in Dutch river management: ‘deep’ institutional change or adaptation strategy? In: *Hydrobiologia*, vol. 565, pp. 327-338. Both texts have been reprinted with permission of the authors.

2.2 The concept of policy arrangements

2.2.1 Introduction

A policy arrangement has been defined as the temporary stabilisation of the content and organisation of a particular policy domain (Van Tatenhove *et al.* 2000, p.54). Daily interactions between policy actors are assumed to gradually develop into more or less stable patterns. These patterns may include the substantive delineation of the problem at stake and of possible solutions, but also the processes of give-and-take between the actors and the formal and informal rules according to which these processes take place. This process of stabilisation is usually referred to as institutionalisation. The structures thus formed in turn shape subsequent behaviour. Such structures are not fixed, however. Like language, structures are ‘used’ by people who are able to change their behaviour. Speakers of a given language may gradually adopt new grammatical or syntactical rules, or invent new expressions. They are not able, though, to change the entire language at once (cf. Giddens 1984).

The structure of a policy arrangement can be analysed along the following four dimensions:

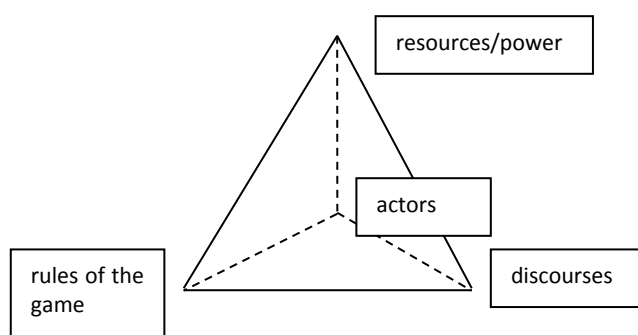
- the *actors* and their *coalitions* involved in the policy domain;

- the division of *resources* between these actors, leading to differences in *power* and influence, where power refers to the mobilisation and deployment of the available resources, and influence to who determines policy outcomes and how;
- the *rules of the game* currently in operation, in terms of formal procedures of decision making and implementation as well as informal rules and ‘routines’ of interaction; and
- the current policy *discourses*, where discourses entail the views and narratives of the actors involved (norms, values, definitions of problems and approaches to solutions).

The former three dimensions refer to the organisation of policy, whereas the latter refers to its content. The four dimensions listed above draw heavily upon, in particular, network theory of the 1990s (e.g. Marsh and Rhodes 1992), enriched with elements from discourse analysis (e.g. Hajer 1995; Dryzek 1997; see further: Van Tatenhove et al. 2000). In contrast with these approaches, however, the crucial point we want to make is that the four dimensions do not just sum up to define a policy arrangement. They are inextricably interwoven. This interrelatedness is essential for understanding a policy arrangement at a given point in time. In addition, it allows us to analyse the dynamics of a policy arrangement over time. The interrelatedness of the four dimensions can be symbolised by a tetrahedron, in which each of the corners represents one dimension (Figure 1).

The symbol of the tetrahedron visualises that any change in one of the dimensions may induce change in other dimensions. The appearance of new actors or a change in the composition of coalitions, for instance, may add new elements to the prevalent discourse or lead to another distribution of resources. Similarly, the introduction of extra resources (e.g. subsidies, knowledge, skills) or their withdrawal may attract new actors, exclude others or instigate new coalitions. A change in formal procedures, such as rules of participation or voting, may have similar effects. Finally, new ideas may enter the tetrahedron through the dimension of discourse. Examples are concepts like ‘public-private partnership’ or ‘sustainable development’. If successful, such concepts may mobilise new types of expertise or legitimacy (i.e. resources) or form the nucleus of new coalitions. As a consequence of the indissoluble interrelatedness of the four dimensions, repercussions across dimensions are likely to occur – even though they do not necessarily have to do so in each and every case. Therefore, the analysis of a policy arrangement should in principle address the entire tetrahedron.

Figure 2.1: The tetrahedron, symbolising the interconnectedness of the four dimensions of a policy arrangement



The argument of the interconnectedness of the four dimensions does of course not imply that policy arrangement are always harmonious, stable and internally consistent. Incongruence among the dimensions of an arrangement or ‘institutional voids’, e.g. the absence of shared rules (Hajer 2003), may in fact result in shorter or longer periods of instability and shock wise changes. Alternatively, as we will see in this as well as the next chapter, it may lead to a policy arrangement which hardly moves at all.

Policy arrangements, moreover, do not operate in a vacuum. They are part of society. This means, on the one hand, that changes in individual policy arrangements may add up to more structural trends, for instance an inclination to involve stakeholders such as business and non-governmental organisations (NGOs) in policy making or the increased use of voluntary agreements in environmental policy. On the other hand, changes in the broader social, cultural, political or economic context or in fact the physical environment may have an impact on, for instance, power relations between actors involved in specific policy processes. Thus, structural change in individual policy arrangements often originates in broader processes of (social, political etc) change. In the present paper, we will not go into these possible drivers of change (for this question, see Arts and Van Tatenhove, this volume). Instead, we will focus on the analysis of the dynamics at the 'micro-level', i.e. the question how the four dimensions within one policy arrangement interrelate and make the tetrahedron 'revolve'.

2.2.2 Analysing policy arrangements: a matter of perspective

The main message of the previous section was that the analysis of a policy arrangement only makes sense if it is comprehensive. That is, if it encompasses all four dimensions of policy arrangements distinguished above – actors/coalitions, resources/power, rules of the game, and discourses – as well as their mutual relations. In practice, the analysis may in fact start at any corner of the tetrahedron, as long as all corners and the connections between them are eventually covered. This is not to say, however, that the choice of a starting point is unimportant. On the contrary, different starting points imply the use of different conceptual and methodological tools. In a more practical sense, this is useful also in terms of delimiting the research task. But perhaps most importantly, different starting points shed different light on the policy arrangement at stake. Which phenomena one prefers to highlight depends on the research question underlying the analysis.

In the present section, we will elaborate our multi-perspective approach to the analysis of policy arrangements with the help of the same case as used in the previous chapter, the development of organic farming in the Netherlands. In the present chapter, however, we will scrutinise the internal dynamics of the arrangement rather than its structural ramifications. We will systematically subject the case to different types of research questions. Where relevant, we will distinguish between scholarly, theoretically informed research questions and more applied policy-oriented research questions. In addition, we will address some methodological implications. It will turn out that the dimension where we start provides the key to 'unlocking' the other dimensions. Thus, by departing from all four corners of the tetrahedron subsequently, we will watch the same policy game from four different angles. The empirical material for this exercise will be taken mainly from Arts *et al.* (2001) and Hofer (2000).

2.2.3 Actors/coalitions

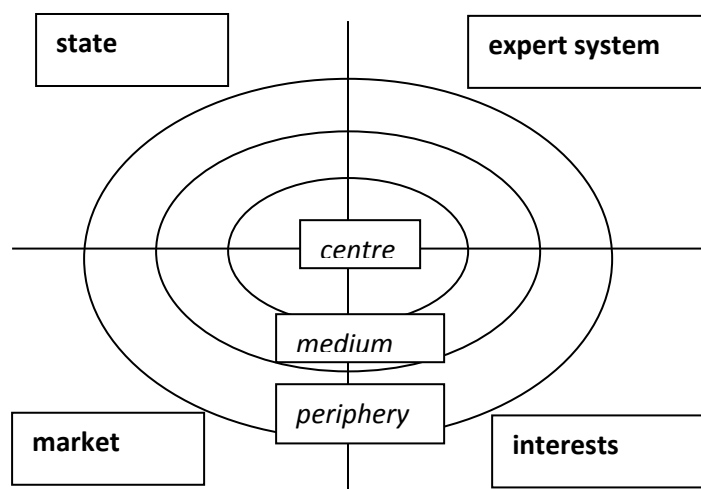
For research questions focusing on the positions and roles of actors in a given policy arrangement, it is most suitable to 'enter' the tetrahedron from the actor/coalitions corner. Questions of this type can on the one hand be theoretically inspired, for instance by macro-theories on changes in the relationship between state, market and civil society. To mention just a few examples: Does, as is often assumed, the nation-state retreat in favour of international and sub-national levels in the face of globalisation and European integration? And if so, how are policies made and implemented in this system of multi-level governance? Does late-modern society lead to the emergence of new coalitions and various kinds of sub-politics by private actors such as firms, social movements and even individuals (Beck 1994)? On the other hand, research into the positions and roles of specific actors can be more practice-oriented. It can help to assess the options available to actors such as government departments, firms or NGOs to improve their stakes in the policy process. For this reason, in fact, a large part of privately funded, strategic policy research implies an actor perspective.

In addition to this, an actor perspective is the most tangible way to get to an overview of the policy arrangement around a given issue. It also comes closest to how policy actors themselves view their own situation. In day-to-day practice, people deal with other people. It is only through them that our other analytical categories, i.e. resources/power, rules and discourses, materialise. They often do so in the form of 'allies' and 'enemies'. Not surprisingly, therefore, many policy studies start by determining *who* is involved in the policy area under consideration. They then go on to analyse the power relations between these actors and the institutional context in which they operate (cf., for example, the Dutch policy network approach, as represented by Glasbergen 1989). This is also how we will proceed in the present section.

Analysing a policy arrangement from an actor perspective starts by identifying the relevant actors and their influence in the policy process. This can be done through the study of policy documents, but also 'in the field'. In doing so, it is useful to distinguish between central and more peripheral actors and to cluster actors that fulfil similar roles in the arrangement.

In the organic farming case, we asked a limited number of experienced, well-informed practitioners to position all relevant players in the field in a figure consisting of three concentric circles, indicating influence, and four spheres or segments, indicating the different roles performed by the actors (see Figure 2; of course the number and character of clusters can be different for other cases). The outcomes were validated with other practitioners. Some minor adjustments were then made, but generally speaking there turned out to be a remarkably high consensus on the positions of the various actors. The original version of Figure 2 (see Arts et al. 2000) contains several dozens of actors. As the figure only serves as an illustration in the present context and in order not to make things unnecessarily complicated, these have been left out here. A figure like this, when properly filled in, provides a good basis for looking at the power relations between the actors involved in the arrangement. When turning to the dimension of power and resources, in other words, our initial analytical starting point in the actor dimension implies a focus on the relative power of actors *vis-à-vis* each other, i.e. 'relational power'.

Figure 2.2: Example of a 'map' of actors and their relative positions in a policy arrangement



Source: Arts et al. 2000 (adapted).

With regard to organic farming in the Netherlands, it is striking that the core of the figure is almost entirely dominated by actors normally associated with traditional agriculture. In the Netherlands, more than in almost any other country, this association with traditional agriculture implies a focus on resource intensity and high productivity. This goes in particular for the 'state' and 'expert system' segments. In the 'market' segment, some organic trade firms and associations appear, but in terms of size and resources they are in fact tiny in comparison with the regular agricultural and food chains operating in the same segment. Only in the 'interests' segment, a truly 'organic' association dominates, i.e. Biologica, the principal umbrella organisation for organic interests. Although this might be taken as an indication of Biologica's exceptionally influential position, this is probably not the case. It rather suggests that Biologica is in fact fairly isolated in articulating the interests of organic farming in the Netherlands. Other 'organic' players, on the one hand, have far more marginal positions in the arrangement than Biologica. Traditional agricultural interest groups, on the other hand, are not eager to promote 'organic' interest. At the same time, they hardly need to invest much of their resources in defending their own 'traditional' interests, as these are sufficiently taken care of by the majority of core actors anyway. This picture reflects, in short, that organic farming does not in fact present a serious challenge to traditional agriculture in the Netherlands.

Having identified the most important actors and power relations in the arrangement, we can add another dimension to the analysis by grouping actors according to the views they have about the policy domain at stake. Such exercise is reminiscent of the work by Sabatier and Jenkins-Smith (1993) on advocacy coalitions. It should be noted, however, that advocacy coalitions are based on shared beliefs (i.e. an attribute of *actors*) which cannot be fully equated with discourses (i.e. 'storylines' or 'narratives' prevailing in a given *policy arrangement*). In our case, two discourses on the future development of organic farming in the Netherlands can be distinguished. A radical one is based on the idea of sustained *competition* between traditional and organic farming. According to this view, organic farming should maintain its own expert system and its own marketing channels, and – as a long term perspective – eventually push traditional products from the market. Not surprisingly, this discourse is not supported by the traditional agricultural community. But perhaps more surprisingly, it is not supported by 'mainstream' organic groups either. They rather cherish a more pragmatic view, by which the existing gap between traditional and organic farming gradually comes to be closed. In this perspective of *assimilation*, organic farming makes increasing use of the traditional agricultural research infrastructure and traditional marketing channels, e.g. regular supermarkets instead of small, specialised shops. This should eventually lead to a reform of traditional agriculture 'from inside'. In a moderate form (i.e. without the explicit long term perspective), this view finds support among many traditional actors too, including the organisations of traditional farmers. Pressed by environmental problems and a steady stream of animal diseases associated with traditional agriculture, they have gradually come to recognise the need of developing more environmentally friendly forms of farming. In this context, organic farming is seen as an option for the future at least worth considering. Commercial considerations may play a role here as well. Organic products now constitute a small but potentially growing niche market. Regular trade and retail are in principle ready to step in, as did supermarkets for instance in Austria and Denmark (cf. Hofer 2000).

As far as the rules of the games are concerned, finally, an 'actor-based' analysis implies a focus on the rules governing the interaction between the actors involved. In the organic 'game' in the Netherlands, those rules are almost fully controlled by the state. Since the 1991 EU Regulation on organic farming (2092/91), for instance, the leading organic food label EKO is controlled by the state (see below). Most major (and traditional) research institutes in the agricultural field, moreover, have recently been privatised, but still maintain close links with the state. In the present situation, finally, the conversion of individual farmers from (highly intensive) regular to organic agriculture is heavily

dependent upon state subsidies. The relatively low conversion rate in the Netherlands may in fact be partially explained by the modest level of those subsidies.

The above 'actor-based' analysis of the policy arrangement around organic farming in the Netherlands makes clear that resources as well as rules are strongly focused around the state and, consequently, around traditional agriculture. As we have seen, the world of traditional farming is not radically against organic farming, but it is not particularly active in promoting it either. Organic food is essentially seen as a small and 'harmless' niche market, or – more positively – as one possible way out of the protracted 'crisis' of traditional agriculture. At the same time, however, organic farmers lack resources to have a strong position of their own. This may explain, among other things, a good deal of pragmatism and the prevalence of a discourse of assimilation among major organic groups, notably Biologica.

2.2.4 Resources/power

The analysis of resource dependencies and power relations is central to several theoretical approaches in political science. One can think of traditional (neo-) Marxism or more recent approaches such as studies of multi-level governance. In the latter case, research questions focus on (shifts of) resources and power between, for instance, the nation state, sub-national levels and the EU. Within these approaches, to be sure, one can choose to concentrate on the resources and power of one particular (type of) actor, e.g. NGOs or the national environment ministry. Questions of shifting resources and power are addressed even more directly in practice-oriented research evaluating the impact of policy interventions. Such interventions often amount to the introduction of certain resources into or the withdrawal of resources from the policy arrangement, for example by way of subsidies, taxes or the production and dissemination of particular expertise. Evaluation can be done either *ex ante*, i.e. in advance in order to estimate the impact of a planned policy intervention, or *ex post*, i.e. afterwards to measure the impact actually realised, perhaps as a basis for revising or fine-tuning the policy.

The empirical analysis to be undertaken in this perspective comes close to what is propagated by the British school of policy network analysis (e.g. Rhodes 1986; Marsh and Rhodes 1992). The core idea of this approach is that actors around a given policy issue are to different degrees dependent upon each other for resources, e.g. money, information, or political legitimacy. In this way, linking the dimensions of resources/power and actors, we are again (as in the previous section) identifying power relations between actors, i.e. 'relational power'. By mapping these resource dependencies, moreover, it will become clear that certain actors are, as it were, driven into each other's arms because they share control over important resources. Thus, we can identify different 'resource coalitions'. It should be noted that in one policy arrangement, money may be the central stake, while in another arrangement, the exchange of, for instance, knowledge and expertise may be crucial. In the policy game, moreover, resources can be seen as 'weapons', i.e. actors attempt to determine outcomes with the help of resources, but at the same time as 'prizes', i.e. during the process actors attempt to improve their situation by changing the distribution of resources to their advantage (Rhodes 1986, p.19-20). In this context, rules play an ambiguous role. On the one hand, they can be used strategically, i.e. as legal resources, in the policy game. On the other hand, they are not, as for instance money or personnel, exclusively controlled by certain actors. Instead, they are part of the actors' mutual knowledge (Giddens 1984, p.17-8). Nevertheless, they can be changed by actors who have the power to do so. The ambiguous nature of rules in this context is further stressed by the fact that such 'regulatory power' is often based again upon formal or informal rules. The same actually goes for discourses. Discourses can be used as 'weapons' for gaining, for instance, political legitimacy, but without being under any actor's exclusive control, some actors may be able to change the content of the narratives prevailing in the arrangement, or even to introduce wholly new ones. Such 'discursive power' is not seldom based upon political legitimacy.

Turning to our case of organic farming in the Netherlands, the position and role of the organic umbrella organisation Biologica can very well be understood from the perspective of resource dependencies. In the first place, Biologica is itself a sort of resource coalition. It brings together a number of smaller organisations representing different strands in organic agriculture (e.g. biodynamic, 'regular' organic, etc). While discursive differences persist within Biologica, the organisation's main aim is to pool resources (money, personnel, legitimacy) so as to increase the political impact of the organic sector. As we have seen, however, Biologica's position vis-à-vis traditional agriculture and traditional farmers organisations is weak. Therefore, Biologica is particularly active in forging strategic alliances with a large variety of other actors, ranging from environmental and consumer organisations and the Dutch party *Green Left*, to banks and supermarkets. The main resource paid into these alliances by Biologica, the single 'authorised' voice of the Dutch organic sector, is clearly political legitimacy.

At a more practical level, a resource based approach may help us to analyse the impact of specific policy interventions, e.g. the impact of state subsidies for conversion from regular to organic farming. The level of these subsidies in the Netherlands is comparatively low. According to Biologica, it is in fact far too low to compensate for the loss of income during the conversion period. The low conversion rates and the low overall percentage of organic farming (2,2% of all agricultural land in the Netherlands in 2003; Biologica 2004) suggest that they have a point here. However, the development of the Dutch home market for organic food has been even slower (in 2003: 1,6% of total food consumption in the Netherlands; Biologica 2004). More than in other countries, both consumers and retailers have been very reluctant to 'go organic'. Who could break this stalemate? The state is basically unwilling to intervene in the market for organic food. This is legitimated with reference to a production and market oriented discourse which has also long dominated Dutch agricultural policy at large (see for instance LNV 2000). This basically liberal discourse has been further strengthened with the present central/right wing government. According to it, the role of the state should not be more than giving an initial push in the beginning of the production chain (i.e. the conversion subsidies), after which the market is supposed to do its work. The organic sector itself, however, lacks the resources for convincing either the general public or the supermarkets, or both, of the merits of organic products.

2.2.5 Rules

As a third 'entrance' into a policy arrangement there is the dimension of rules. Rules are the mutually agreed formal procedures and informal routines of interaction within institutions. Formulated this way, rules have a particularly strong connection with the actor dimension of the arrangement. When rules are connected with the dimension of resources and power, we are back again at the notion of 'regulatory power', introduced above. Turning to the dimension of discourses, finally, we may try to identify the discourses *underlying* the rules of interaction prevailing in the network. As will be further elaborated below, these discourses mainly deal with general ideas about governance, i.e. the relationship between state, market and civil society.

Entering the tetrahedron via the rules dimension is a suitable strategy for studying the influence of institutional change on particular policy areas. From a more theoretical starting point, one could think of the influence of evolving European Union rules on national institutions, often referred to as Europeanization. But one does not have to focus on change to study the impact of institutions. Comparative research allows for institutional analysis at a more general level, addressing not only change in particular institutional constellations, but also by singling out the impact of their more stable, enduring features, e.g. electoral systems, or the division of competences between national and sub-national levels in federal systems. This is typically done by studying one policy issue or policy field in different countries. This perspective may remind one of neo-institutional analysis (e.g. March and Olsen 1989). Furthermore, from a more practical point of view, starting from the rules

dimension can of course be helpful for evaluating (*ex ante* or *ex post*) the effect of the introduction of new rules or procedures on other dimensions of the policy arrangement.

A nice example in the field of organic farming is the differential impact of rules on the labelling of organic food products in the Netherlands and Denmark (Hofer 2000). The Dutch EKO label was developed in the 1980s as a private initiative, while the state deliberately held aloof. Although the name of the EKO label became relatively well-known among the public, its market share remained below 1%. The Danish label, the 'red Ø', was introduced in 1989. It was owned and controlled by the state. Organic farmers were reluctant at first, but changed their minds when they saw that the 'red Ø' was implemented forcefully and that, moreover, the strong state involvement could be used as an argument for the quality and credibility of the label *vis-à-vis* consumers and retailers. In 1996, the market share of organic food in Denmark amounted to 5%. In the course of the 1990s, however, EU Regulation 2092/91 on organic farming had to be implemented. Among other things, it required state control of organic food labels. For Denmark nothing really changed, but in the Netherlands the formerly private EKO label was put under state supervision. Contrary to the experience with the Danish state-owned 'red Ø', however, this did not improve the effectiveness of the Dutch label? This was due first to the fact that the state still hardly associated itself with the label. It limited itself to its technical control task, but did not put any effort in strengthening the label's position on the market. This was of course perfectly in line with the 'reticent state' philosophy referred to above, according to which the market should function by itself. At the same time, it reflected a considerable degree of distrust between the state and the organic sector. In the eyes of Dutch organic farmers, strong state involvement would in fact hardly increase the credibility of the label. This situation was aggravated by the extension to food products, also in the mid-1990s, of the state-owned *Milieukeur* label, which had existed for some years but initially covered non-food products only. The standards of this label were less strict than those of the EKO label. Although the performance of the agri-environmental *Milieukeur* on the market had turned out quite poor, the competition between the two labels increased confusion among consumers and thus posed an at least indirect threat to the EKO label.

The example illustrates that the impact of different or changing rules cannot be studied in isolation. We saw that the question of state control as such did not tell anything about the functioning of the Dutch and Danish organic food labels. Why things worked out so differently in the two countries could only be understood against the background of the other dimensions of the respective policy arrangements, in particular the distribution of public resources and the prevalent discourse around the role of the state.

2.2.6 Discourses

The final dimension of the tetrahedron to be discussed here is discourses. It is important to note from the outset that discourses are relevant at two different levels. The first level refers to general ideas about the organisation of society, particularly the relationship between state, market and civil society, i.e. about the preferred mode of governance. Such ideas clearly exceed specific policy issues or sectors. Through the views of the actors involved, however, they may have an impact on specific policy arrangements. As pointed out above, for instance, they may have important implications for the rules of interaction in the arrangement. The second level concerns ideas about the concrete policy problem at stake, e.g. about the character of the problem, its causes and possible solutions. Discourses at this level imply substantive strategic positions of actors in the arrangement. Groups of actors around one particular discourse, or discourse coalitions, may be discerned at both levels, i.e. not only at the second, issue-specific level, but also at the governance level. Moreover, as we will see below, considerable incongruences between the two levels may exist. If and how such incongruences can be reduced depends not least on the distribution of discursive power in the arrangement.

From a theoretical point of view, entering the tetrahedron through the dimension of discourses may clearly be interesting to study the empirical effects of political modernisation, i.e. changing ideas about governance (see also Hajer 1995). More practically, the role of changing problem perceptions, new scientific insights or public information campaigns may be scrutinised.

As we have seen, at the level of the policy arrangement around organic farming as such, a discourse of assimilation, rather than sustained competition, between organic and regular farming prevails in the Netherlands. However, due to the high cost of conversion, the number of organic farmers lags behind. Moreover, retailers and consumers persist in their wait-and-see behaviour, and research institutes continue to focus on traditional, intensive farming methods. The organic sector itself is not sufficiently powerful to get things moving, while regular farmers and farmers' organisations are only beginning to develop a friendlier attitude towards organic agriculture. Under the present circumstances, in other words, accelerating the process of assimilation would hardly be conceivable without close collaboration with the state. At this point, however, the discourse of assimilation collides with a second discourse, referred to earlier, that is dominant at the governance level. It implicates that the state should limit itself to facilitating the development of the market and refrain from strong and focused interventions. Regardless of the political and ideological arguments involved here, this is important in view of the state's central position in the policy arrangement around organic farming. As discussed above, the state controls a number of central resources in the field. These notably include financial ones but also expertise. For many new initiatives in the field, therefore, the state is a necessary partner. Thus, the incongruence between the assimilation discourse on the one hand and the liberal, 'reticent state' discourse on the other plays an important role in explaining the slow development of the sector in comparison with most surrounding countries (cf. Biologica 2004, p.7).

In more theoretical terms, our analysis shows that substantive discourses at the issue specific level on the one hand and governance discourses at the more general level on the other, are potentially incompatible. If this is the case, actors within the arrangement may be driven by conflicting ideas. This may, as it were, suffocate new developments and paralyse the policy arrangement. The organic farming case also suggests that such immobility may last for a considerable period.

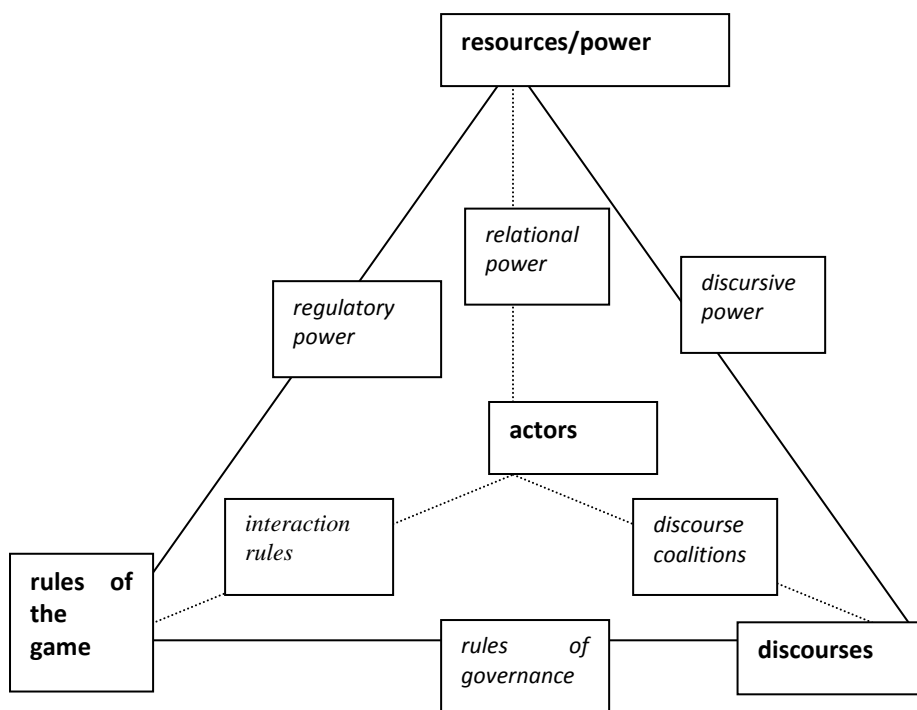
2.2.7 Summing up

In this section we have attempted to show how a policy arrangement can be analysed taking each of the four respective corners of the tetrahedron as starting point. By highlighting different features of the arrangement's dimensions, each perspective emphasised different aspects of the arrangement as a whole.

Analytically, the section can be summarised with the help of an annotated version of the tetrahedron (Figure 3). Looking at the rules dimension from the vantage point of actors, for instance, one focuses on the rules of interaction between the actors in the arrangement. Starting from actors and looking at resources, the focus is on power relations between actors, or relational power. Far from being meant as a blueprint for empirical research, the figure draws the attention to the analytical possibilities implied in policy arrangements theory.

The case of organic farming was used to give some flesh and blood to the four analytical angles. This is an unusual exercise, to be sure, as normally one would prefer to choose just one perspective, dependent on the research questions at stake. And even then, a full empirical analysis would require at least several pages. Our case, therefore, has been no more than an illustration to give the reader an idea of the empirical implications of our argument, which remains essentially analytical.

Figure 2.3: The tetrahedron: analytical perspectives



2.3 Discursive shifts in Dutch river management: ‘deep’ institutional change or adaptation strategy?

2.3.1 Abstract

This section argues that a discursive shift is taking place in Dutch water policy, from ‘a battle against water’ to ‘living with water’ or ‘accommodating water’. Yet we ask ourselves whether this shift is just an adaptation strategy of the existing elite group of water managers, who pay lip-service to new management approaches in order to maintain their vested interests, as some authors claim, or whether it implies ‘deep’ institutional change, e.g. in terms of the emergence of new water institutions, power relations and procedures. While investigating this question, we make use of the ‘policy arrangement approach’, which pays attention to institutional and discursive aspects of policy making alike. Our conclusion is that we are currently observing institutional changes beyond ‘policy talk’, particularly in terms of new legislation and procedures. However, it is too early to speak of ‘deep’ institutional change in Dutch water management, because the former water institutions are still maintaining their power positions, despite the availability of additional resources for policy and research as well as the emergence of several new modes of governance.

2.3.2 Discourse

In our view, a *discourse* refers to a set of ideas, concepts, buzzwords and stories which combined give meaning to a certain phenomenon in the real world (Hajer 1995). An example is the sustainability discourse, which brings together notions such as economic, ecological and social sustainability, sustainable development, a belief in the possibility to integrate economy and ecology, examples of win-win situations, etc. This language gives meaning to a world characterized by poverty and ecological degradation, but also to a world which has the potential to become sustainable after all. Theoretically, a discourse consists of three ‘layers’: *ontological*, *normative* and *strategic* (Therborn 1982). Discourses can be essentially *ontological* in nature, related to questions such as: How do we see reality? How do we define problems? What do we think is taking place? Can we be

certain about our risk management strategies? Here we deal with the ‘paradigms’ or ‘world views’ of policy actors. Whether these (fundamentally) change over time, or not, is our first discursive ‘indicator’. At the same time, discursive space is filled with *normative* expressions, concerning the values at stake and the goals that are set. Here we deal with the ‘utopias’ or ‘ideals’ of policy actors. Whether these (fundamentally) change over time, or not, is our second discursive ‘indicator’. The third layer of discourse consists of the *route* or ‘road map’ from what we see as ‘real’ to what we conceive of as ‘desirable’, from problem to solution. Here we deal with the ‘policy programmes’ of policy actors. Whether these (fundamentally) change over time, or not, is our third discursive ‘indicator’.

2.3.3 Rules

The next dimension, *rules*, consists of ‘legislation’, ‘procedures’ and ‘political culture’ (Giddens 1984; Rittberger 1993). Legislation refers to the formalization and transposition of policy discourses into binding law. Therefore, an important aspect of ‘deep’ institutional change is the extent to which changes in discourse are reflected in changes in legislation. Nevertheless, rules are not only ‘substantive’ in nature, but also ‘organizational’ (Giddens 1984). This latter aspect refers to *procedures*, to how political participation and decision-making processes are codified in ‘the rules of the game’. One might wonder whether discursive shifts in a policy domain also imply that there are new rules to the game, e.g. in terms of the participation of new actors in decision-making. For example, does a cross-border river basin approach (discourse!) indeed lead to the formal participation (rules!) of German policy actors in Dutch water management? Besides the distinction between substantive and organizational rules, one can distinguish between formal and informal ones. The former refers to legislation and procedures, previously dealt with, whereas the latter refers to ‘political culture’. For example, the Dutch ‘polder model’ generally causes other types of policy processes to occur rather than the German ‘formal-legalistic model’ (Haverland 1999). Such national ‘policy cultures’ colour the way in which policies are shaped. Nevertheless, it should be noted that these informal rules do not change that easily.

2.3.4 Actors and coalitions

The third policy arrangement dimension of *policy actors* is analysed on the basis of ‘actor constellation’, ‘interaction patterns’ and ‘coalitions & oppositions’ (based on: Godfroij and Nelissen 1993; Kickert et al. 1997; Marsh and Rhodes 1992). The first ‘indicator’ relates to the set of (key) policy actors in a given policy domain (such as water management). The question is “Who is involved in agenda-setting, decision-making and policy implementation, both formally as well as informally?” And do we see changes over time within this constellation of actors? Secondly, we can perhaps observe changes in the way these players interact, quantitatively and qualitatively. For example: does interaction increase or decrease? Do we see more co-operation or, in contrast, more conflict? As a consequence of these changing interaction patterns, we might find the emergence of new coalitions and oppositions, making this our third ‘indicator’. However, it should be noted that a change of interaction patterns does not *necessarily* imply that new coalitions or oppositions have been established. This is just a *possibility* and, for that reason, we distinguish between this second and third ‘indicator’.

2.3.5 Resources

Finally, the dimension of *resources* is elaborated upon in the ‘indicators’ resource constellation, power relations and political influence (Arts 1998; Huberts and Kleinnijenhuis 1994). The first relates to assets which policy actors have or can mobilize on the basis of which they can exercise power, e.g. authority, money, knowledge or technology. The relevance of these resources and their usefulness may vary depending on the setting and the time span. Generally, these assets are not equally divided among policy actors, which leads to a situation in which not all of the actors share similar capacities to achieve (political) outcomes. Here, unequal power relations between policy actors exist, although

one should acknowledge that these relations are dynamic (in time and space). A further question to be asked is whether policy actors do indeed make use of their capacities to achieve the outcomes they wanted. In other words, power should be ‘operationalized’ into influence. Although intrinsically related, there is no one-to-one relationship between power and influence, as policy actors may decide not to make use of their resources and/or fail to achieve the outcomes they wanted. Although it should be noted that power (in terms of resources and capacities) is relatively easy to assess, whereas it is very hard to measure political influence. Therefore, in the context of this article, we will limit our analysis to resources and power relations.

Table 2.1: Operationalization of the Policy Arrangement Approach (PAA).

Concept	Aspects	Dimensions	Change indicators
Policy arrangement	Substance Organization	Discourses Rules Actors Resources	<u>Change in:</u> *Paradigms *Utopias *Policy programmes *Legislation *Procedures *Political culture *Actor constellation *Interaction patterns *Coalitions & oppositions *Resource constellation *Power relations *Political influence

¹Rules can be both substantive and organisational in nature and hence they pertain both to the substance and the organisation of policy arrangements

2.3.6 Discursive shifts in river management

The traditional discourse in Dutch water management reflects the history of the Netherlands: fighting against the sea, storms and frequent flooding, losing land, building dikes, conquering land from the sea, embanking and cultivating it. Similarly, the Dutch streamlined their rivers, minimised the river basins, closed creeks and small streams and replaced them with canals. Water was mostly viewed as a ‘threat’ and had to be regulated and controlled. (Van de Ven 2004; Van Steen and Pellenbarg 2004) This ‘battle against the water’ discourse led to a river management that was, up until very recently, focused on building dikes in order to keep peoples feet dry (Wiering and Driessen 2001). This perspective was widely spread amongst the civil engineers and (other) policy makers of the Ministry for Transport, Public Works and Water Management and its Directorate-General, as well as the regional water boards. From the beginning of the 1960s up until the mid-1980s river management was politically overshadowed by the coastal works (the famous Delta works) and subsequently upon finishing these major projects, river flooding management was discussed, but had low priority on the Dutch political agenda. A long lasting discussion on water safety norms for river flooding was characterised by Van Eeten (1999) as a ‘dialogue of the deaf’. Dike enhancement was even more slowed down by protest and litigation from river landscape protectors and environmental agencies, who expressed a ‘counter-movement’ distrust in Dutch water authorities as well as a Not In My Backyard (NIMBY)-effect among civilians. Yet, in general and among the public at large, feelings of trust in Dutch water management could be sensed.

From the mid-1980s river management gradually incorporated the upcoming issues of water quality, environmental concern and nature conservation, and evolved towards the concept of integrated water management. In 1985, a memorandum of the Ministry (Dealing with Water; *Omgaan met*

Water) promoted a system-oriented and integrated view on water management, thereby initiating a fundamental discussion during the following years. Gradually appeared a new, 'system ecology' discourse on river management, in addition to the – still hegemonic – 'battle against water' discourse of the hydraulic engineering water manager (Van Hemert 1999; Disco 2000).

The relatively low priority placed on the (river-) flooding management in the Netherlands changed rapidly in 1995. In January of that year, the Dutch faced a near national disaster, as the water rose to extreme levels in the country's major rivers – the river Rhine (and its branches, the Waal, the Lower Rhine and the IJssel) and the river Meuse. Almost 250,000 people, and a large number of livestock belonging to farmers, had to be evacuated. The authorities feared that the dikes would not hold. The economic damage and evacuation costs were great. In the end, the dikes along the Rhine and its branches did hold, but only just. Large stretches of the river Meuse have only quays and natural embankments, and the people living in the southern part of the Meuse floodplain suffered the most material damage. In 1995 the near flood disaster can be considered as a real *shock event* felt in Dutch society.

Initially, this *strengthened the traditional discourse* relation to the division of water and land use. In terms of the different discursive elements we mentioned earlier, we could say that *ontologically*, water proved to be a threat to society (once again) and that there was a strong feeling of urgency in regard to 'defending ourselves against water'. According to tradition, the Dutch named the policy program 'Major Rivers Delta Plan' and 'Major Rivers Delta Act' as a blatant reference to the world famous 'Delta Works' which were constructed as a defence against the North Sea. Discursively, this was an important step. The *normative* discursive space was dominated by 'direct safety first' and the *strategy* was focussed on strengthening the dikes. Accompanied by a strong sense of urgency and by new legislation, the Dutch diking operation between 1996 and 2000 was, in itself, an example of successful project management (Wiering and Driessen 2001).

But this was only a short-term response to the shock event and the near-flooding disaster had also a strong impact on the ideas and plans for *future* water management. Policy makers and scientists began to realise that the former policy of dividing water and land - and marking borderlines with dikes - was not sufficient in order to meet the goals of safety and the reduction of risks in the long run. Some even used the term 'control paradox' (Rommelzwaal and Vroon 2000; Wiering and Immink 2003): by building and strengthening dikes an idea of safety is created, giving way to more social and economic activities behind the dikes. This, in turn, could lead to an increase in the detrimental social effects and the economic damage suffered when occasional flooding does happen and to an eventual increase in the feelings of *insecurity*. Gradually water management seemed to be moving away from merely building higher dikes towards adjusting and extending the flood plains and giving 'room to the river' (Smits et al. 2000) Van Stokkom et al. 1995). After the first policy guidelines in this direction were established, more radical policy plans were published. We could witness a discursive turn towards a new discourse of 'accommodating water'.

This idea of 'accommodating water' was not new. Environmental scientists, some of the hydrological engineers and even planning agencies, already described ways of water management that took natural water systems into account for many years, especially since the 'Dealing with Water' Memorandum of 1985. Disco (2002) stresses the growing importance of the ecological conceptualisation of water systems in the domain of central water agencies and institutions. He called this the 'ecological turn' of Dutch water management in the 1980s and 1990s, an evolutionary development that is also visible in the different planning reports on the national 'water household'. How, then, can we judge these processes by way of our discourse-indicators, changes in ontological, normative and strategic discourse? We predominantly witness changes in policy aims and goals (normative discourse) and route-changes (strategic discourse): 'room for the river', 'space for water'

and a 'good water quality' are the aims of the water managers. These are to be reached by new strategies: a more system-oriented approach towards water issues, implementing regional views on river basin management, new tools for integrated water management, etc. This is also promoted by communicating with society, by raising the public awareness of water related policy problems and by visualising water as a possibility, and not merely as a threat (Smits et al. 2000).

Interpreting the ontological indicator of discourse is, however, more problematic. We see the traditional 'hydraulic engineering' -perspective of the old 'battle against water' -discourse eroding, but the new discourse 'accommodating water' is to be found somewhere in between the old 'battle against water' and the more radical 'system ecology' discourses. Moreover, concepts such as 'room for the river' and 'space for water' are open to interpretation and can inhabit both natural ecosystem-based or more traditional hydraulic 'room for the engineer'-based views of water systems and similar solutions. We can conclude, though, that the traditional fixation on dike enhancement in river basins has made way for a variety of options to combine flooding management with land use and nature conservation (Van Stokkom et al. 2005).

After having given giving, in general terms, the different discourse-elements that involve fighting over hegemony in water management, we can now turn to the other aspects of the policy arrangement. We will focus on a few major processes of change in the three remaining institutional dimensions: rules of the game, power /resources and actors /coalitions. We will discuss these in more detail, in order to give insight into the nature of these changes, and to see if these discursive shifts are actually being institutionalised in the current water management.

2.3.7 Changing rules

The first organisational dimension in which changes have become visible is the 'rules of the game'-dimension. We shall first, very briefly, characterise this dimension. Present-day water management has differentiated into a set of rules, which have been laid down in, for example, the Water Management Act, the Groundwater Act, the Embankment Act, the Pollution of Surface Waters Act and the Pollution of Sea Water Act. Besides these laws, there are formalized or informal rules concerning the jurisdiction and competences of the water policy agents. The water legislation is looked upon as rather complex and difficult for outsiders to penetrate. Because of its complexity and fragmentation, the relevant authorities often rely on informal agreements concerning the division of tasks. Thus, when it comes down to new rules in general, one of the first questions to answer is whether the Dutch water legislation is on the brink of being redesigned.

In 2002 the Vice-Minister of Transport, Public Works and Water Management asked a special Committee to give its advice whether the legal design had to be revised and the course of possible institutional changes. The Committee argued that there were several reasons for changing the legislation. First of all, there is a lack of internal coherence and transparency in the complex set of rules regarding water management. Secondly, there is an indistinct relationship between the rules on general water policy (e.g. 'water household' planning, norms, standards, general procedures) and the specific rules of water management in practise, i.e. dike enhancement and maintenance, dams and embankments. The Committee concluded that a new, Integrated Water Management Act should be created, which would include most existing water legislation, as mentioned above, but would exclude the rules concerning 'Water Chain Management' and the (constitutional) legislation regarding competencies of organisations. It was only quite recently that the Dutch government confirmed that it is indeed necessary to redesign and create such an integrated Water Management Act (Memorandum 2004). Some of the arguments refer to the overall policy intended to streamline rules and to reduce rule density. More importantly, the arguments (of both the Committee and the Cabinet) are connected to the following two policy developments: 1) to anticipate the implementation of the European Water Framework Directive (WFD) and 2) to (better) embed the

central concept of integrated water management (referred to as 'water system management') - and to improve its instrumentation. In other words, the proposed legislative changes are thought to be an important step in switching from a sector-based water management to an integrated 'water system management'. These changes are stimulated by obligations, rules and concepts stemming from the European policy arena (WFD).

Change in procedures

Another sign of institutional change can be found in the instrumentation of water management in relation to other policy domains. An exponent of new procedures is the development of the process instrument of the so-called "Water Test". This procedure is to be considered as a form of 'water impact assessment' (partly resembling the environmental impact assessment) as a result of which water management will change its interrelations with spatial planning in the Netherlands. According to the Water Management in the 21st Century Advisory Committee - in the following referred to as the WB21-Committee - the Cabinet paid insufficient attention to safety and water-related problems in the past. As a result, a great deal of space was gradually reclaimed from the water management system. "New spatial planning decisions may not exacerbate the challenges to safety and leave water-related problems unnoticed" (Ministry of Public Works, Transport and Water Management 2000). The water test explicitly addresses water-related aspects in all relevant spatial plans and must result in a separate section in the explanatory policy document. It considers both water quantity issues (impacts on retention and storage capacity, risks of flooding, drought, and groundwater level) and water quality issues (surface- and groundwater pollution, water sanitation, biodiversity). It must be applied at an early stage in the spatial planning process and water/related problems cannot be passed from one watershed to another. If decisions are taken that have negative effects on the water system, adequate measures must be identified to compensate or mitigate these effects. Seen as a new procedure, the water test is expected to influence both the position of water managers and the responsibilities of spatial planners. Water managers will gradually shift from a 're-active' attitude in the realm of spatial planning to a 'pro-active' position, in which water managers are supposed to be involved in an early phase of policy formation of spatial plans (location, impact on water systems, etc.). Spatial planners are expected to take water aspects into account when making decisions concerning spatial plans.

Change in political culture

By using the term 'political culture' we point at the general patterns of the (mainly) informal and implicit rules of the game which can be associated with certain policy domains or which can characterise national politics. The political culture of water management can be typified by three features, which are momentarily in a state of flux. First of all, the governmental authorities are the ultimate locus of authoritative power in water management. Both market and civil society are relatively weak regulation mechanisms in this field. A centralised water management planning system exists and water agencies exhibit a rather hierarchical organisational culture. The policy arrangement involved here can thus be called 'state-oriented' or 'etatist' (cf. Van Tatenhove et al. 2000). However, this situation was not problematic during the past decades. Water quality and water safety have always been considered to be important public goods in Dutch society; goods which should be provided for by the state.

Secondly, because of the past history of strong sector-based politics, the water managers are used to operating in a rather autonomous and *isolated* policy field. There has hardly been any public support or protest from social groups, citizens or the business community, except, perhaps, for the traditional strong participation of farmers on Dutch water boards (Wiering and Immink 2003). Thirdly, water management is *technocratic* in nature. This is the result of the relatively closed policy domain in combination with specific functional governmental tasks and a specific epistemic community focusing on 'hydraulic engineering'.

In the wake of near-flooding disasters, excessive rain fall, drought problems and expected climate change, the political and societal attention paid to water problems is increasing. Because of this, the etatist, closed and technocratic features of the Dutch water policy arrangement have been contested in recent years. A 'societalisation' of water management has emerged, including an aspiration of the domain itself to act less hierarchically and to decentralise decisions (Van Leussen 2002). Moreover, the water boards are being subjected to a democratisation-process. As a consequence, water management has gradually been forced to 'open up', and to become more transparent to its citizens by abandoning its isolated, expert-based and technocratic policy style. Signs of this transition, from 'government-to-governance', can be found in the explicit need for stakeholders to become more involved. However, one could (again) question the 'depth' of this institutional change, because in truth the basic administrative structures have still remained unchanged thus far, as we will see in the next section.

2.3.8 Policy actors

Actor constellation

Dutch constitution consists of three general administrative levels, the municipal and provincial authorities and the national administration. Only the issue of water management has an additional fourth layer: the water boards. The (public) authority of the water boards is geared towards the management of regional water quantity and water quality and those of the province towards groundwater management and the planning aspects of regional water management. The national General-Directorate is responsible for the main water infrastructure of large rivers and canals, and, finally, the local authorities deal with urban water and sewerage.

Two central policy developments arise, in searching for changes in the actor constellation. In the context of the European Water Framework Directive, the river basin approach asks for co-operation within the four river basins of which the Netherlands forms a part – Meuse, Rhine, Scheldt and Eems – and thus, it asks for cross-border water management. But there is no prescription on *how* to co-operate. Up until now, it has not led to any new arrangements in the Netherlands, but instead the co-operation between the *existing* water authorities, nationally and internationally (such as the International Rhine Commission) continues to predominate (Backes 1999). When it comes to water quantity issues, the successive discussions on the WB21-Committee -proposals eventually led to a so-called 'National Administrative Agreement on Water', between the national administration and the representatives of the municipalities, provinces and water boards. In short, in both water quality (the WFD) and water quantity (the WB21-policies) the Dutch authorities have clung to the *existing* organisational order and division of tasks and competencies. The Vice-Minister of Transport, Public Works and Water Management illustrated this nicely by saying that she did not want "an institutional discussion" concerning the administrative organisation of Dutch water management.

Interaction patterns

Although there are no crucial changes in the actor constellation, there is a change in interaction patterns, namely a move towards decentralising responsibilities and empowering the regional authorities. This can be illustrated by the process architecture of the 'Room for the River' operation. The provinces have obtained an essential voice in this policy process by giving a so-called 'weighty advice', and because of their strong representation in the advisory steering committees. Both the WB21-policies and the WFD have stimulated territorial shifts towards the region and sub-river basins. We can conclude that the interactions between the different governmental layers are being intensified and that this is heading towards more co-operative and horizontal interrelations, accompanied by the empowerment of provinces and – to a lesser extent – the water boards and the municipalities.

Finally, we should pay attention to changes in the way the water management agencies and the public interact. Here we refer to the ways in which communication has changed since the discourse of “living with water” and the ‘disclosure’ of water agencies. It is again difficult to draw conclusions on ‘real’ institutional change. On the one hand, a change has definitely taken place in policy style, through communication campaigns directed at the public, increasing information flow, stakeholder involvement, etc. On the other hand, when it comes to taking critical steps towards formulating and preparing policy, such as the first stages of the implementation of the WFD or the selection of the location of so-called emergency flooding areas (these are strongly contested in some of the preserved areas), stakeholder participation is much less appreciated and the policy style of the Ministry and the Directorate-General for Public Works and Water Management is, once again, regarded as being technocratic.

Coalitions and oppositions

Are we witnessing new coalitions between water management and other policy actors in the field? Here we notice that, since the upcoming ecosystem-based discourse in river management, agencies that were traditionally primarily engaged with nature conservation, biodiversity and forest policies are entering the policy arena of water management. For example, the National Forest Service has presented interesting ideas on how to combine water safety issues with nature development and the ecological management of river basins (e.g. Ministry of Agriculture, Nature Conservation and Food Quality 2003). Even though today’s water policy seems to have been broadened towards issues of environment, landscape and nature conservation, and even though ‘non-water’ authorities are sometimes key players in *ad hoc* co-operation in floodplain projects, we still cannot conclude that this has led to new actor coalitions in the water policy arrangement as a whole. At some points the National Forest Service, with a strong nature conservation perspective, has in fact competed with the Directorate-General for Transport, Public Works and Water Management, which is mainly focussed on safety issues in river reconstruction (Van de Bilt 2004).

2.3.9 Shifts in power

The two indicators ‘resource constellation’ and ‘power relations’ point at the possibility of new and/or more resources in a policy domain on the one hand, and the (re)division of these resources over the different key players, potentially implying new power relations, on the other. With regard to the former, we can observe two important changes: (1) there are more financial resources for water management at our disposal; and (2) there are more resources to further develop the knowledge infrastructure in particular. However, with regard to the second indicator, we cannot see structural changes in power balance. As was previously stated, the organizational structure of Dutch water management has remained rather stable so far. We can observe the same key players and (more or less) the same power relations, as the new resources have strengthened those who were already ‘in power’, neither the smaller parties nor the newcomers. Below we will elaborate on these observations.

Resource constellation

Rudely awakened by the (near-)floods in the 1990s, and triggered by the concern over the effects of climate change, new resources have become available for water management, both internationally and nationally. First of all, extra public money was spent on strengthening the river dikes as quickly as possible (Wiering and Driessen 2001). Secondly, money was set aside in order to design the *new* policies for water management, in line with the results of the WB21-Committee. In the National Administrative Agreement on Water it was stated that, from 2003 to 2015, an amount of 8 billion euros should be reserved; for the period preceding 2050 about 16 billion euros. Over the short term this entails an investment of 1.3 billion euros until 2007 (Ministry of Transport, Public Works and Water Management 2003).

Besides these overall financial resources belonging to the policy domain, another essential resource is its knowledge infrastructure. There are several important power resources in water management. For instance, the supreme technical knowledge possessed by water managers, and - to a large extent - the public trust that has traditionally been given to this functional layer – combined with important legislative powers (e.g. risk norm-setting and water-related taxes). The knowledge system of Dutch water management is dominated by specialised governmental services, knowledge institutes and universities. The most important agency is the Directorate-General of the Ministry itself and within this central agency there are specialised services such as the Royal Institute for the Coastal Zone and Sea (RIKZ) or the Institute for Inland Water Management and Waste Water Treatment (RIZA), which are now operating on a more independent basis. The external knowledge infrastructure of Dutch water management is also extensive; especially WL Delft Hydraulics has to be mentioned here (Delft University is traditionally the cradle of Dutch ‘hydraulic engineering’) but also other institutes are active in the field: the Netherlands Organisation for Applied Scientific Research (TNO), the Meteorological Institute (KNMI) as well as the universities of Wageningen, Nijmegen, Twente and Utrecht. The differentiated knowledge infrastructure (hydraulic engineering, hydrobiology, ecology, policy analysis, rural and urban areas) of these various institutes is gathered together in the Netherlands Centre for River Studies (NCR). Furthermore, there is intensive co-operation between the specialised water management services, the environment and health research institute (RIVM) and Alterra, the research institute for the green living environment. The knowledge infrastructure on water is strong, but it is also narrowly focused on technical issues. A background report on the knowledge for integrated water management (Wisserhof 2000) claimed that the financial impulses for multi-disciplinary projects did not, as a rule, stem from the water sector, but from other ministries such as the Ministry for Housing, Spatial Planning and Environment or the Ministry for Agriculture, Nature Conservation and Food Quality. But, according to Wisserhof (2000), the broadening of the knowledge system and co-operation is ‘in statu nascendi’.

What about the new resources that have become available for research on water management, both in the Netherlands and Europe? Indeed, there are new research programmes, for instance the NWO Research Council for Earth and Life Sciences (ALW) and the Netherlands Foundation for the Advancement of Tropical Research (WOTRO) have launched a research program for fundamental research regarding fresh water; there is a new ICES/KISS program ‘living with water’ (in total 45.7 million euros, with co-financing) and there are many new initiatives which concern climate change, water and space, water and society, etc. In general, more financial means have been allocated to water management and research on the topic.

Power relations

What does these new resources for water policy and -research signify when it comes to the re-allocation of resources and power relations? With regard to the supplementary public funds made available for the traditional policy of strengthening the dikes (after the 1995 near-flooding disaster), the financial resources went to the core players of Dutch water management: the Directorate-General, the provinces and the water boards. With that, the position of the existing authorities, services and institutes was in fact strengthened. Moreover, in the new ‘room for the river’ -policy and other WB21-policy measures, as well as the Water Framework Directive, the existing agencies continue to dominate policy formulation and policy implementation, although there is more involvement of other policy domains and more sensibility towards other political arenas and knowledge resources. Considering the content of various research programs, the focus has (partially) shifted and a gradual broadening of the knowledge infrastructure has occurred, with more attention being paid to ecological research, social science research, policy analysis and socio-economic research. At the moment, it is impossible to identify the changing allocation of budgets and their impact on individual research agencies exactly, but we can sense several new initiatives that indicate

the general trends. First of all, the Netherlands Centre for River Studies is an attempt to integrate the various knowledge sources. The NCR managed a large research umbrella project in the framework of the EU Structure fund (the IRMA- Interregional Rhine/Meuse Action - Programme). We can also witness several new clusters geared towards technical knowledge, i.e. hydraulic engineering (i.e. Delft cluster) and new portals that show co-operation between existing knowledge institutes (e.g. Coordinated Program on Water and Climate). Apparently, the existing well vested research institutes on water, nature conservation, environment, climate and health are searching for new joint ventures in the water domain. All in all, we can observe an increase in public resources for water policy and a gradually broadening of water research, although the division of these resources over the different key players in the water policy field have roughly remained the same. Granted, additional money went to newcomers, e.g. in policy analysis, and ecologists seem to have been strengthened by the broadening and division of resources, but these are minor shifts, compared to the budgets of the 'big players' in the field.

3 Theories on explaining policy and legal change

Dries Hegger, Peter Driessen, Marlous van Herten, Anoeska Buijze, Jean-Baptiste Trémorin, Willemijn Van Doorn-Hoekveld, Marleen van Rijswick and Carel Dieperink

3.1 Introduction

The explanatory framework draws heavily on literature from the field of policy analysis (e.g. Sabatier and Weible 2007; True et al. 2007; Zahariadis 2007) but also uses complementary insights that can be derived from the field of Science and Technology Studies (STS) (e.g. Hughes 1987) as well as from the Multi-Level-Perspective as it has been developed by Dutch scholars analysing socio-technical transitions (e.g.

2002). This appendix provides a general overview of the main theories on which the explanatory framework draws. Theories discussed include the Advocacy Coalitions Framework (Sabatier and Weible 2007), the Multiple Streams Approach (Zahariadis 2007), Punctuated equilibrium theory (True et al. 2007), the Multi-Level Perspective from transitions theory (Geels 2002) as well as some other relevant theories. This discussion results in an overview of potentially relevant factors for explaining policy change. As we indicated in the main text, the following issues should be carefully considered when setting up explanations:

1. **The explanandum** (the dependent variable) (Capano and Howlett 2009; Dupuis and Biesbroek 2013; Howlett and Cashore 2009). The latter is sometimes conceptualised as change in *output* and sometimes as a change in *processes*. From Valman (2012) we can derive different models of change. She distinguishes between displacement, layering, drift and conversion. Displacement happens “when new modes of practice or new rules replace the existing previously taken for granted forms, settings or practices.” (p2) “layering means that new rules are introduced and put alongside existing ones”, “drift takes place when the consequences of existing rules alter due to changes in context” and “conversion differentiates from drift in that the new interpretation of rules is more active compared to when drift takes place”. Within the STAR-FLOOD project, and especially in the step of analysing Flood Risk Governance Arrangements – the step preceding the explanation of stability and change in flood risk governance – it will be necessary to very precisely denominate the degree of stability AND change, possibly by considering all four dimensions of the PAA. It should also be carefully established if change should be seen as **evolutionary** (incremental) or **revolutionary** (radical) (Capano 2009; Capano and Howlett 2009);
2. **Explanatory factors.** A discussion of explanatory factors should include, amongst others, an elaboration of whether change is seen to be **endogenous** change that is change occurring from within a policy system of change that comes from **outside** (Capano and Howlett 2009; Howlett and Cashore 2009). One should also determine the relative importance of the role of **structures vs. agency** (Capano and Howlett 2009), that is the role of individuals vis-à-vis the social structures they are part of;
3. **The explanation behind the explanation.** When assessing the explanation behind the explanation, it should be assessed to what extent change or the absence thereof can be attributed to **chance**. According to Capano (2009:26) “chance” or “serendipity” can explain stability and change, suggesting that there are limitations both to the ability of researchers to theorise changes in flood risk governance and to the potential of actors to purposefully steer developments therein.
4. **Establishing evidence.** When establishing chains of evidence, it is important to compare competing explanations and make use of different sources of data (triangulation). Zittoun (2009) argues that many policy analysts, through the use of their analytical techniques, “distort the object i.e. public policy beyond recognition” (p.65). He therefore makes a case for observing “how the participants produce this identification [of policy change] how they identify and define

problems and transform instruments into action” (p. 80). He also argues that “rather than identifying on one side, the networks and on the other their beliefs, we would like to consider that it is during the experimentation with the connections between belief, problem and public policy that the contingent coalitions are formed which ultimately determine policy content” (p. 80). Indeed, for us as STAR-FLOOD researchers it will likely be fruitful to look at our object of research, flood risk governance, through multiple theoretical lenses, including lenses that presuppose an “objective” reality that can be studied as such and a more “subjectivist” perspective in which we look at perceptions of actors, sense making and attribution of meaning.

The following sections will discuss the main theories on which the explanatory framework discussed in this report is based.

3.2 The Multiple Streams Framework (MSF)

The MSF, originally developed by John Kingdon (1984) is a prominent framework conceptualising policy change. The framework presupposes the existence of three relatively independent “streams”, those of problems, policies and politics (Zahariadis 2007). According to the MSF, often these streams are not connected. Policies are then made in policy subsystems. The main role in connecting *problems*, *policies* and *politics* is reserved for *policy entrepreneurs* who make use of *policy windows* that exist from time to time. A basic assumption behind the model is that individuals involved in policymaking are boundedly rational. They can devote their attention only to a limited number of issues at a time. Therefore, the context in which they operate significantly influences what captures their attention. Key concepts of the MSF are (Zahariadis 2007: 71): the problem stream, the politics stream, the policy stream, policy windows and policy entrepreneurs. These five factors together are expected to determine policy output.

In relation to the MSF, an analysis of Zohlnhöfer (2009) is interesting because it provides – without explicitly referring to the MSF – a more in-depth analysis of what happens when issues do capture political attention. First of all, Zohlnhöfer makes the point that types of policies differ according to the degree in which they tend to capture the attention of high level politicians with decision making power, including the so-called “veto players” (parliaments, presidents in presidential systems). Some policies like those related to taxation and social security systems will grab political attention in most countries most of the time, whereas others, including foreign and environmental policies will do so only from time to time (it seems safe to assume that in most countries Flood Risk Governance will belong to the latter category of policies). According to Zohlnhöfer, who bases himself on empirical research into the role of politics in policymaking more generally, the fate of policies in political processes has some degree of predictability. For instance, governments are very likely to favour the status quo or only moderate departures from it for electoral reasons. Governments may also adopt different types of policies but (Zohlnhöfer 2009: 103): “The adoption of policies may be impeded by electoral considerations unless the government is confronted with problems that put its re-election at risk”. Also (p. 104): “The further the status quo is from the ideal point of the incoming government, the greater the policy change will be all else being constant”. One can also expect, according to Zohlnhöfer, that if radical change is undertaken, it is most likely done at the beginning of a new government period. The current document is not the place to discuss all such theoretical relationships in great detail. It is, however, good to know that some expectations can be raised regarding what happens when policies come into the reach of important political decision makers. For STAR-FLOOD, the question to be addressed is at which moments flood risk governance entered high-level political agendas, whether or not decisions were taken to make profound changes in approaches or legislation, how these dynamics are to be explained and, most importantly, what the relative importance of this political process has been in changes in flood risk governance more generally.

Box 3.1 Examples of questions that could be posed inspired by the MSF

- -Which flood-related problems can be distinguished? When did they get the attention of policymakers? When did they get the attention of politicians? Which factors ensured this (lack of) coupling of problem, policy and politics streams?
- -Did flood problems enter political agendas (both at the lower and higher government levels)? If so, were decisions taken to make profound changes in approaches, policies or legislation? Why (not)? What could be the relative importance of these political processes for changes in flood risk governance more generally?

3.3 Punctuated Equilibrium Theory (PET)

Punctuated Equilibrium Theory (True et al. 2007) seems to be a useful theory for trying to explain stability and change in flood risk governance because the theory explicitly addresses both. PET assumes that most policy processes can be characterised by long periods of relative stability punctuated with short periods of major change. Similar to MSF, PET departs from the assumption of boundedly rational individuals who have to operate in the context of existing institutions. According to PET, most policy issues rarely dominate the political agenda, but at some points in time (True et al. 2007: 158) “some issues catch fire, dominate the agenda, and result in changes in one or more subsystems. The explanation for the same political institutions producing both stasis and punctuations can be found in the processes of agenda setting – especially the dynamics produced by bounded rationality and serial information processing”. Another notion to be derived from PET is that “like earthquakes or landslides, policy punctuations can be precipitated by a mighty blow, an event that simply cannot be ignored, or by relatively minor events that add up over longer periods of time” (p. 160). An important implication for STAR-FLOOD would be that it is probably good to assume that *change is always underway*. Each event at every point in time could be a contributor to change or stability: the devil is in the details.

Related to this, Rayner (2009) argues that what happens in periods of relative stability is probably “to identify a range of strategic possibilities for action. The existence of these strategic possibilities and the use that is made of them by historical agents explains both the durability of the “period” and its ultimate collapse and transition into a new one” (p. 87). It is an important question to what extent change derives from exogenous shocks or from internal contradictions. As Rayner (2009: 91) makes us aware, sometimes [shock events] “tend to close policy windows and inhibit change rather than the reverse”.

There are some other potentially useful notions within the PET framework. PET ascribes an important role to *policy images* or the framing of policy issues and it assumes that there are multiple *policy venues*. The places where policies are made may be as wide-ranging as parliaments, state agencies, ministries, universities, congresses, the media and others. Within STAR-FLOOD, these venues will have to be systematically mapped, to enable the researcher to make an effort to determine their relative importance and influence.

Box 3.2 Examples of questions that could be posed inspired by PET

- What were periods of relative stability in flood risk governance and when did major changes occur? If major changes occurred, can they be related to shock events (e.g. floods, but also major changes in the broader societal context)? If yes, how probable is it that this relationship is a causal one? What minor changes (e.g. actors identifying strategic options for action) have occurred during periods of relative stability and do they contribute to an explanation of later more profound changes?
- In what venues were flood risk related policies made (e.g. parliaments, conferences, state agencies, Ministries, universities, the media)? What was the relative importance of each of them for explaining policy change and stability?
- How were flood issues framed? Which policy images were created and used and by whom? What is the significance of this for explaining stability and change in flood risk governance?

3.4 Advocacy Coalitions Framework (ACF)

The ACF can be seen as an encompassing framework because in principle various other frameworks and approaches are compatible with it. The framework as originally proposed by Sabatier and Jenkins-Smith (1988) and further developed later (e.g. Sabatier and Weible 2007) has been specified to the analysis of wicked problems (ibid). This makes the framework a good candidate for the analysis of flood risk governance, which in some cases possesses some characteristics of wicked problems, including value pluralities, uncertainties and high stakes. Sabatier and others furthermore attach an important role to technical information, something which has traditionally been important in the domain of flood risks. Sabatier and Weible (2007: 192) state: “that researchers, (university scientists, policy analysts, consultants etc.) are among the central players in a policy process”. A question that could be symmetrically considered within the STAR-FLOOD project would then be if various forms of user engagement in research might explain certain policy outcomes and which forms of engagement in which research phase have led to which outcomes (see also Talwar et al. 2011).

At the core of the ACF lies the assumption that in each policy sub-system we may find multiple (at least two) competing advocacy coalitions, that is coalitions of actors that converge in their ideas and compete with other coalitions. Actors within these coalitions have certain policy beliefs as well as a certain amount and type of resources (including 1 formal legal authority; 2 public opinion; 3 information; 4 mobilizable troops; 5 financial resources; 6 skilful leadership) (Sabatier and Weible 2007: 203). According to ACF, the question what actually takes place is an empirical question. The framework does not *presuppose* the existence of different coalitions, but states that their existence and modus operandi should be empirically assessed. For the purpose of the current research, the question is if we do find these advocacy coalitions in the field of flood risk governance in certain countries and cases and at which level, e.g. country or case and most importantly, to what extent stability and change could be *explained* by the presence of these advocacy coalitions.

Within the ACF, it is assumed that policy subsystems are nested within a *broader physical and societal context* which *is* relatively stable and *contributes to* policy stability. In its initial formulation, therefore, within ACF it was assumed that major policy change can come from *external shocks* and *policy oriented learning*. In recent revisions of the framework, two other change patterns have been added: *internal shocks* and *negotiated agreements*. In all cases, the relative stability of policies in the short term was a reason for Sabatier and others to plead for diachronic analyses in which policy developments of at least a decade are taken into account. ACF gives an important role to the so-called *deep core beliefs* and *policy core beliefs* of actors. Changes therein are seen as an important explanatory factor for policy change. Empirical questions for the STAR-FLOOD project derived from the ACF would then be how the deep core beliefs and policy core beliefs of actors involved in flood risk governance can be characterised or whether external shocks have influenced the policy core

beliefs of a dominant advocacy coalition. However, Real-Dato (2009) stresses that ACF's focus on beliefs is one-sided and that both *ideas* (beliefs) and *interests* should be considered when trying to explain policy change.

Box 3.3 Examples of questions that could be posed using the Advocacy Coalitions Framework

- Which advocacy coalitions (if any) can we distinguish in flood relevant policies and at what levels (case, National Flood Policies and Regulations Domain)? How can we characterise the deep core beliefs (e.g. preference for state-led or market-led solutions) and the policy core beliefs of the actors in each coalition? What do these coalitions look like (e.g. which ties can be found between which actors)? To what extent could change or the absence thereof be attributed to these actors' beliefs and to what extent could it be attributed to their interests?
- In National Flood Policies and Regulations Domains, can we find evidence of learning? If so, how did it take place and why? How probable is it that this learning contributed to (the absence of) policy change?
- Can we find examples of conflict expansion within National Flood Policies and Regulations Domains? If so, who initiated this conflict expansion, why and to what effect?

3.5 The Multi-Level Perspective from transitions theory

3.5.1 Introduction

The Multi-Level Perspective from transition theory comes from another body of literature than the explanatory frameworks discussed above. The scholars advancing the perspective focus on the analysis and explanation of transitions in socio-technical systems. Also in scientific literature, only few connections have been made between transition theory and theories from the field of policy analysis. It is, however, useful to discuss the MLP as an additional framework, because from the MLP some hypotheses can be derived regarding relationships between different levels. In the MLP these levels are referred to with different terms than we do in the current report, but there are some similarities between these levels. The macro level within the MLP is what we refer to as 'context', the meso or regime level resembles what we refer to as 'National Flood Policies and Regulations Domain'. Only the niche level does not – as we will show – resemble what we refer to as the case level.

3.5.2 Explaining stability and change from a Multi-Level Perspective

The multi-level perspective (MLP) is a heuristic tool that helps to explain stability and change of socio-technical systems. It is often used in research strands that focus on transitions. A transition is in that strand of literature referred to as a gradual, continuous process of transformation of a societal system. The multi-level perspective draws attention to three levels of reality: the macro-, meso- and micro-level, that are part of a nested hierarchy (see Figure 3.1).

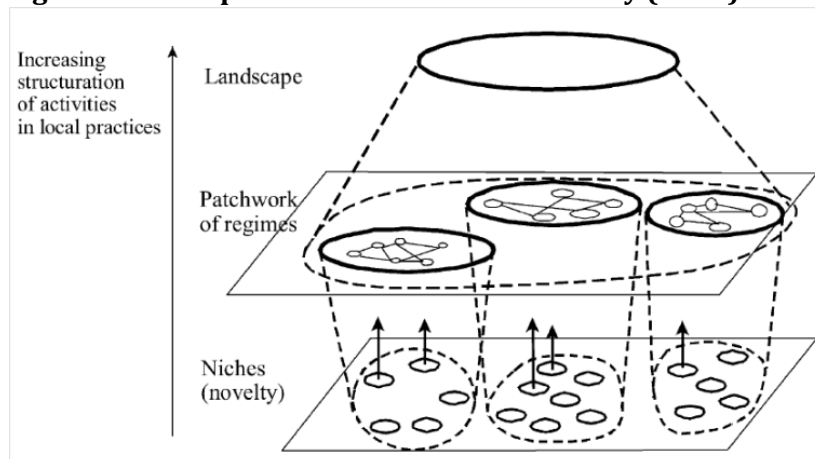
The **macro-level** (by us referred to as context level) encompasses the context, the wider exogenous environment that influences societal systems. Climate change (including global warming and sea-level rise) or economic globalization are examples of elements belonging to the macro-level. Macro-factors are considered to be persistent. On the short term, they are beyond the direct influence of actors and they cannot be changed at will (Geels 2005). On the very long term, persistent mutually supporting initiatives and trends might be able to affect the macro-factors, but the change process is inherently slow. In the MLP approach, the metaphor 'landscape' is used to refer to the macro-level. A landscape can encompass both tangible facets of the natural and built environment (e.g. material and spatial organization of cities, including flood defence infrastructure) and intangible facets (such as national culture, world views, political beliefs, social values).

The **meso-level** refers to the prevailing system installed to deliver a specific societal function. This system can be described as the 'normal' way of doing things. It encompasses a web of interlinked actors, following a set of rules, respecting dominant power relations, and confirming and reconfirming discourses [notice the link with the dimensions of the policy arrangement approach]. In the MLP approach, the so-called 'regime' sits at the meso-level, sandwiched between the macro-level of the landscape and the micro-level of the niche (Berkhout et al. 2003). Within transition theory, it is assumed that change occurs at the regime level incrementally and is geared to achieving optimization, rather than deep change. Inertia is seen as an explanation of difficulties in achieving transitions. In many regimes the vested interests contribute to stability; there is hardly any room for innovation. Shocking events or radical ruptures might offer a 'window of opportunity' for a regime to change. Note that this diagnosis differs from the one offered for instance by the Advocacy Coalitions Framework discussed above, that envisages much more room for change 'from within' than the MLP from transitions theory, through conflict expansion and endogenous learning. This reconfirms the need to actively envisage multiple competing explanations and discuss the plausibility of these.

Whereas the macro and meso level of the MLP resemble our analytical levels of context and National Flood Policies and Regulations Domain, the micro-level in the MLP is something completely different from the case level in our conceptualisation. Within the MLP, the micro-level is the level at which space is provided for experimentation. At this level, we see innovators working, on a small scale, to develop new approaches to providing societal functions (Paredis 2009). Also, at this level, we see the emergence of new practices that deviate from the normal way of doing things (at the regime level). In the MLP approach, 'niches' can be found at the micro-level. Niches are less subject to the influence of the regime and they can act as a safe environment in which breakthrough developments can grow, sheltered from the selection process that occurs at regime level (Geels 2002). A regime may host a range of niches which generate innovations to challenge the status-quo.

Hence, within the MLP, it is assumed that an important driving force for transitions comes from the deliberate creation of protected spaces (niches) which are somehow set apart from the incumbent regime, e.g. through the use of legal exemptions (in the domain of flood risk governance these could include rules for experimenting, exemptions for specific areas or the use of more open norms that provide space for innovation), the setting up of pilot projects or the granting of subsidies. For STAR-FLOOD, it is an interesting empirical question to assess if we can find such examples of 'protected spaces' and what actually constitutes the protection. However, there is no reason to a priori assume that these protected spaces are to be found at case level and if they are, that they can *only* be found at case level. On the contrary, we may very well find out that the case level is the level where most resistance and rigidity is present and where most examples of path-dependency and lock-in manifest themselves. For similar reasons, we also do not a priori assume that, as MLP scholars state, transitions are only possible when there is enough pressure coming from the micro- and macro-level. As said before, change from within, at regime/policy sub system level may very well be possible. It seems safest to assume that at all three levels (case, National Flood Policies and Regulations Domain, and context) we may find both factors that contribute to change and factors that inhibit change.

Figure 3.1 Multiple levels as a nested hierarchy (2002)



Box 3.4 Examples of questions that could be posed inspired by the MLP

- In explaining stability and change, what role is played by the context level? For example, did any shock events occur, and if so, what were their effects? Did they reinforce stability or did they trigger major change?
- Can we find examples of ‘protected spaces’ (niches) ? If so, what constituted the protection (e.g. legal exemptions, pilot projects, subsidies)? Where do we find these niches (e.g. certain geographical regions, certain types of river courses, certain sectors etc.)?
- How ‘receptive’ is the National Flood Policies and Regulations Domain for change? For instance, how stable/dynamics are certain rules and regulations? How open is it to the entrance of new actors? Has a discursive shift occurred or not?

3.6 Some syntheses between the policy analytical frameworks for explaining policy change

When we compare the three frameworks with the help of the six factors denominated in the section on conceptual and epistemological starting points, some syntheses and differences can be identified:

1. **Structure vs. agency.** The three policy analysis frameworks to some extent seem to presuppose a duality of structure (Giddens 1984) in which actors are enabled and constrained by their structural context and have some possibilities to change this context. The MSF attaches a large role to agency by putting policy entrepreneurs who use policy windows forward as the main explanatory factor for policy change. Also the ACF provides much space for agency through its focus on coalitions of actors and their beliefs. The main focus of PET is on policy subsystems. Hence, analyses performed using the PET framework will put more emphasis on structural preconditions. Also PET notions like *policy images* and *policy venues* can be placed closer to the structure side than to the agency side. As opposed to the policy analysis frameworks, the Multi-Level-Perspective leaves very little room for agency: it focuses entirely on systems.
2. **Endogenous vs. exogenous change.** All the three policy analysis frameworks seem to attach importance both to exogenous and endogenous factors. The MLP, on the other hand, only talks about endogenous factors in the sense of referring to a regime that is more or less ‘receptive’ to change. The main origin of the change is expected to come from outside the regime (exogenous change). MSF focuses on how and why policy streams become connected to problem and political streams. PET argues that many developments may be going on within a policy subsystem before large observable change occurs. These changes are, however, expected to be triggered by shock events, which in principle can originate from within and from outside the policy subsystem. In the ACF, various endogenous and exogenous sources of change have been explicitly denominated. The former include internal shocks, policy-oriented learning and negotiated agreements. The latter includes external shocks induced by the wider context of the

policy subsystem. This similarity between the frameworks is acknowledged by Real-Dato (2009) who argues that endogenous change often occurs via learning. According to Real-Dato, the chance that learning takes place is enhanced by the existence of institutional elements within organisations and at sub-system level designed to foster it, such as internal or external evaluations, consultative bodies, professional fora, information systems integrated in policy implementation procedures etc. Learning is believed to be less likely in contexts in which individuals with alternative views are easily marginalised. Another mechanism of endogenous change is conflict expansion (Real-Dato 2009). Contrary to learning, this is expected to be originated by outsider participants. Finally, Real-Dato distinguishes *exogenous impacts* of the types discussed above. These include changes in material conditions, attributes of the community (socio-economic conditions, public mood), focusing events etcetera.

3. **Evolutionary vs. radical change.** MSF and ACF are not very explicit in whether they conceptualises change as evolutionary or radical. On the contrary, PET assumes change to be BOTH evolutionary and radical although the criteria for denominating something as a radical or revolutionary change remain of course arbitrary. The MLP was developed to analyse socio-technical transitions and hence focuses by definition on radical change.
4. The **explanandum.** Looking at the explanandum, the three policy analysis frameworks put different emphases. MSF tries to explain the extent to which connections are made between the three steams of policies, politics and problems. The main question that is posed is why some issues enter certain agendas and why not. Also PET heavily emphasises the issue of agenda setting. It furthermore poses the question of why the supposed punctuation of major and minor change occurs. The ACF does not very specifically denominate what it is that should be explained, but it seems that this is the policy output and policy impacts of policies within policy subsystems. Regarding the explanandum, Real-Dato makes the important point that the time span under consideration to some extent determines if the researcher observes change. It has been shown that different events can be seen as stability and change, depending on the length of the time frame under consideration and, related to this, the baseline situation that is – often implicitly – assumed (change compared to what?) (see also Rayner 2009). The MLP intends to explain changes in socio-technical systems which are assumed to comprise three different levels (micro, meso and macro level).

It seems that issue 5, the extent to which change or the absence thereof can be attributed to chance (B4.7) is not explicitly addressed in any of the frameworks. Issue 6, the extent to which policy change can be objectified (B4.7) is addressed implicitly in the sense that all frameworks seem to presuppose that policy analysis and explanation can be rather objectivistic. However, “subjectivist” factors are included in the PET framework (policy images), while ACF gives the beliefs of actors centre stage.

3.7 The relationship between structure and agency

An assumption central to the STAR-FLOOD proposal and also to the PAA (Arts et al. 2006) is that of a *duality of structure* (Giddens 1984). In the past structure and agency were viewed as a dichotomy. Most modern social theorists, however, now stress the reciprocity between structure and agency. As Giddens (1984: 19) puts it “One of the main propositions of structuration theory is that the rules and resources drawn upon in the production and reproductions of social action are at the same time the means of system reproduction (duality of structure)”.

Individuals are part of wider systems and are influenced by structures but at the same time through their actions they contribute to on-going system reproduction. Especially when we start to delve into the role of change agents, it will be key not to underestimate their role by focusing solely at the level of policy the National Flood Policies and Regulations Domain. But neither will a focus solely on individuals, attaching an almost heroic status to them, be appropriate. The challenge will be to arrive at a balanced analysis.

Giddens (1984) often uses language as an illustration – and an example – of what he means with his notion of duality of structure. When people use a language, they make use of certain rules and resources (grammar, vocabulary). Only if they do so – nearly – correctly, others will understand what they are saying or writing. However, how users exactly speak or write may shift gradually over time. New vocabulary is added, other words cease to be used, the way grammar is used changes over time. If enough people stop using a word, it may even ‘cease to exist’. Vice versa, if enough people start using a word, after some time it may be included in dictionaries and rules on when to use it may become formalised. This example of language highlights that many ‘normal’ ways of doing things are tacitly reproduced.

There are different degrees in the ease with which structural factors can change (see e.g. Williamson’s (2000) distinction between four levels of institutions, some of which can change much easier than others). In some cases language can easily evolve (e.g. new words, slang (street language) etc. emerge every year), but it can also be really stubborn when it is connected to deeper core beliefs. This is the case when language produces a discourse, in which it actually matters **how** something is said. This way of speaking is actively maintained by the advocates of the discourse.

3.8 Change agency literature

Caldwell (2003) has synthesized much literature on change agency. From this synthesis he has developed a proposal for a classification according to four types of change agents: *leadership*, *management*, *consultancy* and *team* models. Each category includes various theoretical streams which we will not discuss in detail here. An important lesson for STAR-FLOOD is, however, that change agents are likely to be found at different places and in different roles. They can be leaders or senior executives, middle level managers, external or internal consultants and they can work at a strategic or at an operational level. Change agents can even be teams. In our empirical research, we will have to address the question which types of change agents we can encounter where, which factors determine their influence or the lack thereof and what their relative importance for explaining policy change is. Also Huitema et al. (2011) have delved into the notion of change agency. They have analysed strategies of policy entrepreneurs in water transitions and found that these individuals can employ various *types of strategies* to reach their goals. These include the development of ideas, the building of coalitions, the selling of ideas, recognizing and exploiting windows of opportunity, orchestrating and managing networks and recognizing, exploiting, creating and/or manipulating multiple venues.

Olsson and Hysing (2012) add some other insights to the literature on change agency. First, they hypothesise that various contemporary societal developments provide individuals with more room for manoeuvre. These developments include a broader shift from government to governance, a trend towards more civic engagement, a growing demand for professionals as expert consultants and a general weakening of democratic power and an increase in bureaucratic power. Olsson and Hysing found a new type of change agent called the inside activist. This is someone who is engaged in civil society networks and organizations, who holds a formal position within public administration and who acts strategically from inside public administration to change government policy and action in line with a personal value commitment.

Brouwer and Biermann (2011), partly in line with Huitema et al. have analysed strategies of policy entrepreneurs in Dutch water management. Contrary to Huitema et al. their analysis was largely focused at the sub-national level. They conceptualised policy entrepreneurs as “risk-taking bureaucrats who seek to change policy and are involved throughout the policy change process” (p.5). Based on a synthesis of existing theories – including MSF, ACF en PET – and after empirical research they identified a range of strategies policy entrepreneurs make use of: (1) attention and

support-seeking strategies, to demonstrate the significance of a problem and to convince a wide range of participants about their preferred policy; (2) linking strategies, to link with other parties, projects, ideas, and policy games; (3) relational management strategies, to manage the relational factor in policy-change trajectories; and finally, (4) arena strategies, to influence the time and place wherein decisions are made. Within each of these four types, a number of concrete strategies were identified as depicted in table 3.2.

Table 3.2: Strategies of policy entrepreneurs found in the Dutch water management domain

Attention and support seeking strategies	Linking strategies	Relational management strategies	Arena strategies
-pilot projects -indicators -focusing events -rhetoric -correlating	-coalition building -selective activation and exclusion -issue linking -game linking	-trust building -networking	-venue shopping -timing

Brouwer and Biermann arrived at a number of conclusions of relevance for the research goal of STAR-FLOOD. First, their findings suggest that policy entrepreneurs are quite common and through their characteristics are relatively easy to identify – at least in the context of Dutch water management. It was found, for instance, that within Dutch water management 339 policy entrepreneurs can be found within 491 Dutch local government bodies. Second, these policy entrepreneurs are generally hard working people with entrepreneurial skills. With this, the authors mean that policy entrepreneurs want to achieve change and are willing to take risks to achieve these. Third, policy entrepreneurs have been found to use a mix of strategies in a way that is by Brouwer and Biermann referred to as “juggling”. The policy entrepreneurs are depicted as both streetwise and boundedly rational. Fourth, contrary to much of the literature discussed above, Brouwer and Biermann are relatively positive about the potential of policy entrepreneurs to achieve policy change. Based on their research, they question, for instance, one of the basic assumptions of the MSF, being that the streams of problems, policies and politics are relatively autonomous, are rarely coupled and if they do so are generally coupled ‘by accident’. Brouwer and Biermann go against this by claiming that policy entrepreneurs are essential key actors that are continuously working at the linking of the three streams.

A lesson for STAR-FLOOD is that we should continuously ask ourselves the question how much room for agency there is in achieving policy change in flood risk governance. While most of the discussed literature argues that the margins are small, Brouwer and Biermann, but also Olsson and Hysing provide arguments for the thesis that there might be more levers for action than is generally assumed. It is an empirical question how this works out within the countries and cases of STAR-FLOOD. To be able to address this question, it is necessary to look critically at the actual actions of individuals – including policy entrepreneurs. Who are they? What do they do exactly (which strategies do they use)? In which contexts do they operate and to what effect? When making this analysis, it is important to bear in mind that individuals cannot only act with the aim of establishing change, but on the contrary, that they may also use their entrepreneurial skills for maintaining the status quo.

Box 3.5 Examples of questions that could be posed inspired by literature on change agency

- What types of change agency can we find in the domain of flood-risk related policies? Who were the change agents (e.g. senior executives, middle level managers, external or internal

consultants)? Where do we find these change agents? Which strategies did they use (e.g. can they be described in terms of Huiteima et al.'s conceptualisation, or in terms of the four types of strategies depicted in table 1? In which contexts do they operate and to what effect?

- How can we characterise the relative room for manoeuvre for change agents in a specific country or case? Has this room for manoeuvre increased or decreased in the course of the years and why?

3.9 Relationship between explanations by policy analysts and legal scholars

The legal discipline, contrary to the social sciences, does not have dedicated theories, models or methods to explain the data provided by (empirical) research. This does not mean, however, that legal scholars cannot contribute to explanations for stability or dynamics, on the contrary. For one, even in the absence of dedicated theories, models and methods, legal scholars are familiar with making explanations. When they describe legislation, legal scholars generally tend to also describe or explain why this legislation has been established (e.g. positive law study). By identifying the purposes of new legislation, legal scholars can identify factors that may cause change or at least reflect a desire for change. Second, many of the factors explaining policy change have a legal component and one can generally expect policy change and legal change to be inextricably linked. In some cases, explanatory factors may be the cause of changes in legislation and case law as well as in policy. On the other hand, changes in legislation and case law may in turn be explanatory factors in themselves. An example of a dramatic legislative operation is the Dutch crisis and repair law, which on the one hand was inspired by the economic crisis and aimed to accelerate large infrastructural projects to combat the said crisis, and on the other hand includes some radical changes to existing legislation which prompts a certain response in actors. In sum, we expect the explanatory factors described in chapter 4 to be recognisable and useful both for social scientists and legal scholars.

The legal system both at NFPR and case study level is part of the Flood Risk Governance Arrangements whose stability or dynamics we intend to explain. As we have shown in chapter 3, legal factors are predominantly to be found at the rules and discourses dimension of the PAA. Since the legal system at the NFPR level is part of the explanandum, changes and stability in legal systems need to be explained. At the same time, the STAR-FLOOD researchers will have to determine the extent to and ways in which the legal system influences societal changes in turn.

In this respect, we can sketch the following potential interrelationships between changes in policies and changes in legal systems:

- One can logically expect that in many cases *legal systems respond to societal changes and change in turn* (e.g. through teleological interpretation, the adoption of new legislation or the contra-*legem* application of legal principles. Different legal systems will do this in different ways. This could be an explanatory factor for change in flood risk governance, because it might affect how change occurs. E.g.: differences in amount of public support needed before the legislator takes action or when courts take action.
- *Legal systems can to some extent obstruct, hinder, or contradict societal changes*. Put in other words, the ability of the legal system to accommodate change may not always be sufficient and law can be more conservative than society. In that sense, a rigid legal system could be an obstacle to change. This is also related to path-dependency. Once a legal system is in place, it may limit what further developments are possible: existing rights tend to be protected.
- Endogenous change within the legal system can take place, but this can be expected to be of limited importance overall. For civil law systems, there is a tendency of the legislator to respond to unforeseen events with new regulation tailored specifically to that event, resulting in increasingly detailed regulation (Van Rijswick and Salet 2012). There is a similar tendency in

common law systems to produce increasingly detailed rules in case law. In case of unmanageable levels of detail there may be legislative operations to codify and simplify (Van Rijswick and Salet 2012).

- Law can be used as a tool to instigate change. This requires action by rule makers either following democratic debate or for other reasons. It does not cause this change by itself, but the way legislation is drafted may affect whether the desired change is accomplished. The effectiveness of law as a tool to cause societal change is debated, because legal norms are not necessarily obeyed. (see also the remarks on effectiveness in chapter 5)

3.10 Natural law, positive law and their relevance for explain stability and dynamics in flood risk governance

Many explanations for change or stability in the legal framework can be understood through the concepts of natural law and positive law and their avatars in the legal framework. Natural law (or *lex naturalis*) is a system of law determined by nature: “it is a view that certain rights or values are inherent in or universally recognizable by virtue of human reason or human nature, while positive law is the legal tradition whereby certain rights or values are legally cognizable by virtue of judicial recognition or articulation”.

This opposition between natural law and positive law echoes the role attributed to law in society. Once again two opposite visions can be identified and are related to the formerly described distinction. On the one hand, law can be viewed as a tool to promote and convey a certain type of social model, according to some specific values or some overarching principles (fundamental rights for instance). On the other hand, law can be viewed as a responsive instrument to societal change, sticking to a social reality and implementing legal norms needed by these changes. As an illustration, France’s recent debate on gay marriage illustrates very well this opposition. Indeed, most of the debate was evolving around a clash between two perceptions:

- **An idealist vision of law** (natural law): giving or not giving the right to gay people to get married depends of what model of society you want to build in the future. Values related to the traditional family are important and giving the right to get married is going to rattle these values and models. Here law’s role to play is about carrying an idea of what society should be;
- **A more pragmatic vision of law** (positive law): gay people live together in France, there are many of them who live as if they were married and this is just the way things are. Society is made of this reality. In this view law’s part to play in society is to give them the right to get married in response to this specific societal evolution.

It is important to keep in mind this conceptual opposition as it underlies most of the explanatory factors for stability and change in the legal framework.

3.11 Stability and dynamics in the legal framework

3.11.1 Introduction

Legal explanatory factors can be seen as important structural factors. One may argue that legal explanatory factors are more relevant or prominent to explain stability than change. Indeed, the legal framework has a tendency to absorb change and, in some ways, to digest and assimilate new rules, laws or treaties much more slowly than any other systems. If we do find such inertia, it will be important to understand why it is there and to try to explain it.

3.11.2 Stability in legal frameworks

An important explanatory factor for stability in legal frameworks can be identified through the theory developed by the Austrian legal scholar Hans Kelsen, i.e. the **hierarchy of legal norms**. Hans Kelsen defines the static theory of law as a hierarchy of laws where the individual laws were related to another as either being inferior, or superior with respect to each other. This hierarchical aspect of the legal framework ensures some sort of stability and a hypothesis could be that the higher the legal norm is, the more stable it should be. The highest principles at the top of the pyramid can be written down in a Constitution as in France, or can be identified by the judges as in an Anglo-Saxon common law model. The level of importance given to each of these principles is usually revelatory of a legal and political tradition.

Concerning floods, on the top of this pyramid there are some founding principles for each national legal framework that are directly linked to this matter. **Property right, security and liability** stand as the most common founding principles shared by most of western European countries. Indeed public administration has generally three main duties: to respect everyone's property right, to protect each citizen from natural disasters and to provide a compensation system taking into account the responsibility of each actor. The main challenge for all legal frameworks is to find a balance between all these principles. This tension between them limits the innovative potential and the discretionary power (decisions and interventions) of the administration. More generally, fundamental principles can be seen as a strong explanation for stability from a legal point of view. They provide, as superior norms that can be referred to, a steady ground for the legal environment. These fundamental principles are fairly intangible but adjustment and balance between them can be a factor for change.

The practical consequence of this for STAR-FLOOD is that each consortium country could try to identify these principles, try to see what their legal status is as well as their place in the pyramid of the legal norms. Having a clear view on which principle seems to be dominant and which ones are being more and more rattled could give some good keys for understanding stability and change in legal frameworks. Moreover, comparison between each case study should be done while being aware of these differences.

3.11.3 Instability of the legal framework

Apart from these fundamental principles, a strong instability exists in legal frameworks, due to some legislative overproduction in most European countries. This procedural instability is due to a very high frequency of legal norms production. For instance, in France's urban planning legislation, since 2000, the frequency of legal norms has boomed and it is once or even twice a year that some new legal dispositions are adopted in this specific field. This instability proves to be very problematic for the stakeholders and the legal environment can become too fluctuant. Many actors, and not only in the field of flood management, find it too complex or even completely abstruse. This lack of readability tends to create a lack of transparency in public decision making.

This characteristic of legal norm production can help legal scholars and policy analysts to find a common ground for explaining change, as the legal framework is strongly influenced by exogenous factors such as shock events or societal evolutions. Rushing for the adoption of a new law in reaction to a major flood is probably something that every European country has already experienced. Another explanatory factor for legal instability is the disconnection between electoral mandates and a sensible frequency for legal norms production. Changing of government or of parliament representatives always sees a bunch of new rules going out as a buzzing effect and acting as smoke and mirrors. Cycles of legal production are too close to each other and operate on very short term. More generally, there has been a strong acceleration in the frequency of legal norms production over the past 10 to 20 years.

3.11.4 Dynamics in legal frameworks

This high level of inconstancy does not necessarily mean that major changes in the legal framework will occur automatically. Indeed, a distinction should be made between volatility/instability and change. Legal overproduction mostly reveals a procedural problem rather than a sensible evolution in the legal framework.

Yet, some major long term motions can be identified from a legal perspective and can be considered as deep changes starting to produce some effects on the legislative power and on the appreciation in tribunal decisions. For instance, changes such as the emergence of soft law in European legal frameworks, or the increasing obligation to evaluate legal norms, or the “contractualisation” phenomenon are all long-term trends that can be used to explain change. Moreover, some new principles tend to emerge such as the precautionary principle, transparency, information, public participation and more generally good administrative practices. A more in-depth analysis of these changes should be done as WP3 will advance. The description of the flood policy and regulation domain for each country could be a good occasion to evaluate the importance of these changes in each STARFLOOD country.

The constitutional structure of a country is another good example of a set of structural factors that can contribute to explaining stability or change. Although laws can be changed, some laws are easier to change than others. Constitutional norms are resistant to change, and when these norms relate to the way in which competences are distributed and which actors are expected to initiate change, they are even more so. The level of centralisation or decentralisation in a given jurisdiction will affect how change occurs as well. In decentralised countries legislation might be easier to change on a local level, when local circumstances dictate. In a centralised country it is possible that change in legislation is more time-consuming, and the threshold for change is higher, because the legislation processes at national level tend to be more complex, but when a change is made, it applies to the whole country and has a large influence.

3.11.5 Conclusion

Fundamental principles give a fairly stable and unmoving base to the legal framework by playing the role of reference for the rest of the legal norms. These principles are mostly shared by western European legal frameworks. But from a procedure point of view, legal frameworks seem to be very unsteady and unclear. The STARFLOOD consortium countries are all characterized by a very high level of law production frequency. As a conclusion we may advance the hypothesis that change occurs in legal frameworks according to different temporality and different type of cycles. Short term cycles show a very reactive side of the legal framework, strongly influenced by exogenous factors such as elections or shock events. These short term cycles can be identified and analysed within some more long term cycles of general trends such as the emergence of new fundamental principles.

4 Theoretical background to the evaluation framework

Sally Priest, Meghan Alexander, Colin Green, Anoeska Buijze, Willemijn Van Doorn-Hoekveld

4.1 Introduction

This chapter provides some further discussion and theoretical grounding of the concepts discussed in Chapter 5 of the theories and methodology report related to the evaluation framework. It is important to set out the purpose and rationale for the development of the STAR-FLOOD evaluation framework. It also provides some examples of how governance arrangements have been evaluated in previous projects. This document should be consulted in conjunction with Deliverables D1.1.1 and D1.1.2 from STAR-FLOOD work package 1 which identify and explain the characteristics of governance and in particular flood risk governance.

4.2 The purpose and logic of evaluation

4.2.1 Introduction

There is much literature discussing the processes and principles of evaluation (including Mark 2005; Cojocaru 2009; Rossi et al. 1999; Stufflebeam and Shrinkfield 2011; OECD 1991) which can inform and shape the evaluation framework adopted in STAR-FLOOD. We begin here by looking at some of these more general principles about evaluation, its key elements and its purpose before discussing its logic and determining those criteria for use in STAR-FLOOD. A useful starting point perhaps is to define evaluation. There are of course numerous potential definitions which could be included but a comprehensive one which includes both defining its aims and outcomes is provided by the OECD (1991) (Box 4.1).

Box 4.1: A definition of evaluation

“An evaluation is an assessment, as systematic and objective as possible, of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process.” OECD (1991; 4)

4.2.2 The purpose of evaluation

It is clearly essential to start by defining what we are trying to do in evaluation. In this project, it might be argued that we are seeking to identify what forms of governance are more appropriate in some circumstances than others. For instance, in terms of FRM, we may define inappropriate governance forms as: Those that exclude desirable forms of intervention either singularly or in combination; and/or are poor at delivering particular forms of intervention.

Examples of the first weakness are areas which do not provide a planned provision of emergency flood storage and, clearly, those that do not provide a multi-layered FRM strategy. Instance of the second weakness include development or building control in flood plain which is ineffective and a flood warning system which fails to reliably deliver valid flood warnings. These two criteria imply that a purpose of the evaluation criteria is to be diagnostic. In turn, diagnosis requires causal analysis when applied to *ex ante* evaluation. This purpose implies something about the evaluation criteria we may usefully adopt. But such criteria would require that we know in advance what the important diagnostic indicators are. But in the case studies, unless we measure the candidate

indicators it will not be possible to draw, at the end of the project, any reasoned conclusion as to their significance.

There is then a logical chain involved:

1. What are we trying to achieve?
2. What is the problem?
3. How do we measure the degree of achievement?
4. What are measurable indicators of the degree of achievement?
5. What are the variables that determine and predict the degree of achievement?

To be useful, in this project we need to be able to define the last set of variables in sufficient detail that stakeholders know how to apply them to deliver the appropriate framework of governance. For example, it is not sufficient to state that an accountable system of governance is required; it is necessary to state what actions will achieve this.

But the fifth stage cannot be derived from the fourth: a leap of imagination, critical thought and a review of the empirical literature are required. As examples of the last, Hooper (2005) sought to draw conclusions as to the variables which determine the success of integrated river basin management, and there have a number of other studies on this issue in the USA (Imperial and Hennessey 2000; Adler and Straube 2000; Sabatier et al. 2005) and elsewhere (Farrington et al 1999; Warren 1998). Hence, the fifth stage is one where it is necessary to state hypotheses as to what are the parameters which will make a difference in delivery.

4.2.3 A logic of evaluation

Evaluation is a means of testing the comparative performance of two or more alternatives. As a comparison, it may be made solely against each other or against some baseline (e.g. to contrast, in the first case, A is preferable to B; in the second case, neither A or B are satisfactory when considered against the baseline). The first requirements are therefore:

1. a set of objectives which it is desired to achieve;
2. a set of yardsticks, criteria, by which to measure the performance of each of the alternatives against each of those objectives.

This implies a degree of hierarchy, it being possible for two or more criteria to be required to be satisfied if a single objective is to be achieved. There then might be subordinate levels for each of those primary criteria. Table 4.1 identifies seven logical requirements or ideals for evaluation which might be implemented at each level.

Table 4.1 Seven ideal logical requirements for evaluation criteria

1. That the criteria are exhaustive and mutually exclusive (and thus to avoid double-counting).
2. That subsidiary criteria are necessary and sufficient conditions for satisfying the primary criterion. This is the concept of validity in measurement (Carmines and Zeller 1979).
3. Performance against each criterion can be measured at some level of measurement (categorical, ordinal, interval, ratio, absolute). Critically, measurement at the ratio level or above requires that it be possible to define an absolute zero point rather than to arbitrarily define some point as zero. This is critical because the mathematical operations that can meaningfully be performed depend upon the level of measurement and multiplication and division or the taking of a power are only meaningful if measurement is at a ratio scale or above. In many cases, measurement will only be possible at an interval level or below; this severely restricts the ways in which different measurements can be combined.

<p>4. Measurement at interval scale or above requires the definition of a zero point. In interval scale measurement, this is arbitrarily determined; achieving ratio or absolute scales of measurement requires that a zero can be defined unambiguously on theoretical grounds. A crucial distinction here is between the units of measurement and the measurement of effects. For example, in economic analyses, the units of measurement are ultimately money units where an unambiguous zero can be defined. But the baseline, the practical zero point, is more difficult to establish.</p>
<p>5. In turn, in any form of aggregation, the rules of dimensional analysis apply. Whilst all mathematical operations are permissible on numbers, the validity of those operations depends upon what the numbers represent in the particular case. The validity of applying mathematical operations depends upon the nature of units of measurement (e.g. in the second example, it is not valid to multiply length by mass and read the dependent variable as a length measurement).</p>
<p>6. This is also to require that measurement can be undertaken rigorously (e.g. with a high degree of inter-judge reliability (Carmines and Zeller 1979)): that is, if four judges are asked to categorise, rank or otherwise measure several alternatives against one criterion, there is a very high degree of agreement, consistency or reliability. When the argument for 'objective' measurement is made, this is argue that the first two points above can be achieved coupled to a claim that consistent measurement can be achieved at least at an interval scale of measurement. Secondly, that the variable used as a measure can be validated either against some physical variable or that there is a consensus as to the appropriate variable to use.</p>
<p>7. In order to test performance against an objective or the performance of some alternative against a higher level criterion by examining its performance against lower level criterion, some form of aggregation is required. This is to require the definition of some functional relationship be specified at each hierarchical level. What functional relationships are meaningful then depends upon the level of measurement achieved for each lower level criterion (and also on the units of measurement being such as to be consistent with the requirements of dimensional analysis). The alternative is to aggregate through pattern e.g. as in non-parametric Multiple Criteria Analysis (Green 2003).</p>

However, in the type of evaluation (required by STAR-FLOOD) such ideals are unlikely to be achievable for a number of reasons. Some of the objectives or criteria are likely to be antagonistic: one cannot be achieved except at the sacrifice of another. Sen (1992) noted that this is the case for equality. It is similarly commonly asserted that equity and efficiency are often antagonistic. It might be claimed that decisions only become difficult when there are at least two or more antagonistic objectives. Particularly in the case of objectives, there is commonly not a universal consensus either as to what should be objectives, how each is to be defined, or as to the relative importance that should be attached to achieving each. We can all agree on the importance of 'fairness'; the argument in each case is what constitutes fairness in that individual case.

Whilst the criteria should be mutually exclusive, the performance of one alternative against one criterion need not be independent of its performance against another. Conversely, one lower level criterion may contribute to the achievement of more than one higher level criterion. The inability to define one or more criteria by which to assess the achievement of a particular goal, or a rationale for aggregating those criteria appropriately to define achievement of that goal. The absence of reliable data with which to evaluate individual cases against a criterion. The latter two problems result in the need to use indirect or surrogate measures to estimate what in psychology is termed a 'latent variable' (Everitt 1984). For example, personality and intelligence tests rely upon the use of indirect measures to derive statistical estimates of what are considered to be inherently latent variables.

Consequently, these limitations are used present when selecting and applying criteria to be used for evaluation as they may impact on how criteria can be applied, the rigour of the output and how criteria are useful for comparative purposes.

The legal literature in particular differentiates between substantive and procedural tests; the former covers what must be done, essentially in terms of outcome, whilst the latter defines how the decision must be taken and the consequent actions implemented. The Water Framework Directive (Official Journal of the European Communities 2000b) and the Floods Directive (European Parliament and the Council 2007) to an even greater degree involve procedural requirements. Legal writers (e.g. Keesen et al. 2013) are critical of procedural requirements on the grounds that procedural correctness does not necessarily result in particular substantive outcomes. But it might be argued that where there are multiple and potentially conflicting objectives, and a large measure of differing conditions, a substantive criterion is potentially a Proscutean Bed. It is perhaps also relevant to note the difference between this legal framing and the economic framing; for example, the economist will argue that once we get the prices correct then whatever happens is the optimum. This is the basis of the economist's preference for economic instruments over command and control instruments which specify what is to be achieved or what is to be done. The economist does not look at the outcome but only questions whether the prices are those that would fall out of a perfectly competitive market. For example, a World Bank study visit to Germany concluded (Briscoe 1995) that there was less pollution and less leakage than would be efficient, those levels that would result if the expenditure on reducing each were reduced to the level of benefits resulting from reducing each. The economist will similarly question whether 'good ecological quality' justifies the resources required to achieve it in the particular local conditions. In turn, the economist is likely to argue that the issue of 'disproportionate' costs is misstated: the economic question is whether the benefits are more than proportionate to the costs, the standard benefit-cost ratio test. If the benefits cannot be confidently shown to be greater than the costs, then the costs are disproportionate and the action should not be undertaken. Given the scarcity of resources taking action when the benefits are less than the costs will squeeze out other action where the benefits are greater than the costs. For example, action to reduce air pollution will be precluded because the resources have been expended on reducing water pollution.

In case law, there have been notable attempts to reconcile economic framing with the substantive approach of law; Posner (1995) being the lead example. Again, in the English law of the tort of negligence, there has developed a causal approach. Hart and Honore (1990) concluded that three grounds for imposing responsibility for some consequence may be argued:

1. Is someone responsible because of their specific actions or irrespective of their actions?
2. Must it be shown that their actions resulted in the harm?
3. Are they responsible for the consequences of their actions only if they were at fault or irrespective of the nature of, or reasons for, their conduct?

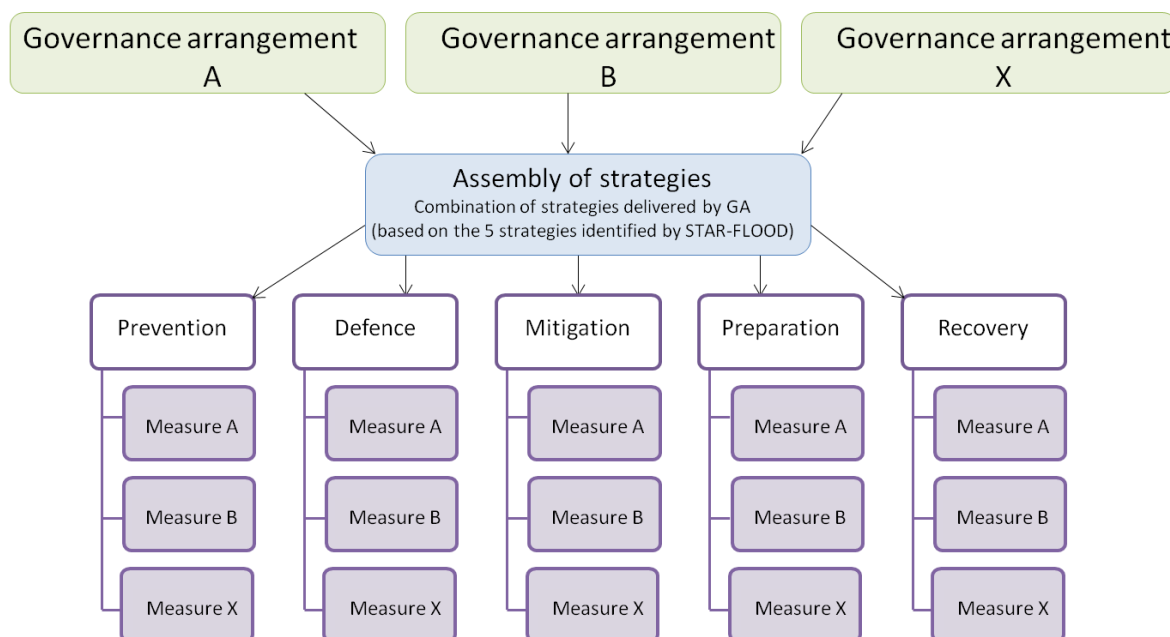
The combinations of these three grounds generate at least seven possible grounds for assigning liability to the defendant.

4.3 Evaluating governance

4.3.1 The object of evaluation

Whilst distinctions are made between measures, strategies and the assembly of strategies, it is the overarching governance arrangement(s) that constitutes our primary object of analysis and evaluation (Figure 4.1).

Figure 4.1: Clarifying the object of evaluation: Distinguishing governance arrangements, strategies and measures (where “X” indicates a number of possibilities). Note that different arrangements may be linked to different strategies and not all strategies are necessarily implemented.



Case study researchers should regard the measures and FRMSs as an entry point towards understanding the overarching governance arrangements. The evaluation process needs to examine the measures, the strategies and the assembly of strategies as a process of understanding how governance arrangements are both implemented and function in practice and where there are multiple governance arrangements how they intersect. But the actual object of evaluation is the arrangement itself. We are therefore aiming to answer the following questions;

- Is the FRGA(s) resilient? (i.e. Does it enable a broadening of FRMSs? Does it provide scope for learning and innovation?)
- Does the FRGA(s) deliver strategies which enhance the flood resilience of the area?
- Is the FRGA appropriate? (i.e. Does it deliver an assembly of FRMS which are effective, legitimate and efficient?)

To support evaluation, each criteria and sub-criteria are conceptualized according to **process** and **outcomes**, and further distinguished in Table 4.2. This distinction is justified on a pragmatic level and emphasises the importance of being holistic in our approach. It also to some extent maps onto the PAA framework discussed in Chapter 3.

Table 4.2: Distinguishing stages for which evaluation criteria can be assessed

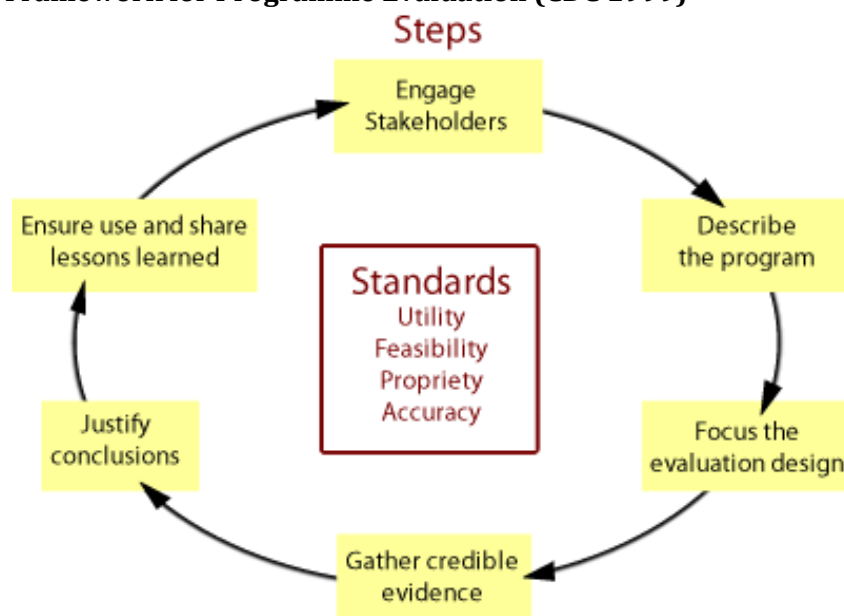
Stages at which criteria can be assessed:	
<i>Process</i>	Input – use of resources and actor involvement in the decision making process
	Output – Result of the decision making process; also involving consideration of alternative courses of action
<i>Outcomes</i>	Outcome – Implementation of output (e.g. has a regulation been implemented)
	Impact – Resulting effect of decision making process and outcome

4.3.2 Previous examples of evaluating governance

There are of course many examples of evaluation in the literature and in particular numerous texts on policy and programme evaluation as well as on project appraisal. This section will provide some examples of evaluating governance from this literature particularly focussing on examples from the development and environmental literature and where possible from flood risk management and water governance.

Examples in the literature often involve the appraisal of institutions, development programmes and measuring the impact that interventions have made. Examples of this nature include OECD (1999); OECD (2011); PAIB (2009); Garcia (2011); UNDP (2004). An influential model of the *ex post* evaluation process (Figure 4.2) is that adopted for the evaluation of health intervention programmes in the USA (CDC 1999). This provides a clear sequential approach for evaluating these types of programmes and the standards against which they should be measured. The OECD has developed a series of publications (e.g. OECD 2001; 2010; 2013) as part of their Development Assistance Committee (DAC) Network of Development Evaluation which provide both methods and criteria for governance evaluation. They suggest that the following five criteria provide a framework for evaluating programmes: Relevance, Effectiveness, Efficiency, Impact and Sustainability (OECD 2010).

Figure 4.2 A Framework for Programme Evaluation (CDC 1999)



Abrams et al. (2003) have adopted and revised both the UNDP (1997) criteria and those developed by the Institute of Governance (2002). Although their methodology for evaluating governance relates to the activities of NGOs in protected areas some of the processes they have developed are equally pertinent in other situations and have relevance for evaluation in STAR-FLOOD. They suggest that there are five principles of good governance: Legitimacy and Voice, Accountability, Performance, Fairness and Direction (Table 4.3). Many of these criteria relate to issues of the process in which decision-making is undertaken and contribute to an overarching notion of answering a question about whether a process is fair and legitimate. Although many differing concepts are presented and potentially overlap, STAR-FLOOD researchers might utilise Abram et al.'s criteria to inform qualitative and/or quantitative evaluation.

Table 4.3 Criteria to evaluate Abrams et al.'s (2003) principles of good governance

LEGITIMACY AND VOICE
<p>Representation of Interests and Concerns. The governing bodies accurately and reliably represent, directly or indirectly, the interests and concerns of all relevant actors.</p> <p>Public Participation. Strong participation in numbers and contributions are obtained in relevant consultations and decision-making forums at various levels (legislation, system of PAs, individual PA).</p> <p>Commitment to Multi-party Processes. The governing bodies provide the support necessary to build and maintain strong multi-party processes of consultation and decision-making. The free expression of views is promoted, with no discrimination related to gender, ethnicity, social class, etc.</p> <p>Subsidiarity. Decisions are taken at the lowest level compatible with relevant capacities.</p> <p>Checks and balances. Civil society groups and an independent media act as a check and balance on the exercise of the powers granted to PA political leaders and managers.</p> <p>Support for Organizational Capacity of Various Relevant Actors. The governing bodies provide assistance to various relevant actors (e.g. local indigenous communities, NGOs) to develop their organizational capacity. This allows them to better represent their interests, participate in collaborative processes, and engage in activities.</p> <p>Variety of Institutions. Various types of PA institutional settings are recognised as legitimate. In particular, Community Conserved Areas are recognised and supported but not forced to conform to either “conservation” or “development” as defined by non-local people.</p> <p>Responsiveness to Power Sharing. The governing bodies demonstrate responsiveness to new ideas and institutional arrangements that explore constructive forms of sharing their governing powers.</p> <p>Effective Consensus Processes. The relevant actors in the governing bodies are able to entertain an effective dialogue among themselves, arrive at mutually satisfactory decisions and prevent the process from producing only “lowest common denominator” agreements (vague, generic decisions).</p> <p>Credibility. Governing bodies honour internal and external commitments (e.g. following strategic plans, addressing issues of emerging importance, aligning practices with stated values). Their decision-making processes inspire confidence because they are unbiased, fair and open.</p> <p>Continuity in Membership of the Governing Bodies. Members of governing bodies are maintained for long periods of time and there is a smooth hand-over process when replacements occur. Stability and consistency in representation occurs, allowing for the building of trust and collaboration amongst the members themselves, their constituencies and other relevant actors.</p> <p>Ownership of institution. The management rules are respected because they are “owned” by the relevant actors and not solely because of fear of repression.</p>
ACCOUNTABILITY
<p>Appropriate Roles and Responsibilities. The roles and responsibilities held by the governing bodies’ and other relevant social actors are within the scope of their influence and ability to carry them out.</p> <p>Clarity of Roles and Responsibilities. There is a clear identification and assignment of roles, authority, responsibility, rights, rules and accountability in all aspects of PA management. Clarity is critical in being able to answer the questions “who is entitled to what under which conditions?” and “who is accountable to whom for what?”</p> <p>Effective Reporting System. Effective mechanisms exist and are used to provide relevant PA information to the public at large and, in particular, to the most</p>

<p>directly concerned social actors.</p> <p>Guaranteed Access to Information. There is guaranteed access to information that is adequate in terms of quantity, quality and completeness regarding the governing bodies, the management process and results and the accountability of each decision/ result. Information is guaranteed as a right of citizenship.</p> <p>Public Concern and Demand. The civil society (e.g., communities, NGOs, unions, associations, business groups, leaders and individuals at large) is active (e.g. through the media and the legal system) and effective in obtaining the accountability of the governing bodies.</p> <p>Performance Evaluation. There is an effective, on-going evaluation of the PA governance, fostering improved performance and information sharing.</p> <p>Independent Accountability Institutions. There are independent public institutions of accountability with the authority and capacity to oversee and question the actions of the governing bodies. Examples of such institutions are the legislature, the judiciary, auditing agencies, ombudsperson, and human rights commissions.</p> <p>Rewards and Sanctions. Accountability is reflected, as appropriate, into concrete and appropriate rewards and sanctions</p>
PERFORMANCE
<p>Capacity. The governing bodies have sufficient human, technical and financial capacity to carry out their required roles, responsibilities and accountability over time.</p> <p>Programme Design. The governing bodies establish and maintain a balanced and responsive overall design to structure their own activities and those of other PA actors. They are capable of embracing and merging conservation and other objectives (e.g., supporting sustainable livelihoods, strengthening cultural identity, and satisfying spiritual needs).</p> <p>Co-ordination. PA governing bodies effectively co-ordinate with other governing bodies, technical bodies and relevant actors, in particular at ecosystem/landscape and regional levels.</p> <p>Cost Effectiveness and Efficiency. The governing bodies are cost effective and efficient in achieving the PA management objectives, on the basis of a sound planning and implementation system that prevents damaging delays.</p> <p>Attainment of Management Objectives. The governing bodies are able to demonstrate progress towards the PA management objectives and/or attained performance. [Protected area objectives may include protection of habitat, nature interpretation and outreach, protection of cultural resources, maintenance of ecosystem functions, etc.]</p> <p>Performance Information to the Public. Governing bodies provide sufficient and timely information to allow assessments of their performance by interested parties and the public.</p> <p>Responsiveness. Governing bodies are responsive to complaints and public criticism of their activities. Appropriate changes are made to meet expressed needs.</p> <p>Internal Evaluations. Governing bodies are capable to undertake internal programme evaluations and respond to their own findings.</p> <p>Robustness and Resilience. Governing bodies identify key potential threats facing them (e.g. funding shortfalls, legal attacks, political sabotage, changes in national protected area policies and leaders), successfully cope with them and learn from the experience.</p> <p>Advocacy and Outreach. Governing bodies successfully inform the public and interested parties about their own functions and roles and are able to influence decision-making processes in other components of a country's governing system.</p> <p>Policy Learning. Lessons learned from the experience of the specific PA are successfully fed back into the policy of governing bodies.</p>
FAIRNESS

Decency. Governing bodies make sure that management activities and conservation in general are undertaken with decency, without humiliating or harming people.

Impartial Enforcement of Rules. The rules that restrict the use of PA's territory and resources are transparent and sanctions over infractions can be appealed.

Shared Decision Making. Mechanisms for sharing relevant PA decision-making with local and indigenous people exist and are appropriately utilised.

Fair Management of Conflicts and Past Injustices. Governing bodies provide fair avenues for conflict management and, eventually, non-discriminatory recourse to justice. They take full advantage of a variety of formal and informal means to promote fair and effective dialogue and development of agreements among conflicting parties. **Opportunities.** Governing bodies provide support to conservation and development initiatives benefiting stakeholders.

Non-discrimination and Consistency. Governing bodies do not discriminate against any group on the basis of gender, religion, ethnicity, social status, political affiliation or other, with respect to the membership of policy and decision-making bodies or other PA-relevant functions. PA rules are applied consistently.

Distributional Equity. Governing instruments (e.g. PA management policies, rules, conflict resolution mechanisms, funding opportunities, etc.) are used in an impartial fashion to distribute fairly the costs and benefits generated by the protected area.

Rule of Law. Governing bodies hear cases and enforce sanctions for all potential infractions according to established rules and regulations, without any form of discrimination and consistently through time. These laws are transparent, enforced fairly and there is a right to appeal for the transgressors.

Integrity in of Staff Management. PA Staff receive positive or negative rewards in fair proportion to the results of their performance (merit-based).

DIRECTION

Leadership. The governing bodies generate new ideas and launch innovative processes (e.g., cultural sensitivity training, conservation covenants) to address and resolve difficult issues. They support other social actors engaged in innovative, promising work.

Collaborative Learning. The governing bodies provide or support initiatives to increase and improve the use of collaborative learning in various forums: policy and decision-making, conflict resolution, etc.

Policy Direction. The governing bodies provide clear policy directions for the main issues of concern to the protected area e.g., sustainable commercial development, use of local ecological knowledge, gender balance in programmes, public participation in decision-making, etc.

Guiding Values. There is a set of agreed-upon values that guide the protected area's governing bodies' processes and activities.

Vision. The governing bodies provide an inspiring vision of the protected area's future based on values shared by its main relevant actors and society at large. They are able to mobilise support for that vision.

Consistency between Values and Practice. The protected area's governing processes, objectives and plans are clear, viable and respect the values collectively agreed upon.

Shared Perspective on Good Governance. There is a broad, shared perspective between the governing bodies and concerned social actors on what is needed to create a system of good governance.

Consistency with International Obligations. The PA directives are consistent with their international obligations, such as the Convention of Biological Diversity, the World Heritage Convention, the Ramsar Convention, the ILO rulings on Indigenous Peoples, etc.

Abrams et al. (2003; 34-37).

A number of authors have proposed criteria specifically for environmental governance (e.g. Cadman 2012; Lockwood et al. 2010; Thomas 2008) and additionally criteria were proposed for water management (e.g. Rogers and Hall 2002; Moench et al. 2003; Menard and Saleth 2011; Pena and Solanes 2003; UNDP 2004). It is not possible to discuss all of the criteria presented in this short document, however some of the more useful and common criteria for evaluating governance are discussed. For example, Lockwood et al. (2010) proposed eight principles for natural resource governance as highlighted in Table 4.4. There is a great deal of overlap between Abrams et al.'s (2003) criteria presented above and those presented here by Lockwood et al. (2010). There is again a great deal of emphasis placed on the process and its legitimacy. This reinforces the importance of STAR-FLOOD investigations of legitimacy within an evaluation framework.

Table 4.4 Lockwood et al.'s eight principles for governance in natural resource management

Governance principle	Explanation
Legitimacy	Legitimacy refers to (a) the validity of an organization's authority to govern that may be (i) conferred by democratic statute; or (ii) earned through the acceptance by stakeholders of an organization's authority to govern; (b) that power being devolved to the lowest level at which it can be effectively exercised; and (c) the integrity with which this authority is exercised.
Transparency	Transparency refers to (a) the visibility of decision-making processes; (b) the clarity with which the reasoning behind decisions is communicated; and (c) the ready availability of relevant information about governance and performance in an organization.
Accountability	Accountability refers to (a) the allocation and acceptance of responsibility for decisions and actions and (b) the demonstration of whether and how these responsibilities have been met.
Inclusiveness	Inclusiveness refers to opportunities available for stakeholders to participate in and influence decision-making processes and actions. Governance is regarded as inclusive when all those with a stake in governance processes can engage with them on a basis equal to that provided to all other stakeholders.
Fairness	Fairness refers to (a) the respect and attention given to stakeholders' views; (b) consistency and absence of personal bias in decision making; and (c) the consideration given to distribution of costs and benefits of decisions.
Integration	Integration refers to (a) the connection between, and coordination across, different governance levels; (b) the connection between, and coordination across, organizations at the same level of governance; and (c) the alignment of priorities, plans, and activities across governance organizations.
Capability	Capability refers to the systems, plans, resources, skills, leadership, knowledge, and experiences that enable organizations, and the individuals who direct, manage, and work for them, to effectively deliver on their responsibilities.
Adaptability	Adaptability refers to (a) the incorporation of new knowledge and learning into decision making and implementation; (b) anticipation and management of threats, opportunities, and associated risks; and (c) systematic reflection on individual, organizational, and system performance.

After Lockwood et al. (2010; p991-997)

Expanding on this discussion, Cadman (2012) identifies two criteria which affect global governance (Meaningful participation and Productive deliberation) and four criteria (Interest representation, Accountability and transparency, Decision-making, and Implementation) which can be used to

measure them. Table 4.5 illustrates Cadman’s (2012) key principles, how they interact with the criteria and also some indicators for evaluation. Similar, to the above STAR-FLOOD researchers may choose to adopt the indicators presented within their evaluation. Cadman (2012; 24) goes on to operationalise this and presents an institutional model of the quality of global governance and suggests adopting a scalar (high, medium and low) scoring system to evaluate, compare and contrast governance systems. In their study on water governance Rogers and Hall (2002) agree with many of the described principles for effective water governance described above. They divide the principles into the Approaches (which should be Open and transparent, Inclusive and communicative, Coherent and integrative and Equitable and ethical) and Performance and operation (which should be Accountable, Efficient and Responsible and sustainable).

Table 4.5 Hierarchical framework for the assessment of governance quality (Cadman 2012: 21)

Principle	Criterion	Indicator
Meaningful Participation	Interest representation	Inclusiveness
		Equality
		Resources
	Organisational responsibility	Accountability
		Transparency
Productive deliberation	Decision making	Democracy
		Agreement
		Dispute settlement
	Implementation	Behavioural change
		Problem solving
		Durability

Whereas the previous examples have centred on legitimacy criteria, Ostrom (2006; 10) has provided a series of evaluation criteria for measuring the outcomes that are being achieved under current institutional arrangements (Table 4.6). These concepts take a much broader view and might be used to both measure whether the particular outcomes have been achieved as well as the process of reaching the outcomes. In particular, the author introduces additional criteria to those presented thus far. She places emphasis on the efficiency and the ability of the institutional arrangements to learn and adapt in the face of change (i.e. which might be considered to be linked to resilience; discussed in later sections).

Table 4.6 Evaluation criteria for social and institutional systems (adapted from Ostrom 2006: 10)

Economic Efficiency	Economic efficiency is determined by the magnitude of the change in the flow of net benefits associated with an allocation or reallocation of resources. The concept of efficiency plays a central role in studies estimating the benefits and costs or rates of return to investments, which are often used to determine the economic feasibility or desirability of public policies. When considering alternative institutional arrangements, therefore, it is crucial to consider how revisions in the rules affecting participants will alter behaviour and hence the allocation of resources.
---------------------	--

Fiscal Equivalence	There are two principal means of assessing equity: (1) on the basis of the equality between individuals' contributions to an effort and the benefits they derive and (2) on the basis of differential abilities to pay. The concept of equity that underlies an exchange economy holds that those who benefit from a service should bear the burden of financing that service. Perceptions of fiscal equivalence or a lack thereof can affect the willingness of individuals to contribute toward the development and maintenance of resource systems.
Redistributional Equity	Policies that redistribute resources to poorer individuals are of considerable importance. Thus, although efficiency would dictate that scarce resources be used where they produce the greatest net benefit, equity goals may temper this objective, and the result is the provision of facilities that benefit particularly needy groups. Likewise, redistributional objectives may conflict with the goal of achieving fiscal equivalence.
Accountability	In a democratic polity, officials should be accountable to citizens concerning the development and use of public facilities and natural resources. Concern for accountability need not conflict greatly with efficiency and equity goals. Indeed, achieving efficiency requires that information about the preferences of citizens be available to decision makers, as does achieving accountability. Institutional arrangements that effectively aggregate this information assist in realizing efficiency at the same time that they serve to increase accountability and to promote the achievement of redistributional objectives.
Conformance to General Morality	In addition to accountability, one may wish to evaluate the level of general morality fostered by a particular set of institutional arrangements. Are those who are able to cheat and go undetected able to obtain very high payoffs? Are those who keep promises more likely to be rewarded and advanced in their careers? How do those who repeatedly interact within a set of institutional arrangements learn to relate to one another over the long term?
Adaptability	Finally, unless institutional arrangements are able to respond to ever-changing environments, the sustainability of resources and investments is likely to suffer. Rural areas of developing countries are often faced with natural disasters and highly localized special circumstances. If an institutional arrangement is too inflexible to cope with these unique conditions, it is unlikely to prosper. For example, if an irrigation system is centrally controlled and allocates only a specific amount of resources to annual and periodic maintenance; it may not be able to meet the special needs associated with a major flood that destroys a section of the canal system.

Also of interest to STAR-FLOOD researchers is the work by Newig and Fritsch (2009) who provide an exploration of multi-level environmental governance and whether the existence of multiple levels of governance (which are also likely to exist when managing flood risk) hampers the ability to deliver high quality policy through participatory decision-making. They have undertaken a review of a large number of case study examples and adopt a coding system to test eleven hypotheses arranged under five overarching criteria headings (Table 4.7). These provide some interesting example of hypotheses that STAR-FLOOD researchers might want to consider and adapt when evaluating governance in their case study areas. Of particular importance and relevance from Newig and Fritsch's (2009) findings was that many concepts concerning governance which are currently articulated in literature fail to take full account of context and further analysis of the case studies in this research highlighted that "the underlying causal relationships were different or more complex than those predicted by theory" (p210). Both of these issues provide powerful insights for STAR-FLOOD researchers when designing the case study research processes.

Table 4.7 Hypotheses developed by Newig and Fritsch (2009) to investigate the complex relationships between participatory, multi-level and scale-adapted governance and the effectiveness of environmental policy.

Criteria	Hypotheses
Participatory Versus Top-Down Modes of Governance	H1a. Participation of non-state actors leads to more ecologically rational decisions than top-down modes of governance. H1b. Participation of non-state actors leads to improved compliance with decisions and thus better outcomes and impacts in ecological terms than top-down modes of governance.
Spatial Relevance of Actor Interests	H2a. Citizens living in close spatial proximity to a natural resource tend to favour its economic exploitation, whereas those living farther away tend to favour its conservation. H2b. Where decision competences regarding environmental issues are on lower levels of governance, a stronger and more influential participation of citizens with economic interests can be expected.
Local Scale Versus Higher Scale Decision-Making	H3a. The environmental effectiveness of decisions correlates <i>positively</i> with the scale of the governance unit. H3b. The environmental effectiveness of decisions correlates <i>negatively</i> with the scale of the governance unit.
Spatial Fit Between Governance Scales and Natural Scales	H4a. Governance of natural resources on natural scales leads to more ecologically rational outcomes than governance on territorial scales. H4b. Governance of natural resources on natural scales leads to less ecologically rational outcomes than governance on territorial scales. H4c. Participation improves the fit between natural and governance scales and thus improves environmental outcomes.
Polycentricity of the Whole (Multi-Level) Governance System	H5a. The more levels and actors involved in a policy implementation process the lower its effectiveness. H5b. A large number of horizontal and vertical, quasi-autonomous decision points is better able to adapt to external change than hierarchical modes of governance, leading to more sustainable resource use.

Another example from the environmental governance literature is that provided by Falaleeva and Rauschmayer (2013) who undertook an outcome and process evaluation of a biodiversity project in Belarus. The outcome part of the analysis related to whether the specific goals of the project (such as the expansion of protection areas and further scientific research) were met. The process evaluation is perhaps more relevant to the process in STAR-FLOOD and examined the following variables: the integration of knowledge and information, how legitimacy was supported, how social dynamics were promoted and the cost-effectiveness of the scheme. The author illustrated three different evaluation models which might be used to analyse both the processes and outcomes of the project each reflecting differing perspectives. The first approach involved the international organisations and local partners evaluating the project from their own viewpoints and standards. The second approach involved the different groups partially sharing evaluation standards with the final approach involving the full integration and the adoption of a common evaluation model. Falaleeva and Rauschmayer (2013) argued that the second model manifested over the longer term in this case study; however the failure to fully integrate into a common approach contributes to the limited long-term sustainability of investment.

Many of the environmental examples related to governance however relate to the evaluation of strategies, rather than governance arrangements *per se* and therefore these examples may be of use for undertaking *ex post* evaluation of the FRGS, but are of more limited interest to a more overarching *ex ante* evaluation of FRGAs. For instance, Paneque Salgado et al. (2009) undertook a

participative multi-criteria evaluation of urban water supplies in a Spanish case study. This example highlights the value of the engagement of social actors within an evaluation process, however like many other examples of environmental evaluation focuses primarily on evaluating different options, strategies of policy interventions, rather than overall governance arrangements. Therefore, although we are able to learn from the techniques and criteria presented for evaluating FRMSs their applicability and use for evaluating governance is limited. The authors have highlighted the difficulty in establishing rigid quantitative criteria which are also sufficiently meaningful to enable the comparison on alternative options. Of the eleven criteria (including social, environmental, economic and institutional) that were selected, perhaps of most interest to the STAR-FLOOD project are the *Degree of social acceptability* and *Degree of institutional difficulty*. In this instance the former highlighted the issue of interscalability and the fact that the problem was perceived in very different way by the different groups. The latter issue enabled the practical constraints in dealing with a range of different alternatives to be considered.

The practical problems with all such criteria and one that must be confronted in STARFLOOD are:

1. Developing criteria that can actually be tested;
2. Identifying the means by which each can be satisfied.

Otherwise criteria are no more than invocations of motherhood and apple pie with the practical problem that all are likely to be contested. The concept of fairness is an obvious example; small children learn early that a complaint that 'it isn't fair' is more likely to be attended to than one of 'but I want it'. It is much easier to state a criterion than to actually specify what this criterion means in practice. However until that can be done, the proposed criterion is essentially meaningless. The following section explores some of the methodological issues associated with adopting criteria for evaluation.

4.3.3 Methodological reflections for evaluating governance in STAR-FLOOD: the challenge of using indicators

Before describing specific criteria in more detail and after establishing both the purpose and logic of evaluation and examples of previous evaluation, this section will explore in more detail methods for evaluation and in particular consider the potential role of indicators for evaluating flood risk management governance. It is important to consider carefully those data collection and interpretation methods which should be used to evaluate (both *ex post* and *ex ante*) FRGAs alongside the selection of those criteria for evaluation.

Arguably in an ideal situation it would be possible to define and utilise rigid indicators from which to measure the different evaluation criteria. Indicators may be defined in various ways: a useful definition however is "a qualitative or quantitative parameter characterising the current condition of an element of the environment or its change over time" (Aubry and Elliot 2006; 175). The authors go on to suggest that indicators should have three basic purposes which are:

- To simplify;
- To quantify;
- To communicate.

Indicators in this context are considered in a hierarchical way and we distinguish between the overarching evaluation criteria which we are using to provide the evaluation and indicators which may be used to measure or categorise them. Therefore, ideally each evaluation criterion may have more than one indicator. Abrams et al. (2003) have identified the characteristics of good indicators (Box 4.2) which they have adopted from a range of other literature (including Centre for Coastal Management 1993; Briggs et al. 1996; Abbot and Guijt 1998; Margoluis and Salafsky 1998).

Box 4.2: Characteristics of good indicators (Abrams et al. 2003: 38)

A good indicator is:

- **Significant.** Reflects changes or aspects of importance at meaningful spatial and temporal scales;
- **Sensitive.** Changes proportionately in response to actual changes in the condition or item being measured;
- **Measurable.** Can be recorded and analyzed in quantitative or qualitative terms;
- **Precise.** Defined the same way by all people;
- **Simple and measurable at low cost.** Easy to measure and cost-effective in terms of data collection, analysis and interpretation;
- **Practical.** Can be collected, analyzed and reported on in a timely fashion;
- **Comparable.** Defined and measured in a way that allows it to be compared to and combined with other indicators (e.g. percent achievement of a standard).

The use of quantitative indicators has the benefit of introducing some degree of rigour to a comparison between partner countries. However, a key importance emerging from case study examples (e.g. Paneque Salgado et al. 2009) is the need to encourage iteration into any evaluation and adopt an approach which has learning at the centre. This process may be difficult if only indicators are used as Conway (2007) argues indicators tend to oversimplify. Additionally, not all parameters are able to be quantified and more importantly using indicators (particularly quantitative ones) may not provide us with the information that we need to better understand governance arrangements. This is particularly important when considering the inherent complexity in understanding the relationships between the social, economic and environmental systems integrated within flood risk management. Consequently, as McFadden et al. (2008: 2) suggest that in a social science context “knowledge is mediated, situated, incomplete and contested” and adopting methodological approaches which recognise these characteristics is essential. Therefore, to evaluate a FRGA, including how it has performed, the opportunities and barriers it has afforded and how it might be transferred elsewhere, there is the need to adopt a range of methodologies for evaluation. Examples of different methods of qualitative data collection are discussed briefly in Chapter 6.

It is particularly necessary to consider evaluation as part of the data collection process. Abrams et al. (2003) provide an example of how both quantitative and qualitative indicators may be used to evaluate different governance criteria and the types of information that may be necessary to collect (Table 4.8). This approach includes both quantitative information (e.g. the number of participants or the percentage of meetings) and also qualitative interpretation (e.g. the level of stakeholder knowledge, the role that stakeholders play).

Table 4.8 Example of a framework for organizing agreed governance principles, criteria and indicators (Abrams et al. 2003; 55)

Principle	Criteria	Indicators	Source of Information
1. Participation and consensus orientation	1.1 Fair representation of interests and concerns	1.1a Number of different social actors with membership on the governing bodies	Legal framework or document indicating membership.
		1.1.b Existence and use of mechanisms that promote power sharing in governing bodies' meetings (e.g. a rotating chair)	Internal regulations of governing body and tracking record of meetings, minutes.
		1.1.c Percentage of meetings of the governing bodies for which the agenda was developed collaboratively with various concerned actors	Tracking record of meetings, minutes of meetings.
		1.1.d Existence and use of mechanism(s) by governing bodies for gathering direct information from stakeholders (especially grassroots and marginalized groups) about their own interests and concerns	Internal regulations, documents recording consultation exercises, grass-roots proposals, tracking record, minutes etc.
	1.2 Public participation	1.2.a Number and range of participants at PA consultation and decision-making forums	Minutes, attendance sheets and tracking record.
		1.2.b Level of stakeholder knowledge of consensus processes convened by governing bodies (e.g. issues to be discussed, specific agreements reached)	Participatory exercise with focus group
		1.2.c Number and types of protected area operations in which local stakeholders play an active role (e.g., as salaried staff, key advisors, evaluators)	Documents and minutes of meetings for quantitative data; Venn Diagram with focus groups for perceptions.

A participatory approach to evaluation is required as much of the richness and value in understanding governance arrangements will emerge directly from those involved in making flood risk management decisions and also those affected by their outcomes. As Abrams et al. (2003: 6) argue that the growing recognition of the "importance of, and value placed on the rights, responsibilities and interests of stakeholders" within governance models is another rationale for including stakeholders within the evaluation of governance arrangements. Therefore, evaluation should not only be undertaken by researchers as a process separated from data collection, but where possible be integrated into that process. O' Conner (2006) contends that this learning is the most valuable component of analysis and evaluation and "it is the inter-subjective process of argument and dialogue with its affective as well as informative dimensions that engenders new insights (learning) and, more particularly, builds (or undoes and rebuilds) alliances, modifies motivations and thus permits the exploration of contradictions and emergence of new solidarities".

The design of the NFPR and case study protocol in STAR-FLOOD allows for an iterative approach to both data collection and evaluation. Studying both the context of FRGAs and implementation at the

NFPR level first will allow the setting of hypotheses and the development of a data collection approach at the case study level which will integrate evaluation concepts and enable these to be explored with those implementing Flood Risk Management or those affected by it.

4.3.4 Examining criteria for evaluation: What are we trying to evaluate and why?

Our target is to evaluate FRGAs; however FRGAs can also be seen as a linking mechanism. In this regard, the function of a governance arrangement(s) is to produce a FRMS that is appropriate to local conditions. In doing so, it has to respond to the nature of that problem. We sought in WP1 to define what constitutes an appropriate FRMS and notably we argued that an appropriate FRMS was likely to be multi-layered. To be able to deliver an appropriate FRMS (or assembly of strategies), the FRGA has to reflect the familiar problem of fit: do the rules, the boundaries created by those rules, the powers available actually enable the appropriate FRMS to be delivered? So we may ask why France has a functioning system for flood insurance but why this is not the case elsewhere.

But the FRGA also has to address the overarching objectives of enhancing well-being in the context of delivering sustainable development. To be effective, it has also to be successful in the way it approaches the problems that governance must address as those are defined in the UNDP (1997) definition: “... *through which citizens and groups, articulate their interests, exercise their legal rights, meet their obligations and mediate their differences.*” In particular, it has to resolve conflicts in a way which is deemed by society to be acceptable. Implicit within the UNDP definition is that co-acting is an essential part of governance. Hence, the appropriate governance arrangements involve the incentives for co-acting and the mechanisms for doing so. These include what have been called ‘bridging’ mechanisms. Whether the individual groups then have the skills to do so may be argued to be a distinct problem.

Confronted by change and simultaneously by the need to shift to sustainable development, a practical requirement for the governance arrangements is that they promote learning and innovation and have an adaptive capacity (resilience). In turn, change creates uncertainty which is a condition STAR-FLOOD researchers should consider when evaluating flood risk governance.

In addition, in the same way that decisions are ultimately tested by the actions they produce, institutions, are commonly manifested in the form of organisations. So, in North’s (1990) distinction between institutions and organisations, between the rules of the game and the players, a failure in governance might be the consequence of a failure of the system of rules to be appropriate or the failure of the organisation operating within those rules. Barcelona is not more successful than Accrington Stanley football club at football because they play different rules but because of differing skills in the team and its organisation. Hence, at some point the problem becomes how to design an appropriate organisation(s) and what skills or capacities that organisation requires to fulfil its role. Ostrom’s (2000) criteria for the design of institutions need to be complemented by those for the design of organisations. For example, there is an on-going debate in the Netherlands about the redesign of local government which is concerned with such questions as economies of scale and scope in organisations versus the issue of local identity. In turn, it must be remembered that governance is the product of people interacting through different forms of symbolic systems. An organisation is one such system of people interacting and its internal rule framework is intended to maximise the productivity of these interactions.

Since we have adopted the PAA framework, the evaluation criteria need to both tie back to the PAA framework and the PAA needs to provide some explanation in terms of the evaluation differences between different FRGAs. The PAA provides an exceptionally powerful analytical framework but its role is equivalent to a textbook on anatomy for a doctor. It may be noted that anatomical textbooks largely drew upon dissections of the dead; the problem for the doctor is to reach a diagnosis and to identify a therapy for the living. Anatomical knowledge is an essential precondition for surgery:

removing and repairing. However, many therapies depend upon the use of pharmaceuticals such as insulin and antibiotics. Similarly, surgery in the form of the replacement of hip joints and the fitting of heart pacemakers require additional knowledge beyond that of anatomy. So, the PAA framework may be anticipated to require supplementation by other tools.

4.4 Introducing evaluation criteria

As discussed in previous sections, there is a need to select criteria to guide the evaluation of FRGAs and facilitate comparisons across partner countries. Many potential criteria have emerged from the reviewed literature discussed thus far. In informing the evaluation framework for STAR-FLOOD we need to exercise caution in being on the one hand thorough; but also developing a procedure that is achievable operationally. From the literature, we have selected two broad criteria to steer evaluation. These encompass many of the elements/indicators/concepts previously mentioned. STAR-FLOOD evaluation criteria are therefore:

1. Resilience;
2. Appropriateness – including the sub-criteria of effectiveness, efficiency and legitimacy.

The following sections review these in turn, providing theoretical discussion and justification for the selection of these criteria.

4.4.1 A theoretical overview of resilience

Resilience is one of the selected criteria for evaluating FRGAs and is considered on an empirical level in Chapter 5. The purpose of this section is to provide a theoretical context for resilience. Here we observe a number of theoretical fracture lines, revealing the contested nature of resilience and remark on a number of “take home messages” that can be applied to our evaluation goals in STAR-FLOOD.

What is resilience?

Resilience can be thought of as the desired state of key systems (social, economic, ecological etc.). Although widely applied, the term ‘resilience’ is contested between different disciplinary and epistemological positions, such that there is no agreed definition of its meaning. The concept originates from ecology and the influential work of Holling (1973). However, beyond the study of ecological systems, resilience science has expanded and authors have adopted ecological principles in metaphorical discussions of social, political, economic, legal and institutional systems (e.g. Westley et al. 2002; Garmestani et al. 2008). Essentially, resilience can be conceptualised in the following ways;

1. **A measure of resistance:** The ability of the system to resist change;
2. **A measure of return:** The ability of the system to recover and return to a pre-existing state prior to a disturbance;
3. **A measure of absorption and functioning:** The ability of the system to absorb and persist at its current level of functioning before it is required to adjust to a new state of change;
4. **A measure of adaptation:** The ability of the system to learn and adapt towards a more desirable resilient state.

These theoretical fracture lines reflect different perspectives on system stability¹ (Gunderson and Holling 2002). Firstly, resilience may be regarded as a measure of *resistance* and *return* to a pre-existing state prior to the disturbance (point 1 and 2 above). This is sometimes referred to as **engineering resilience** and is focused on the speed to which the system is able to ‘bounce back’

¹ For a thorough account of the evolution of the resilience paradigm the reader is referred to Folke (2006).

(Pimm 1984; Holling and Meffe 1996; Folke 2006). This standpoint on resilience is strongly related to notions of optimal *efficiency* and linked to the control-over-nature approach typically adopted in resource and environmental management (Gunderson and Holling 2002). This paradigm is based on a once dominant assumption of steady-state systems, operating near or around equilibrium; i.e. when a disturbance occurs it is assumed that the system is able to self-organise and return to this state of equilibrium.

In the early 1970s this perspective was challenged, as Holling (1973) revealed the existence of multiple stable states, acting as “basins of attraction”. When a disturbance occurs, resilience becomes a measure of the ability of the system to absorb and persist at its current level of functioning before it is required to adjust to a new state of change. Resilience thus refers to the buffer capacity of the system and its ability to maintain its function, structure and fundamental identity. When this capacity is exceeded, the system may flip into an altered state and gravitate towards a new basin of attraction, either suddenly or gradually over time (Carpenter 2003; Walker and Meyers 2004). This paradigm is often referred to as “**ecological resilience**”. Whereas *engineering resilience* is concerned with the constancy of the system, *ecological resilience* focuses on the variability and non-linearity of system behaviour (Holling 1996; Gunderson and Holling 2002; Folke 2006). Diversity in functional groups of the system (e.g. predators, prey, pollinators) and corresponding response diversity is essential for sustaining ecological resilience (Gunderson and Holling 2002; Folke 2006)².

This perspective has drawn attention to dynamic systems, operating in unpredictable futures and opened discussions into **complex adaptive systems** (Berkes et al. 2003; Walker et al. 2004). Such systems are characterised by cross-scale interactions, fast-slow behaviours, non-linearity and the existence of multiple stable states (or domains of attraction). In this context, *panarchy* has emerged as influential theoretical and heuristic device, demonstrating complex system dynamics (Holling 2001; Gunderson and Holling 2002; Gallopín 2006). Panarchy is a theory that captures *unpredictable hierarchies*, moving away from the “top-down” connotations of traditional hierarchy theory and instead emphasising cross-scale interactions between nested sets of **adaptive cycles** (Holling 2001; Holling et al. 2002). Ecological functions can be organised into the different stages outlined in Table 4.9.

Table 4.9 Ecosystem functions organised into the different stages described in the “adaptive cycle” (Holling 2001; Holling et al. 2002)

Stage in the adaptive cycle	Description
r stage	Exploitive phase – characterised by rapid growth and accumulation of resources.
K stage	Conservation stage – characterised by slower growth and a sustained plateau of accumulated resources. <i>Resilience is low during this stage.</i>
Ω stage	Release stage (“creative destruction”) whereby accumulated resources become increasingly fragile (i.e. over-connected in the K stage) and are suddenly released.
α stage	Reorganisation phase, re-combinations can unexpectedly seed experiments leading to innovations in the next cycle. This stage reflects adaptive capability.

These phases influence the *potential* (inherent in accumulated resources) and *connectedness* among controlling variables of the system (Figure 4.3). Resilience is a further property influenced through

² Note that STAR-FLOOD has similarly adopted this principle and hypothesised that diverse, multi-layered FRMSs are required to enhance societal resilience to flooding.

these phases, decreasing and increasing between the K and α phases respectively. Extending this concept into the framework of panarchies, the notion of nested adaptive cycles means that a phase at one level can cascade up and influence larger and slower levels. In particular, Holling (2001) draws attention to the cross-scale connections of “revolt” and “remember” (Figure 4.4). “**Revolt**” may cause a critical change to cascade up from the alpha phase at one level, to the next larger and slower level, and potentially trigger change (especially if the level above is in the K phase). In turn, accumulated potential in a larger, slower cycle may provide opportunities or constraints for renewal (from K to α), described as the “**remember**” pathway. This framework highlights the importance of the alpha phase in stimulating variety and experimentation. Thus disturbance is seen as an important part of development, regeneration and the creation of opportunities for renewal (i.e. adaptation). Nested adaptive cycles mean that changes occurring at one level may cascade through the cross-scale interactions of “revolt” and “remember”, both critical for maintaining adaptive capability.

Figure 4.3 Adaptive cycle and key ecosystem functions (r, K, Ω , α) (from Holling 2001).

Spacing between arrows indicates the speed in the cycle, where closely space indicate slow change. The “front loop” (r to K) is a slow phase of growth and accumulation; whereas the “back loop” (Ω to α) is a rapid phase of reorganisation and renewal. Uncertainty and unpredictability are characteristic of the “back loop”, but is here that opportunities for innovation may be created.

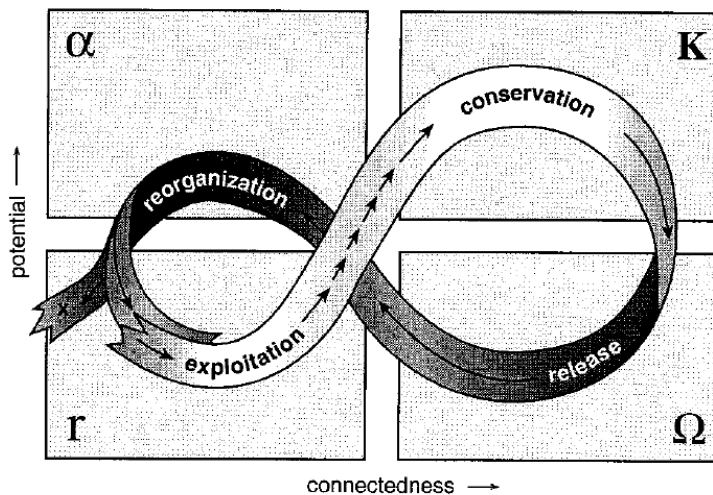
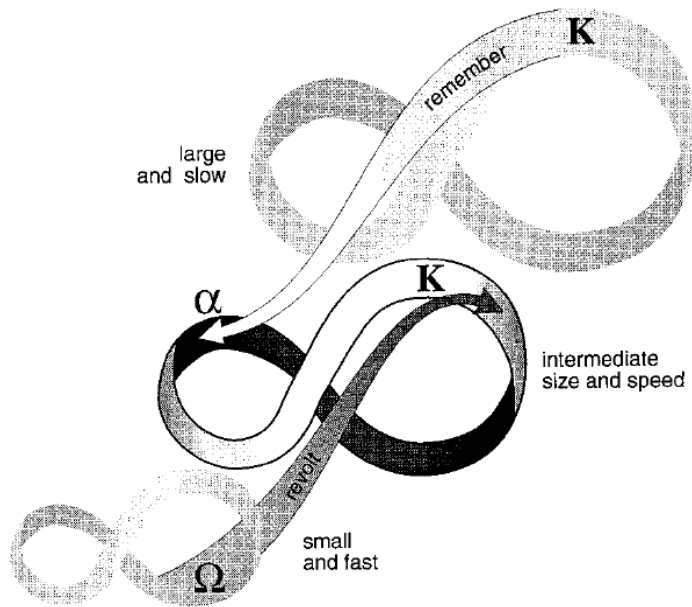


Figure 4.4 Panarchical connections (from Holling 2001).



Resilience science has expanded far beyond its initial focus on ecological systems and has filtered into discussions of social, institutional, economic, policy and legal systems (Gunderson and Holling 2002; Brock et al. 2002; Garmestani et al. 2008). Embracing complexity, there is a growing body of research into the coupling between social and ecological systems (SES) (Carpenter et al. 2001; Young et al. 2006; Gallopín 2006). On one hand, ecological resilience is a key concern for resource-dependent communities and may partially account for spatio-temporal patterns of resilience. On another, political economy theorists have highlighted how these patterns may be shaped through societal structuring, the role played by institutions (including habituated behaviour, norms, rules as well as formal and informal organisations); as well as the historical and cultural context in which these institutions evolve (Adger 2006; Oliver-Smith 2004; Wisner et al. 2004; McLaughlin and Dietz 2008). **Social-ecological resilience** refers to the ability of the system to reorganise (rather than “recover”) and develop following a disturbance. From this perspective, to be deemed truly resilient a community cannot simply ‘return to normal’ or merely persist after a stressful event. Processes for adjustment, learning and adaptation are emphasised in the step towards a more desirable resilient state. Even if this is not intentional, it is argued that societal systems will unavoidably change and cannot return to the status quo prior to the event (Steinfuhrer et al. 2009). This perspective emphasises the role played by processes and feedbacks across multiple spatio-temporal scales, as well as the potential for ‘regime shifts’; these may result from human actions (e.g. soil erosion or anthropogenic climate change) and in turn impact livelihoods and other aspects of the social system (Folke 2006). In this light, it is argued that social and ecological domains should be assessed hand-in-hand. Notions of *adaptability* and *transformability* are integrated into discussions of resilience and refer to the capacity of people to either build resilience through collective action, or to create a new socio-ecological system (Folke 2006). It is in this latter context of transformability that discussions of adaptive governance have emerged.

From the reviewed literature it is clear that adaptive capacity is central to discussions of resilience. Indeed, Holling (2001) defines adaptive capacity as the “*resilience of the system*” (p394), describing adaptation as a continuous cycle of transformation operating at different scales. Holling further argues that identifying points where system change is possible could create “*leverage points to foster resilience and sustainability within a system*” (p392). Other authors have argued that adaptive capacity is a component of resilience, reflecting the aspects of social learning (Carpenter et al. 2001). In contrast, resilience has also been portrayed as an expression of adaptive capacity (Walker et al.

2004). **From a STAR-FLOOD perspective, it seems that we cannot evaluate governance arrangements without consideration of adaptive capacity and adaptation.**

Numerous classifications for adaptation exist, relevant to time, intent, spatial scope, form and degree of transformation (Smit and Wandel 2006; Gallopín 2006). For example, Adger and Kelly (1999) define as adaption as the deliberate strategies for social and policy learning that enable evolutionary responses to environmental change. Other authors draw a temporal distinction between the concepts of coping and adaptive capacity; arguing that former is a reactive and short-term, whereas the latter reflects long-term adjustments (e.g. Folke et al. 1998; Kaspersen et al. 2005; or vice versa, see Lorenz 2013). **Here, we propose that adaptation should be regarded as the deliberate strategies that are intended to transform the system into a more resilient state in response to changing environmental conditions.**

The relationship between resilience and vulnerability:

To maintain the analytical value (and avoid confusion) in the evaluation framework we focus primarily on resilience and little on the concept of vulnerability. This is justified in D1.1.1 by the argument that the term vulnerability has become so amorphous that it has lost its analytical value. However, it is arguably inappropriate to conduct a discussion of resilience without a brief mention of vulnerability, seeing as the two are often debated hand-in-hand. Furthermore, it is important to recognise that this relationship is one that is highly contested. Authors' attempts to define this relationship can be summarised into four categories. Firstly, tautological, binary conceptualisations present vulnerability and resilience as alternative expressions of one another (Downing and Franklin 2004; Villagren 2006). In this context, vulnerability is simply the 'flip side' of resilience. Secondly, some authors distinguish this relationship across a temporal axis. For instance, Gallopín (2006) argues that vulnerability can be understood as predisposing characteristics of the at-risk community (prior to and during hazard impact), and resilience a feature of response. Similarly, Buckle et al. (2001) agree that resilience is a feature of recovery, but contend that vulnerability is relevant throughout the disaster management cycle (Steinfuhrer et al. 2009). For some authors, vulnerability and resilience are conceived as separate but related constructs, relevant for the study of different social units. Whereas vulnerability is conceived as a feature of the individual or household, resilience is described as a feature of the community and their institutions (Wilson 2008). Finally, there is a selection of authors conceptualising resilience as an integral and internal component of vulnerability (Pelling 1998; Clark et al. 2000; Bohle 2001; Cardona 2003).

For our purposes it may not be relevant to continue this conceptual debate. **Ultimately, regardless of how vulnerability and resilience are aligned it seems that the two are intrinsically bound in a causal chain of events; thus to enhance resilience is to reduce vulnerability.**

How can resilience be used to evaluate current FRGAs?

For the purpose of this project, we have adopted the definition of resilience proposed by the EU SPICOSA research project (<http://www.coastal-saf.eu/>). At this stage, this fairly broad definition is adopted (Box 4.3).

Box 4.3 Definition of resilience

"The ability to absorb disturbances, to be changed and then to reorganise and still have the same identity (retain the same basic structure and ways of functioning). It includes the ability to learn from the disturbance".

In determining whether FRGAs are resilient there is a need to consider how these serve to promote or even prevent resilience of systems individually (i.e. social, economic, ecological etc.) and as a whole. The challenge is demonstrating how resilience is altered through governance arrangements, rather than simply focusing on the influence of FRMSs and before/after comparisons. In the

evaluation framework presented in Chapter 5, it is suggested that partners begin by assessing the commitment and representations of resilience in policy and legal frameworks. Furthermore, it may also be possible to evaluate the resilience of FRGAs through *ex post* examination of documented experiences with flooding; for example, there may be evidence highlighting the ability of communities to recover from past floods.

Tensions may emerge between current policy and legal frameworks and the aspirations of resilience and adaptive governance. For example, Green et al. (2013) analyse the EU Water Framework Directive and inherent difficulty of balancing the need for flexibility and simultaneously need for robust and enforceable standards in adaptive water governance. Garmestani et al. (2008) also contend that current policy and legal frameworks are insufficient to address the cross-scale dynamics of socio-ecological systems. A number of authors have argued that transformations must occur within current policy and legal systems, which constrain adaptive forms of ecosystem management (Ruhl 1999; Garmestani et al. 2008). Arguably, there is a conflict between legal certainty and socio-ecological uncertainty, as well as a mismatch between scales of decision making and socio-ecological processes (Garmestani et al. 2008). Moreover, opportunities for innovation and learning are varied between different legal systems. Scholars contend that a multi-tiered nested framework is required to implement scale-specific policies and include important mechanisms for monitoring the implementation and management of environmental policies (Dietz et al. 2003; Ostrom 2006). In this context, Garmestani et al. (2008) emphasise that such policies should really be regarded as hypotheses subject to testing and revisions.

From an evaluation perspective and in the context of FRM specifically, we might evaluate the degree to which resilience is formerly considered and represented in policy and legal frameworks and which model of resilience is emphasised. The extent to which opportunities for adaptation and transformations are integrated can also be evaluated, as well as observing whether there exist any barriers to these goals. As stressed by Keesen et al. (2010), there is also a need to examine how definitions and proposed strategies of resilience are shaped by social contexts and normative values. For instance, Keesen et al. review how definitions of public interest and divisions between public-private responsibilities vary between different political orientations; including rights (limited or extended), utilitarian, socialist and communitarian. Moreover, normative shifts may also be observed in the analysis and evaluation of these texts.

Section 5.3 in the evaluation framework in Chapter 5 of the framework and methodology report provides a more detailed series of prompts for STAR-FLOOD researchers to consider in their evaluations. Whilst resilience is proposed as a specific criterion to be adopted by STAR-FLOOD researchers to facilitate later comparisons across partner countries; we recognise that there must be an inherent flexibility within the evaluation framework to enable new criteria to emerge. To this goal, it should be borne in mind that different interpretations of resilience may come to light from interaction with relevant stakeholders, or within different policy and legal frameworks. In this context, our evaluation framework must in itself be resilient in terms of its flexibility and opportunities for learning.

4.4.2 Appropriateness

The choice of *Appropriateness* for the second criterion has been derived primarily from the governance literature and is based on the assumption that the **implementation of a diverse (and resilient) set of FRMSs in a certain area is only possible if these strategies and their coordination are appropriate (legitimate, effective and efficient), i.e. properly institutionally embedded** given the opportunities and constraints of their physical and social context.

Governance that fails to resolve internal conflicts with a policy arrangement, or poorly delivers particular forms of intervention for example, may be considered to be inappropriate. There is

considerable literature about what governance arrangement should be adopted and what makes a better or poorer approach. In STAR-FLOOD we have chosen to try to avoid classical distinctions of good/bad governance and 'objective' determination of whether one FRGA is 'good' or 'bad.' Instead we are adopting a context-specific evaluation in line with the "*Logic of Appropriateness*" (March 1994; March and Olsen 2008). Additionally, later we will be drawing on legal perspectives of appropriateness in terms of its relevance for the principles of 'good' or 'proper' administration (e.g. Langbroek 2003).

March and Olsen's (2008) perspective on appropriateness considers how the rules organized into institutions determine what is seen as natural, rightful, expected and legitimate. Rules essentially provide codes of meaning that facilitate interpretation and shape human action. Beyond rational actor theories, it is argued that actors define what is appropriate for a given situation, thus matching the problem-solving action to a problem situation. Appropriateness is therefore reasoned through cognition and normative values. According to March and Olsen (2008), actors are driven to fulfil the obligations embedded in *roles, identities, membership in a political community or group, and the ethos, practices and expectations of its institutions*. In similarity to the discussions on resilience (see *Appendix to Chapter 5*) March and Olsen explain how institutions, roles and identities may grow or decay, and how rules can create order and stability, as well as flexibility and adaptiveness (p12). The formation of rules is seen as dynamic, with opportunities for transformation through experience and social learning.

In relation to STAR-FLOOD, the *Logic of Appropriateness* draws attention to the relationship between "the rules of the game" and human action. In our analysis and evaluation it is important that we examine this relationship thoroughly, considering the factors that strengthen or weaken this relationship in the context of FRM. In particular, we should be interested in how lessons acquired through experience shape the formation of these rules and how decisions are negotiated between potentially conflicting actors. Potential examples of a 'lack of appropriateness' involve the failure to resolve internal conflicts within a policy arrangement. So the starting point of case study research might involve examining whether a FRGA performs in practice how it was intended to perform and, where possible, examine reasons why it is not able to. For instance, new strategies or working procedures may be incompatible with explicit rules and there may be institutionalised ideals which can never be realized in practice. This may be because there is a lack of faith in institutions, or intra- or inter-institutional tensions are present between organizational and normative principles.

Appropriateness is ultimately established by "*fitting a rule to a situation*" (March and Olsen 2008: 9). In order to evaluate this, three sub-criteria have been identified, which are discussed and reviewed both in the following sections. In each case, the theoretical background is outlined to justify the underlying rationale for the selection of these sub-criteria;

- *Effectiveness;*
- *Efficiency;*
- *Legitimacy.*

4.4.3 A theoretical overview of Effectiveness

There are many differing perspectives about the concept of effectiveness, its meaning and how it might be evaluated. This section explores some of the literature about effectiveness and its given definitions and considers how STAR-FLOOD researchers might use this concept to evaluate flood risk management at different levels. In addition, it briefly considers how effectiveness links to the other appropriateness sub-criteria.

In terms of evaluation, at a most basic level effectiveness requires an assessment of whether FRGAs and delivered FRMSs have achieved their intended purpose. Effectiveness thus provides a measure of the success to which arrangements or strategies successfully function or address the problem they were designed to address. Therefore, effectiveness is applicable to discussions of *process* and *outcome* (Young 1994; Bäckstrand 2006). To evaluate this concept, requires comparison against a standard of success (Underdal 2002). In order to achieve this, Underdal poses the following questions;

1. What constitutes the object being evaluated? (i.e. FRGAs/FRMSs)
2. Against which standard is the object to be evaluated?
3. How do we go about comparing the object to this standard (i.e. what kind of measurement operations do we perform to attribute a certain score of effectiveness to a certain regime)?

To the first point, there is a need to define what we are measuring. Evaluations of effectiveness require a distinction between the regime formation processes and the resulting consequences of these processes. In the latter context, Underdal further distinguishes consequences in the form of change to human behaviour (*outcome*), compared to those that result in changes to the environment itself (*impact*). Collectively, Underdal describe effectiveness as “*a function of the stringency and inclusiveness of its provisions, the level of compliance on the part of its members, and the side effects it generates*” (p6).

One of the key challenges in evaluating effectiveness is establishing what we are measuring against. Herweijer (2007) argues that in order to determine what is effective can only be achieved via empirical research requires a clear comparison of the situation before the measure/strategy/governance arrangement was in place with the situation afterwards. This necessitates that we define a standard or point of reference against which governance approaches and strategies should be evaluated, as well as a standard metric of measurement (Underdal 2002). To the first, effectiveness can be conceived in terms of **relative improvement** caused by the regime (i.e. a measure of success) and assessed in comparison to a hypothetical state that would have existed if the regime had not been implemented. The challenge with relative improvement evaluations of efficiency is the inherent requirement for some form of baseline in which to measure this improvement. It is suggested that this baseline may be deduced from the hypothetical state. Alternatively, effectiveness could be evaluated in terms of an “ideal” situation. Here, the **collective optimum** is defined as “*one that accomplishes, for the group of members, all that can be accomplished – given the state of knowledge at the time*” (Underdal 2002: 8). This approach essentially considers the ‘distance’ between what has and what could have been accomplished; thus, it requires some decision of what constitutes the *maximum*. This also calls into question what is politically and practically feasible. Relative improvement and the collective optimum are notions that are both relevant to the evaluation of effectiveness.

The third question considers the challenge of scoring effectiveness. This may be more problematic when criteria do not lend themselves to quantitative assessment and draw from subjective interpretations; here, there is a need to include steps in the research method that actively seek to ensure transparency and reliability in qualitative assessments across multiple researchers. To this goal, Underdal (2002) proposes a pragmatic method that is based on ordinal level measurement; distinguishing between four and five levels of effectiveness in the assessment of the collective optimum and relative improvement, respectively. These scales constitute a yardstick in which to measure effectiveness.

However, given the disparate interests of these two notions, we can be critical about the extent to which it is possible, and indeed advisable, to aggregate these assessments to deduced a level of overall effectiveness. To this end, Underdal (2002) proposes that an effectiveness coefficient can

express *actual* improvement as a fraction of *potential* improvement. However, there is also a question of whether different definitions of effectiveness can be aggregated into a single measure. Indeed, is it appropriate to use the same model to assess the effectiveness of *problem-solving* (or “output legitimacy as described by Bäckstrand 2006), as well as *process effectiveness* (Young 1994)? This is something that will warrant further debate in the context of STAR-FLOOD.

Before discussing more about the approach proposed in STAR-FLOOD, it is important to consider how effectiveness is considered in the legal realm. Chambers (2004: 519) argues that there is a significant difference between a legal definition of effectiveness which may only consider the legal instrument (e.g. a treaty) and a broader social one which considers aspects such as the “‘principles, norms, rules, procedures and programmes’ that comprise the regime.” From an instrumental perspective in law, effectiveness has often been considered to be related to the functionality of the legal approach and whether a rule is followed and does what it intends to do. Moreover, there will also clearly be a difference in the way in which effectiveness is viewed between different legal approaches. By way of illustration, in French administrative law the concept of effectiveness is often used to mean “constitutive effectiveness” or “evaluative effectiveness” as per Young’s dimensions. In a great deal of ways effectiveness in a legal sense can be considered to be synonymous with the performance of a measure. STAR-FLOOD researchers need to consider this when undertaking evaluations of effectiveness and assess whether merely evaluating performance adequately captures all of legal notions of effectiveness within their legal regime or whether other elements also need to be considered. A traditional legal review therefore does not take into account effectiveness *per se*, but rather assess whether it is ineffective. If an approach or system is not shown to be ineffective then it will pass a legal test of effectiveness. This provides only a very narrow evaluation but one which may provide more clarity in measurement as Chambers (2004) argues that a regime analysis and the introduction of many variables may cloud the situation. Therefore, there is a potential challenge here for STAR-FLOOD researchers to tackle the trade-off between complexity and simplicity in the evaluation approach.

Through an analysis of the legal literature Buijze (2008) argues that more recent considerations of effectiveness in the legal realm. She argues that some legal researchers have moved beyond the recognition of the instrumental function of administrative law and highlights three differing distinction of legal effectiveness which is:

- The purpose of specific fields of administrative law is to accomplish their goals as often and as well as possible (Van Wijk, Konijnenbelt and Van Male 2005)
- The administration must strive to accomplish its goals (in the general interest), while respecting the law. This emphasizes the notion that the rule of law is a duty and that one of the characteristics in democratic law rests on the state to act as effectively and efficiently as possible (Michiels 1998)
- A pragmatic concept of effectiveness is that a goal is set needs to be attainable. The definition by Blomberg and Michiels (1997) therefore consists of three related components: that an attainable goal has to be set, that this goal has to be achieved and that it needs to be achieved as efficiently as possible.

There may also be an overlap between the idea of Proportionality and effectiveness. For instance, Buijze (2008) argues that this principle is primarily concerned with the balancing of differing interests and it has an effectiveness component in the sense that it may be used to establish the effectiveness of administrative decisions. It is important to recognise the potential problems in measuring effectiveness when considered from a more traditional legal perspective.

Chambers (2004) argues also that effectiveness in the legal context is changing and increasing in importance due to an increased attention on international law and a growing acceptance that

multilateral approaches are required to deal with environmental problems. It is suggested that although compliance (and therefore the performance of a legal aspect) is important and a relevant measure of effectiveness; relying solely on compliance is inadequate to capture all of the potentially relevant aspects of effectiveness. He contends that lawyers have maintained a rigid and narrow view of effectiveness and in order to move beyond this there is the need to draw on other disciplines as well as from within the legal sphere. For instance, Chambers (2004: 508) states that effectiveness from social legal theory is concerned primarily with “social change and the congruence of law with societal norms.” He also presents, among others, an ‘economic legal model (which includes the concepts of efficiency and cost-effectiveness) and a ‘natural legal model’ (which introduces the notion that norms should be based on natural laws or universal principles). There is scope within the STAR-FLOOD project for legal scholars to explore the concept of legal effectiveness from these differing perspectives and in particular those which aim to provide a multi-disciplinary overview.

Within this project we have decided to adopt and utilise Young’s (1994) six dimensions of effectiveness as presented below:

- Effectiveness as problem solving;
- Effectiveness as goal-attainment;
- Behavioural Effectiveness;
- Process Effectiveness;
- Constitutive Effectiveness;
- Evaluative Effectiveness.

Young (1994) considers these different elements to be distinct which means that the different dimensions can be considered separately and that in theory an approach or strategy can be considered to be in some way effective even if one (or more) of the dimensions are not fulfilled. Indeed, Young (1994) argues that there is no basis for assuming neither that there is any co-variance between the dimensions nor that this be easily predictable. However, in reality there may be a great deal of interconnection and overlap between the dimensions. For instance, in principle there is the potential for there to be a high degree of overlap between ‘Process effectiveness’ (i.e. the extent to which an overarching provisions is implemented at the local scale) and Young (1994) also argues that there are likely to be “causal links” between ‘Goal-Attainment effectiveness’ and ‘Problem-Solving effectiveness’. A negative element of these evaluation measures is the fact that there are varying degrees of subjectivity associated with the approach and the result may depend upon who is making the judgement. For instance, the goals of a strategy or measure may be clearly articulated within some related documentation and it is against this that ‘Goal-Attainment effectiveness’ can be assessed, however when evaluating Problem-Solving Effectiveness whose view of the problem is being evaluated? In this regard, STAR-FLOOD researchers need to seek to explore issues of effectiveness with stakeholders and thereby better understand differing perspectives on aspects such as the nature of the problem.

Young’s (1994) approach has the value that it creates a much more comprehensive evaluation of the concept of effectiveness. It moves beyond merely evaluating the solitary question of whether an approach reached its stated goals and provides a broader understanding of the other potential intended and unintended consequences of an approach and how these might contribute to effectiveness. Through some of the dimensions of effectiveness considered it also is consistent with legal viewpoints of effectiveness.

In his explanation of the variables which contribute to the effectiveness factors, Young (1994) initially distinguishes between decision variables (those subject to conscious control and may include the institutional arrangements and the decisions taken in designing and managing governance systems) and structural variables (those not subject to conscious control and include the more contextual

factors which belong to the wider, physical, biological and social environment). Young argues that a full analysis will require both of these elements to be considered for a thorough evaluation and the STAR-FLOOD project recognises this through the three levels of analysis: context, NFPR and case study. However, Young (1994) goes on to propose a related characterisation of those determinants impacting on *Effectiveness* and considers a division into the following categories:

- **Endogenous** variables – These internal components emerge from the characteristics of properties of the governance arrangements themselves and may include (amongst other things) the study of processes of decision-making, resource allocation and sources and organizations;
- **Exogenous** variables – A vast collection of external variables may impact upon the effectiveness of a regime and may include aspects such as power, available knowledge and the presence of absence of different interests;
- **Linkage** variables – The nature of the problem and including the administrative organization of the member state (e.g. strongly centralized vs decentralized).

Young highlights a number of variables which are useful for STAR-FLOOD researchers to bear in mind in the analytical framework and may contribute to effectiveness these include the presence of change (and therefore also stability and the robustness of the system), transparency (and therefore also the formulation of rules), and adaptability and flexibility within the governance arrangements. He argues that long-term effectiveness requires an in-built capacity to change and adjust.

As a success criterion, the different dimensions of effectiveness of FRGAs or FRMSs might be measured by indicators such as the changes in risk, such as the alleviation of flood losses or a reduction in the number of people or properties exposed to flooding or how the behaviour or actions of institutions or those at risk have changed. From a legal perspective, the extent to which a legal instrument results in the action that was intended by its introduction, could also be evaluated. Equally important is the perception of effectiveness amongst relevant stakeholders. Indeed, do the defended public feel safer and is this perception of safety in line with scientific assessments of safety?

Although these questions may more easily facilitate comparison and provide tangible measures of effectiveness, there is a need to exercise caution. Whilst FRMSs are delivered through FRGAs and assumedly reflect the effectiveness of FRGAs, we must also examine other aspects of the arrangement profile and link back to the Policy Arrangement Approach (PAA). There are clearly multiple avenues of effectiveness to consider. STAR-FLOOD researchers should evaluate both single outcomes and the overall effectiveness of the governance approach, distinguishing between arrangements (process) and strategies (outputs).

Ex ante evaluation is particularly difficult and requires us to consider what are the desired outputs; indeed, what are we trying to achieve? Our goal is to propose resilient and appropriate FRGAs that ultimately enhance the resilience of vulnerable urban agglomerations to flooding. Therefore, our measure of effectiveness is one that achieves this goal. What makes for a resilient and appropriate arrangement is subject to hypothesis testing in this research, thus the specific inputs and outputs will emerge from this study. Insights from *ex post* evaluation, should highlight the causes and barriers to effective implementation of FRGAs and FRMSs, and therefore highlight where opportunities exist. Exemplars of effectiveness and ineffectiveness are therefore equally valuable. Box 4.4 highlights some higher level reflections on the evaluation of Effectiveness and provides some illustrative examples of potential comparisons between countries.

Box 4.4 Some higher level reflections on effectiveness and comparison

The analytical and evaluation frameworks need to ensure that sufficient data is collected within the case studies and partner countries to permit comparison and therefore, where possible in the protocols for WP3, we should try to begin to identify comparative questions we might want to answer. These might include:

- At the individual intervention option level we could ask, for example, does the combination of PPR(1)s and other plans in France deliver floodplain development control? How does it compare in terms of success with the PPS25 approach in England and the Water Assessment framework in the Netherlands?
- France is particularly interesting because of the evolutionary changes that have taken place, why did they occur, what were the problems experienced and how were they resolved? How and why did evolution in policy or practice occur or not occur in other partner countries?
- We might approach other issues (e.g. SUDS in urban areas, rural runoff management and flood warnings) in the same way and begin to develop a series of comparative questions to ask within later work packages to ensure that relevant data is collected within WP3.
- A meta-issue is that sustainable flood risk management depends very much upon changing people's behaviour: what and when are the most effective ways of doing this?

When considering effectiveness we should also not view this sub-criterion in isolation from that of either legitimacy or efficiency as there may be many inter-linkages between these concepts. Paavola (2008) demonstrates the link between social justice and the effectiveness of environmental decision making. The author makes the argument that a process which promotes social justice (which he considers to be underpinning legitimacy) is more likely to have outcomes which are complied with and in this way can be seen to be contributing to the effectiveness of the decision or action taken. Indeed, Paavola (2008) also suggests that "the lack of legitimacy may prevent the making of these decisions in the first place, and impair the implementation of decisions that have been made". In this way a lack of legitimacy can be seen as a barrier to effectiveness.

Bell and Tobin (2007) in their study examining the impact in the US of using the flood risk policy benchmark of the 100 year flood term and as a public communication tool provides an example of an operationalisation of an evaluation of effectiveness. In their approach, the authors query the rationale that communication should be considered to be effective if it influences a receiver's attitude and behaviour and support the notion that effectiveness in communication requires understanding: a notion tested in their study. In this way they have considered two different facets to the effectiveness of the communication strategy: understanding of uncertainty and persuasion. This study also provides some insight about the potential interconnections between efficiency and effectiveness in practice. Bell and Tobin (2007) argue that the "initial goal of adopting the 100-year flood criterion was not effective communication of flood risk policy but efficient administration and implementation" (p302) and that in reality the design of communication tools necessitates the trade-off between using terminology and approaches that are more effective but that do not contradict the ability of it to be used in large nationwide communication programmes.

4.4.4 Challenges of evaluating effectiveness

The discussion above details potential measures and indicators for the evaluation of *Effectiveness*: however these do not provide the full picture. Indicators are able to provide some broad indication about whether a goal is met or a problem is solved but in reality fully substantiating the true effectiveness of a governance arrangement or FRMS is more complex and these indicators are only partially fulfilling an evaluation. In order to fully determine *Effectiveness* there is the need to establish causality between the adoption of an approach, strategy or measure and the resulting outcomes and impacts. For instance, STAR-FLOOD researchers need to try to understand whether the adoption of a governance approach has directly led to a positive or negative impact on the

problem of flooding or whether the outcome has been an unintended consequence or the consequence of another set of circumstances.

This process is assisted to some extent by the comprehensive approach to the evaluation of *Effectiveness* recommended here. Adopting Young's (1994) different dimensions of effectiveness may provide evidence to determine this causal link. For example, understanding behavioural change (within an evaluation of 'Behavioural effectiveness') and tracing how and why different behaviours have emerged may provide insight into the reasons for an approach being *Effective* or *Ineffective*.

Evaluating *Effectiveness* and establishing a causal link should be easier to achieve the smaller the unit of analysis. It is likely that the link between a measure and its outcome and impact will be easier to demonstrate than that of a governance arrangement which may oversee a number of flood management aspects and strategies. Within this consideration and in particular when researchers also choose to evaluate *Effectiveness* of not only the governance approach, but also instruments at different scales, is how these results can be integrated. There will be a distinction between evaluating single outcomes and the effectiveness of FRGSs and translating this into the evaluation of a regime or a whole governance approach to flood risk management. As mentioned above, tackling the former is much more straightforward than the latter. In addition, in order to be able to corroborate or contradict the effectiveness of an approach the analysis should also seek insight into the causes or barriers to the effective implementation of FRGAs or FRMSs. This aspect is critical to any comparison of approaches and best practices and for providing recommendations for the design of FRGAs. This may also highlight where opportunities exist. Where FRGAs have not fulfilled their intended goals, understanding why is critical to evaluating the *Effectiveness* of an approach.

4.4.5 A theoretical overview of Efficiency

Efficiency is adopted here as a sub-criterion of Appropriateness. The simplest and most rigorous definition of efficiency is that given in physics, where efficiency is the ratio of the output of work to the input of energy. Efficiency can be measured in terms of the conversion and in terms of the proportion of energy that is not being used (i.e. wasted). Alternatively, efficiency can also be assessed in the context of alternatives and each corresponding combination of objective(s) to input(s).

Whilst efficiency has been discussed in physics and other disciplines, the concept is rooted in economics. Numerous terms surround discussions of economic efficiency and these are summarised in Table 4.10. Cost-benefit analysis (CBA) is a widely-applied method for testing efficiency, requiring the benefits of a proposed flood risk measure for example, to exceed to calculated total costs. This is a relative assessment of efficiency driven to the least costly option and considered to be a rigorous way to judge actions with multiple consequences. Essentially, this approach is an application of the Kaldor-Hicks principle (Table 4.10; Hanley and Spash 1993). An example of this approach is presented in Hanley and Black (2006).

Table 4.10: Key principles of economic efficiency (Winch 1971)

Term	Definition
Allocational efficiency	An output level where the price equals the marginal costs of production. Based on the rationale that the price consumers are willing to pay is equivalent to the marginal utility that they will receive.
Production efficiency	No additional output can be obtained without increasing the amount of inputs. Production therefore proceeds at the lowest possible per-unit cost.
Pareto efficiency	Refers to a state of allocation of resources in which it is impossible to make any individual better off without making at least one individual worse off. Under Pareto efficiency, an outcome is more efficient if at least one person is

Term	Definition
	made better off and nobody is made worse off. An allocation is defined as “Pareto efficient” when no further “Pareto improvements” can be made.
Kaldor-Hicks efficiency	An outcome is considered more efficient if a Pareto-optimal outcome can be reached by arranging sufficient compensation from those that are made better off, to those that are made worse off. The desired result is that no one should be made worse off than before. An outcome is regarded as more efficient if those that are made better off could <i>in theory</i> compensate those made worse off. This criterion is widely applied in welfare economics and managerial economics.

Although these principles of economic efficiency are useful, for the purpose of STAR-FLOOD we are proposing to view efficiency from a very overarching practical perspective, rather than taking a purely economic approach. Arguably, economic notions of efficiency are too narrow and based on multiple assumptions. Therefore to avoid constraining our evaluation of efficiency, we will propose a definition that reflects broader discussions of *resource efficiency*.

Furthermore, we acknowledge that evaluations of efficiency may have both *descriptive* and *normative* aspects. For example, a local flood alleviation scheme may be efficient in reducing flood losses (i.e. it is *descriptively efficient*). However, the scheme may increase flood losses elsewhere or result in significant environmental costs to the catchment; in this instance, the scheme may be considered to be *normatively inefficient*. From this normative standpoint, it is claimed that efficiency judgments ought to be based upon wider concerns, beyond the simple measure of inputs and outputs. This example highlights the potential for conflict between evaluations of efficiency. Indeed, long-term efficiency measures required for sustainable FRM may be at odds with short-term assessments of FRMSs.

In light of this conflict and the multiplicity of potential outputs, there is a question of whether measures of efficiency are compatible. Potential tensions between descriptive and normative, short-term and long-term goals, makes an aggregated measure of efficiency particularly problematic. Can they (and should they) be aggregated and if so, are they additives or substitutes? Should we distinguish between descriptive and normative facets of efficiency in our evaluation?

As soon as more than one output and a single input is considered, the problem becomes how to take into account differences between ratios and whether to seek to collapse these into a lower number of ratios. The benefit-cost ratio illustrates the problems of collapsing multiple ratios into a smaller and more measurable number of ratios. Benefit-cost analysis often involves the comparison of the incommensurate; i.e. the consequences of the different options, differing in who is affected, how they are affected and when the effect occurs. The use of discounting in benefit-cost analysis highlights the problem of comparing consequences that are distributed over time. However, the virtues of CBA in this instance are that it provides an explicit, rigorous and consistent means of taking these differences into account. However, the discounting approach adopted can be argued to be founded on simplifying assumptions that are unsound; notably, on assumptions about individual time preferences that Modigliani (1966) and Friedmann (1957; 1980) argue to be false. More generally, the discounting approach also takes no real account of the time profile of benefits and costs (Green 2003).

How can we evaluate FRGAs and FRMSs from an efficiency perspective?

In order to evaluate efficiency in the context of FRM, it is necessary to define both the *desired* inputs and outputs. Returning to the *Logic of Appropriateness*, the notion of “desired” inputs and outputs should be integrated in our definition of efficiency to emphasise the role played by actors in determining this. Rather than adopting an economic-centered view on efficiency, we have adopted a

pragmatic definition (Box 4.5) which considers more broadly, the efficient use of public and private resources.

Box 4.5 Efficiency can be defined as

The use of both public and private resources in a resource-efficient manner. More simply put, it may look at the ratio of some desired output(s) to some input(s).

From an evaluation perspective, STAR-FLOOD partners should examine whether there is any indication that efficiency has been considered in FRM decision-making. This may be reflected through a rigorous process by which alternative FRMSs have been considered, in which case we may ask which actions were considered? What was the basis for selection? What determines when actions should be put into place?

Whilst CBA is the most articulated form of rigorous analysis, it is not the only example and the existence of some form of rigorous, formal comparison is the test. As a comparison, it informs the basis for decision making about which option to adopt. It is not a formal means of legitimating a decision that has already been made (as benefit-cost analysis is quite often used). An indication of whether the analysis was undertaken as part of the decision-making process or a ritualistic validation of a decision made on another basis, requires us to examine whether the analysis reports the parameters that have had a critical affect upon what option should be adopted (i.e. some form of sensitivity analysis). Moreover, we might consider the extent to which efforts were made to try to improve estimates of these identified variables. In the case of benefit-cost analysis, a simple indicator is the number of significant figures to which the benefit-cost ratio or Net Present Value is presented. Reporting more than two significant figures is a good indicator of a lack of thought, as giving an ratio to more than one decimal place implies a degree of accuracy in the predicting the future. Such recklessness, rather than rigor, is likely to permeate the entire analysis.

The fundamental problem in assessing efficiency is whether there exists an unidentified or undiscovered option that is more efficient than any of the options that have been considered or adopted. Therefore the real test is whether there are better options. In turn, it has been argued that a civil servant should always present their minister with three options – 1. An option which will lose their party the next election, 2. An option which will result in the minister losing their seat in the next election, 3. The option which the civil service has concluded is the best option. In this light, a key test is firstly the range of options that are considered and secondly, how these options were *identified or created*.

From a legal perspective it is recognized that efficiency may be more problematic. There are a number of objections to be mindful of when it comes to legal principles of efficiency and these are explored in more depth in Buijze (2008). Firstly, there is a concern that the emphasis on maximising utility neglects matters of distributive justice and the allocation of costs and benefits. In this context, the efficiency criterion could be seen as conflicting with Legitimacy. Secondly, from a moral standpoint, efficiency (and effectiveness) can be conceived as *consequentialist* in nature and therefore only partially represent moral reasoning, which must also include *rules of right*. The State as a moral actor must clearly balance the two. In defence, Buijze (2008) contends that considerations of efficiency actually enhance the decision making process, improve its transparency and provide opportunities for parties to challenge decisions. Furthermore, such assessments as CBA can be used to establish citizen rights and basic entitlements to flood protection measures for example. Therefore, efficiency can in fact enhance Legitimacy. Moreover, efficiency need not be constrained by economic reasoning and could include forms of ethical inquiry.

In fact, efficiency is already implicit in current principles and practices. For instance, at an EU level, some form of efficiency review is already practiced in legal reviews of compliance through the

necessity component of the proportionality test. In contrast to CBA which demands that the benefits *outweigh* the costs, this prescribes that the costs should not be *disproportionate* to the benefits. In operationalising this sub-criterion, legal scholars are referred to the series of prompts presented in Chapter 5.

In addition, the efficiency of both governance and legal arrangements may also be subject to evaluation, as the decision-making process itself carries a cost and use of resources. Applying the efficiency criterion to the evaluation of governance arrangements directly may be more difficult. Here, we might observe how resources are currently used or are potentially changing. Indeed, in many countries procedural law is being simplified to minimise decision-making costs.

Given that the question of efficiency in flood risk management can be framed in terms of any combination of output(s) to input(s), this is a more difficult question and requires definition of the *desired* output(s) and the input(s). For example, we can say that the requirement is that long-term efficiency should be considered so that resources are to be used sustainably i.e. so that human life can be sustained on this planet for the indefinite future. In turn, the multiplicity of potential outputs raises the question of whether they are all compatible with each other; can we have all at once? The expectation must be however, that they may be incompatible with each other as Sen (1992) observed, the achievement of one form of equality typically precludes the achievement of another form of equality.

It should also be borne in mind, that discussions of efficiency can be thought of as having both *descriptive* and *normative* aspects. For example, a local flood alleviation scheme may be efficient in reducing flood losses (i.e. it is *descriptively efficient*). However, the scheme may increase flood losses elsewhere or result in significant environmental costs to the catchment; in this instance, the scheme may be considered to be *normatively inefficient*. From this normative standpoint, it is claimed that efficiency judgments ought to be based upon wider concerns, beyond the simple measure of inputs and outputs. This example highlights the potential for conflict between evaluations of efficiency. Indeed, long-term assessments of efficiency required for sustainable FRM may be at odds with short-term assessments of FRMSs.

In light of this conflict and the multiplicity of potential inputs and outputs, there is a question of whether assessments of efficiency are indeed compatible. Potential tensions between descriptive and normative, short-term and long-term goals, makes an aggregated evaluation of efficiency particularly problematic. Can they (and should they) be aggregated and if so, are they additives or substitutes? Should we distinguish between descriptive and normative facets of efficiency in our evaluation?

Evaluating the efficiency of an FRGA and delivered FRMSs requires understanding of the balance and division of resources, responsibilities and roles between different institutions and individuals in both the public and private sector. Evaluation using this sub-criterion therefore needs to refer back to that of *Effectiveness*. For instance, in Young's (1994) conceptualisation, efficiency was included within the notion of 'Evaluative effectiveness'. Whereas, *ex post Effectiveness* evaluates whether the intended consequences were achieved; here the objective is to assess whether these consequences were achieved in a mode that was as efficient as possible (Arts and Leroy 2010). In terms of *ex ante* evaluation researchers might question which potential alternatives will effectively realise flood risk management goals while using scarce resources as efficiently as possible.

4.4.6 A theoretical overview of Legitimacy

Why is legitimacy difficult?

As a basis for potential criteria, legitimacy is a difficult area for good reasons. Ultimately, legitimacy concerns the use of power over others; or to the individual concerned, the application of power over

them. As such, legitimacy concerns social relationships both as they are ('descriptive legitimacy') and as they ought to be ('normative legitimacy') (Fabienne 2010). Ideally, the two questions can be separated. For example, the Netherlands Code for Good Public Governance defines legitimacy as an applicable principle essentially in terms of *descriptive legitimacy*: "the executive body takes the decisions and measures that it is empowered to take and that are in accordance with applicable legislation and regulations" (Ministry of the Interior and Kingdom Relations 2009: 12). In practice, the distinction is less easy to make since the legislation and regulations will not be wholly prescriptive so that alternative decisions will be possible within the letter of the legislation and regulations.

The normative question is obviously one that is open to contestation. It is a relationship and hence a dynamic rather than a state or characteristic. As a social relationship, discussions of legitimacy rapidly engage other equally difficult and contested issues, notably *justice* and *equality*. Again, since it concerns what social relationship either do exist or more especially, what ought to exist, the question of legitimacy has engaged some of the major thinkers in human history. That in turn suggests that whilst being ambitious in our intentions, it would be wise to be modest in our expectations that we will provide a more conclusive answer than Aristotle or Weber, to take only two examples of those thinkers. Equally, that those answers we do propose, and the criteria we adopt, need to be grounded in or reflect upon that existing body of thought.

A further problem is that legitimacy is a recursive problem; where one person or body can either expect or requires another person or body to follow the requirements of the first, then the question of legitimacy arises. So, for example, if the legal system is a means of establishing compliance with the requirements or expectations of one party, then the question of the legitimacy of the legal system and its components, notably judges, immediately arises. Hence, a large part of the philosophy of law (in English usage, 'jurisprudence') is concerned with the legitimacy of the law itself and of the actors within that system (e.g. Dworkin 1986).

A particular problem for us in this project is if, as elaborated later, legitimacy in a modern European state ultimately depends upon the claim to democracy, the decisions of the *demos*; then only the *demos* have the right to make normative judgements as to what ought to be both the *form* and *content* of the relationships between individuals or groups. Consequently experts of no hue have a claim to be able to determine what the form and content of those relationships ought to be. So in seeking criteria to test the legitimacy of governance arrangements, those criteria can only be either those established as common European norms through legislation, or reflect the views of those parties engaged in a particular decision. It would be otherwise illegitimate for us to seek to set criteria which seek to define the criteria for legitimacy of governance processes where these have a normative content. Since this is a very restrictive condition, and notably fails to provide any advice to a *demos* who seek to set up a governance arrangement that delivers legitimacy, this annex also looks at the empirical evidence of the conditionalities that in particular contexts seem to determine whether the result is legitimate. But the empirical evidence, and more general thought, can only provide indicators; it is the decision of the relevant *demos* to adopt those conditionalities that would turn them into criteria.

What is necessary and possible is to provide a clear analytic framework.

The legitimacy of 'what'?

Early discussions of legitimacy (e.g. Weber 1947) were concerned with the legitimacy of the nation state. Thus, discussions of legitimacy were tied to the question of authority (Dyzenhaus 2001; Fabienne 2010; Hegtvedt and Johnson 2000; Lake 2006; Tyler 2006): who has the right to expect others to obey their power and consequently is entitled to use coercive power to enforce their demands? In turn, the discussion of legitimacy has been based upon the assumption of a nation

state and one that is hierarchically organised. Thus, discussions on legitimacy typically start with reference to Hobbes and Locke and have been concerned the conditions that create authority in a nation state (Fabienne 2010). Definitions of legitimacy of this time were thus centred upon the conditions under which a nation state or more generally an aspect of the nation state had the right to use coercion to impose its decisions upon its members.

However, in a modern society, the question of legitimacy is much wider. Thus, Scharpf (1999) has addressed the question of the legitimacy of EU governance arrangements and Keohane (2007) those of formal international organisations, such as the World Trade Organisation, more generally. The legitimacy of informal international organisations, such as voluntary agreements to promote sustainable use of resources, has also been examined. Whereas the practice now in Europe is to rest claims of legitimacy finally upon the *demos*, that claim cannot be made for arrangements between governments or between non-governmental organisations. In those cases, an alternative basis for establishing legitimacy must be derived. Since legitimacy is a recursive question, the legitimacy of sub-national governance arrangements may also be questioned. An area which has been widely studied is the legitimacy of the police service and the justice system more generally (e.g. Myhill and Quinton 2011). Equally, stakeholder engagement itself raised questions of legitimacy (Mackintosh 1977).

These questions raise issues about which assessment of legitimacy is to be made and the nature of legitimacy. Hence in contrast to earlier definitions, Scharpf (1999) gave the following definition of legitimacy: *“a socially sanctioned obligation to comply with even if these violate the actor’s own interests or normative preferences, and even if ... sanctions could be avoided at low cost”*. In contrast to the early definitions of legitimacy in terms of authority, Scharpf’s definition is in terms of consent: i.e. under what conditions does the individual have an obligation to consent to the decision and action taken by some agency? Since he is writing within the context of a democracy, this can be interpreted as *“when does the individual have an obligation to consent to a decision taken by the collective?”* Scharpf’s definition also implies that this obligation is socially constructed, whereby legitimacy is dependent upon a social norm to consent. In turn, such a social norm is potentially socially enforced. Here, we might question the conditions and circumstances in which an individual has an obligation to comply with collective decisions. Rather than being a general obligation, it may be an obligation that is delimited to particular circumstances and conditions. For example, there are parts of Europe when spatial planning regulations restrict development but much development takes place in violation of those regulations. Hence, the question is whether in those areas there is only a weak social norm to consent or land development is seen as an area where the general social norm does not apply.

Early research on the legitimacy of the nation state was concerned with *authority*. This was a question of the legitimacy of a ‘thing’, the nation state, and defined legitimacy in terms of a characteristic or state. That is, legitimacy once created then legitimised all decisions and actions. The alternative view is that legitimacy rests on individual decisions and actions; those that the legitimacy of the thing, the actor in question, is determined by and dependent upon the legitimacy of each decision and action taken. The critical question addressed in that early work is, however, still relevant: *when is there an obligation to comply?* Hence, there are three questions to be answered in the case studies:

- What is the organization whose legitimacy is in question?
- Is it the organization or a specific decision or action whose legitimacy is in question?
- What is the basis of its claim to legitimacy?

In some cases, a distinct organisation will have been established to carry out FRM in whole or in part. In other cases, a more ad hoc composite body will have been established (e.g. a catchment

management committee). The ad hoc body will have been given no clearly defined legislative authority but will take decisions which are then implemented using the powers available to its constituent members. For example, the catchment management committee may include a water management agency and a spatial planning authority. Together they may decide that restrictions on development on flood plains should be introduced and those restrictions are then imposed by the spatial planning authority.

The early work on the legitimacy of the nation state focused on the legitimacy of the agent, the nation state. But it may be argued now that legitimacy is determined on a case by case basis, where the question is about the legitimacy of each specific decision and action.

Legitimacy, the demos and the role of the courts

In modern European societies legitimacy is ultimately determined by the *demos*, through democratic processes, including the 'rule of law' (Dicey 1915); this will be adopted here as a normative assumption. The questions of what constitutes a democracy and whether some states are more democratic than others (e.g. can or should Switzerland, with its system of direct democracy), be considered to be 'more' democratic than, say, the UK or the Netherlands?) will not be considered here. But in discussions of a legal perspective on legitimacy, an issue that has emerged is the relationship between the legislature(s), the court system and the executive(s). In practice, there are at least three different articulations of the relationship between the legislature and the court system:

1. Germany is an example of a country where there is both a written Constitution and a Constitutional court which can determine whether legislation complies with the Constitution;
2. Countries, such as the Netherlands, which have a written Constitution but where the Constitution specifically forbids the courts from making judgments about whether legislation complies with the Constitution;
3. The UK; where there is no a written Constitution and hence no Constitutional court and where the convention is that the courts will not make judgments that over-ride legislation: "*Parliament has the right to make or unmake any law whatever; and further, that no person or body is recognized by the law of England as having a right to override or set aside the legislation of Parliament*" (Dicey 1915; xxxvi).

In turn, the court systems in each country have to varying degrees to make judgments that imply that the law, or some normative principles such as natural justice, is superior to the primary legislation (i.e. that made by the national Parliament in a non-federal country); or rather, since a Constitution is itself legislation, whether one form of legislation is superior to another; or to make a judgment based upon invoking some principle outside of legislation. The result is that there is very limited scope at present for defining *the* legal perspective; rather there is *a* legal perspective specific to individual countries. *The* legal perspective is restricted to the judgements of ECJ and European Court of Human Rights, though even here there are questions whether those judgments are universal across Europe. For example, the question of which takes precedence when European court judgments are inconsistent with the Constitution of Germany is held to be as yet to be determined (Foster and Sule 2002). What seems to be generic is that the courts are prepared to test whether the legislation is properly applied, whether decision makers have the powers claimed to make the decisions or take the actions being challenged. Thus, procedural legitimacy is clearly part of remit of the courts.

Here, we might consider how English Law illustrates some of these generic issues. Courts in England have in some cases refused to make a judgment that extends beyond a straight interpretation of current primary legislation, arguing that it is for the legislature to make wider judgments. For example, in a recent case in which a defendant argued that her beliefs prevented her from removing her niqab in court, the judge stated that he could only decide what was necessary for the proper administration of justice in a court. It was then the business of the legislature to make wider

judgments about the wearing of niqabs in other circumstances began. Conversely, in other countries, courts have been more willing to make normative judgments (De Poorter 2013).

In England, the courts have been willing to reach judgements about the procedural legitimacy of those acting under the powers created by primary legislation e.g. Ministers, municipalities and other public bodies. The courts have been content to reach judgements about whether or not those delegated to have powers had the powers claimed and used them within the limits set. However, they have been reluctant to second-guess Ministers and public bodies in terms of the judgments expressed in the decision by that individual or body. That is, since the legislation creates the power for some executive or secondary legislature to make regulations or decisions, they are reluctant to replace their judgments of what decision should be for the decision made by the decision maker created by the primary legislation.

However, all human systems are also fallible and so some decisions will be made erroneously, incompetently, perversely or the decision maker may seek to use the primary legislation for purposes for which it was not intended. The legislation itself may be poorly drafted creating ambiguities, be insufficiently clear, and create conflicts with other legislation. More widely, legislation creates general principles but these principles then have to be applied to specific cases. Whether legislation can be sufficiently detailed to provide a mechanical rule to dictate the outcome in every possible case may be questioned. Bingham (2011: 7) quotes rather critically a legislative attempt to rule out ambiguity: *“Any reference in these regulations to a regulation is a reference to a regulation contained in these regulations.”* Conversely, if legislation simply refers to rivers, then the question of what is a river has to be decided by the courts.

The result is an ambiguous area where a decision reached by a public actor is argued by the courts to be so perverse as to be outside of the powers created by legislation. A traditional test in English law was the concept of reasonableness – in contrast to irrationality (Woolf et al. 1999); would a reasonable person take the decision made by the actor on the basis of the evidence in front of that actor? German law developed a different but related test: that of proportionality (Foster and Sule 2002). Through European law, the proportionality test has now entered English law (Woolf et al. 1999). The use of the proportionality principle does involve second-guessing the actor whose authority to decide comes from legislation. Again in judicial reviews of decisions by public actors, judges have been known to refer to principles of natural justice, to make what are in effect specifically normative judgments. Here it might be argued that it is the responsibility of the legislature to ensure that the public actors act within the scope and intent of the legislation. However, in many countries, the executive and the legislature are not clearly separated, Ministers being simultaneously part of the legislature. This is not to argue that this is inappropriate but that it raises issues which have to be resolved (and so does separating the executive from the legislature).

Here, judges can make two arguments in support of their use of normative judgments. Firstly, court judgements in these regards may be regarded as tentative since they could be deemed not to be the law by an Act of Parliament which stated that a contrary position should be held in future. Secondly, they may assert that the normative principles on which their judgments were made were universally applicable and those implicit in the nature of a democracy under the rule of law.

Where does this leave us?

- Pan-European law is limited at present; there are different national frameworks of law so there is not *the* legal perspective in formal terms. But these differences offer insights into how legitimacy is interpretable;
- Court law is ultimately subservient to the *demos* and hence to legislation in the first instance;

- At the same time, court law is always about specific instances whereas legislation is more concerned with general principles. A legal perspective can then highlight the issues in applying legislation: what are the limits to a rule? What is extent and reach of a power?
- Procedural legitimacy is separable from normative legitimacy.

Testing legitimacy

There is a growing body of empirical work on legitimacy in different contexts. One of the questions raised is, “is legitimacy context dependent?” For example, do the conclusions from studies of resolution of employment conflicts (e.g. Thibaut and Walker 1975), also apply to spatial planning decisions? In general, the statistical analyses of studies of the legitimacy of the justice system and the police generally find (e.g. Hough et al. in press; Mazerolle et al. 2010) that procedural justice is a major determinant of judgments of legitimacy and the individual’s propensity to cooperate with the police. Here the question of effectiveness has a much weaker influence on those judgments. The problem with regression analyses is that the sample strength of a regression coefficient is a composite of the strength of the individual members of population and the agreement between the members of the population about what, in this case, determines legitimacy. So, there are two possible interpretations of the dominance of procedural justice over effectiveness in these studies. Either, there are differences with the population on the basis by which they determine legitimacy; or alternatively, procedural justice is more important than effectiveness.

In other circumstances, there are differences between identifiable sub-groups within the general population determined not simply by socio-economic, demographic and ethnic differences but also by differences in their wider view of the world. Thus, there is the possibility that what constitutes legitimacy is different between different populations. For instance, there may be some groups who do consider that legitimacy is determined by judgments of effectiveness; although there is a larger population who base their judgments upon assessment of procedural justice. This is a not unreasonable expectation in that a follower of Nozick (1974) might be anticipated to have a different view of when the individual is obliged to consent to the decisions of the collective than, say, someone who adopts a Benthamite (Bentham 1970) utilitarian view of the world.

This discussion is relevant in that Scharpf (1999) and others (Bader 2008), who introduced an approach to the assessment of legitimacy based upon **input, output** and **throughput legitimacy**. Scharpf defined *input legitimacy* as “government by the people” and *output legitimacy* as “government for the people”. Both procedural and substantive justice requirements were then loaded into the input legitimacy requirement whilst output legitimacy was defined in terms of effectiveness. However, the empirical evidence discussed above does not provide support for this conceptualisation of legitimacy; at the very least, that evidence implies that a trade-off between legitimacy and effectiveness is part of the decisions that have to be made.

A second question concerning the input-output definition of legitimacy is whether it is a logical framing. Logic basically requires either an AND or an OR linkage between procedural justice and effectiveness in determining legitimacy. Either legitimacy is created by procedural justice AND effectiveness or by procedural justice OR effectiveness. But the latter formulation implies judgments such as that the fascist regime in Italy was legitimate because Mussolini made the trains run on time. The AND linkage implies judgments that a legal system that is not as effective as alternative systems is illegitimate because it is ineffective.

What is legitimate?

In this modern polycentric society of diffuse and contending power, it can be argued that the question of legitimacy is a different one. The questions now concern who or which interests have a legitimate claim to a share in power? Who is entitled to be a stakeholder and what in each case is the degree of their entitlement to power? How must the assemblages of interest decide what to do and

how may they act? Those involved are generally not representing themselves but some organisation such as an interest group. This is the **'principal-agent' problem**: indeed, is the representative actually presenting the interests of those they claim to represent, and how can this be demonstrated, or their own interests? This is the problem of **accountability** and the task of defining who is accountable, to whom and for what (Lloyd 2005). So, traditionally, firms were only accountable to their owners and had no responsibility to their customers, employees or the wider community. They simply had to obey the law defining what duties they had in those respects. The question of accountability is particularly difficult in regard to interest groups whose interest is an abstraction rather than an identifiable group of people: an environmental group can obviously not be held accountable by the environment whose interests it claims to represent.

Those assemblages do not themselves normally take coercive authority directly (where coercive power includes the non-voluntary payment and the seizure of property, coercion being the use of physical force). But they make use of those coercive powers pre-existing in their membership: the use of tax monies raised by a municipality or central government, the rights to make rules (all rules are more or less coercive otherwise it is not a rule; it is the nature of the coercion that differs), or the right to acquire property compulsorily. The legitimacy of using that coercive authority is then thrown back on to the legitimacy of the use by originating body of that particular usage.

This shift also highlights the question of whether legitimacy is a **characteristic, state or relationship**. Discussions of legitimacy as authority tended to focus upon legitimacy as a characteristic; thus, Weber (1947) gave a series of reasons why authority can exist. Secondly, a common distinction drawn in discussions of legitimacy is between **procedural** and **substantive legitimacy**. Substantive legitimacy tends to imply that legitimacy is either a characteristic or perhaps a state. Discussions of legitimacy distinguish between **descriptive and normative legitimacy**. The former is simply whether an organisation or person has the power and uses it properly. The normative question is who ought to have which powers for what purposes in what circumstances: what is the rationale for so concluding.

If legitimacy is considered to be a relationship, then the questions become: how is legitimacy gained or lost? (Johnson et al. 2006). A second order question concerns whether a particular action is itself legitimate. In legal discussions of legitimacy, the focus is largely upon procedural legitimacy most obviously in the concept of ultra vires: does the party concerned actually have the power or authority to do what they did? An arguably more recent development in English law is the concept of judicial review of decisions by the executive such as Ministers (Bingham 2011; Lord Steyn 1999). This concerns whether they used the power that they have properly both for the purposes intended and by the procedures prescribed. Not infrequently, judges now conclude that a Minister acted 'irrationality', in deciding to act they were required by the legislation to take account of the evidence or seek consultation, but failed to do so. Similarly, judges have ruled that if the Minister's decision implies that the legislation upon which they claim to rely would make lawful a criminal act, then the Minister's decision must be wrong, unless the legislation specifically stated the intention to make lawful an otherwise criminal act. In England, in the lack of a formal constitution, a court has no formal test of whether legislation is itself lawful except the Human Rights Act.

Returning to first order question of how legitimacy is gained or lost, legitimacy is also contested; indeed, Cleaver (2001) has referred to the consequences as 'bricolage'. Notably, within a nation state, the different forms of government contest legitimacy with each other. Those contests may be between different legislatures or between different branches of the executive. For example, in the Bideford sea defence proposal, the Town Council opposed the proposal put forward by the Environment Agency but the EA's proposal was supported by the District Council (Town Councils, which do not exist in most parts of England, are the lowest tier of local government and very limited powers and resources). In this case, the ministry, Defra, and the Environment Agency found

themselves essentially as 'piggies in the middle' and felt compelled to support the views of the District Council. Again, the French Agence Bassin essentially have a tax raising power and the Ministry of Finance wanted to replace this by central taxation and making a grant to the Agence Bassin. The Agences Bassin, generally seen as most effective catchment management agencies, resisted this move; the result was that the Minister of the Environment sent in the police to physically remove the chief executive of the Agence Seine-Normandie from his office. Similarly, whilst the Water Boards in the Netherlands are run by elected representatives, some environmental NGOs have argued that they should have a greater say in the actions of the Water Boards.

But in common with the old concepts of legitimacy, the newer concept retains the focus on the consequence of a conclusion, namely that some 'thing' is legitimate or illegitimate. In both old and new concepts, the apparent implication is that if there is legitimacy then the 'thing' ought to be accepted. Conversely, a conclusion that something is not legitimate eliminates the obligation to comply and creates the entitlement to resist, potentially including the use of physical violence. For example, the US Constitution includes a duty to resist a tyrannical government. Similarly, the Nuremberg doctrine expressed the principle following orders was not a defence; formal authority did not create a duty to comply. Legitimacy is thus different from simple agreement or disagreement; simple disagreement does not remove an obligation to comply.

Such conflicts are not restricted to groups within a 'state' (a term which has proved almost impossible to define within the political science literature) but also within what is now loosely termed 'civic society'. For example, there may be a conflict between a local community, an anglers' association, an environmental NGO, and the local association of farmers. There is almost always a conflict between recreational boaters, canoeists and anglers, notably the last do not want the waters disturbed by boats. Similarly, in what is equally loosely termed 'the market' (which from an economic perspective may be termed systems of voluntary exchanges but which tends to be used instead to refer to the producers and deliverers of goods for private profit), there can be conflicts between small scale farmers, especially tenant farmers, and industrial farmers, and with the producers of seeds. An example of such a conflict was the long history of controversy over the Enclosure of the various Commons in England.

As these latter examples illustrate, the argument over legitimacy is about what *share* (particularly if power is considered to be a zero sum game) of power should be given to different interests. If power is now diffusely distributed, interests are even more widely distributed. In particular, any single individual has multiple interests, and consequently ways in which they may be affected by any decision that is taken. For example, one individual may live in a flood risk area, have children and thus be concerned with their safety, work in a local business, be a member of an Environment NGO, and be a taxpayer. Thus interests will often conflict in a decision. What is then termed 'civil society' are then organisations of interest, interests which they claim to represent.

These multiple potential conflicts of interest may also suggest that the old concept of the 'state' as the co-coordinating or directing body is now out-dated. The modern state, particularly the legislative arm, has instead a role closer to that of a mediator, seeking to resolve the multiple conflicts of interest represented by different bodies of interest such as those of market and civil society. Part of its problem then is how to weigh the different interests and otherwise take account of those interests in decision making. In particular, Mackintosh, both an eminent political scientist and a Member of Parliament, early drew attention to the clash between legislatures elected by universal suffrage and stakeholder engagement.

In this melange of competing interests, legitimacy may be won or lost by how these conflicts are handled. A particular concern is then whether the loss of legitimacy is contagious; whether if one action is construed as being illegitimate then the legitimacy of all actions taken by that agent or

person are now anticipated to be unlikely to be legitimate. Here, a failure is likely to have greater effect than a record of legitimate action. This is to argue that legitimacy in practice is about reputation.

Here, there is considerable overlap between the concepts of legitimacy and of **justice**. For example, Zelditch's (cited in Jost and Major 2001) widely used definition of legitimacy ("*..... something is legitimate if it is in accord with the norms, values, beliefs, practices and procedures accepted by a group*") might almost be used as a definition of justice. It again highlights the inherently contested nature of legitimacy as well as of justice: what those norms and values ought to be is always open to challenge. So, for example, the ownership of slaves was almost globally seen as legitimate until the norm was overturned. Equally, in both discussions of justice and of legitimacy, there is debate about the relative importance of procedural aspects and the outcome.

The concept of a rule of outcome justice, substantive justice, is immensely attractive for many reasons, notably that there is simply a rule, and hence a simple test. There is no scope for argument or debate. The obvious problem with the pursuit of a rule of substantive justice (and equivalently of substantive legitimacy) is that it requires that there is a single rule which will provide a unique answer to every possible instance, including those questions which are not foreseen. Here it may be noted that in English law, the law of Equity developed to deal with those cases where the law was not considered to deliver justice (Worthington 2003). The adequacy of a concept of justice which is restricted to a rule of substantive justice may be illustrated by taking two proposed rules of substantive justice: the Rawlsian criterion and the Hicks-Kaldor Compensation Principle (or Potential Pareto Improvement Principle), the latter being adopted in benefit-cost analysis. In the case of Rawls' criterion, it would be just if men took all decisions provided that they ensured that the outcomes were the same for women as for men. Under the Hicks-Kaldor Compensation Principle, it would be sufficient if men took all decisions but ensured that the outcome was such that the gains to men were sufficient to compensate women for their losses (although it is not necessary to actually pay this compensation).

This throws the problem back on to the procedural aspect: Can justice or legitimacy be established by process? In turn, what are the conditions for such a process? Zelditch's definition of legitimacy obviously does not answer the practical question of what are these norms, values and so forth. However, the question of what constitutes procedural justice is quite well studied (e.g. Thibaut and Walker 1975) and recent work on legitimacy (Mees et al., in press) has adopted terms which echo some of the principles of procedural justice that have been identified (Lind and Tyler 1988). In both cases of justice and legitimacy it is often easier to define each concept by exclusion (i.e. to define what is excluded). For example, the oath of office for judges in the UK includes factors which will not influence the judge's judgement. In seeking to define what constitutes the 'rule of law', Bingham (2011) shows that much of UK constitutional history similarly developed what was to be excluded if justice was to be delivered. Similarly, Levi et al. (2010), in a statistical analysis of social surveys of beliefs about the quality of governance in a number of African countries found that beliefs about the inclusion of factors seen as illegitimate had significant effects upon overall judgments of legitimacy. What may be questioned in their study is the inclusion of *effectiveness* as a determinant of legitimacy. Here, we may draw a distinction between **effectiveness as an indicator of legitimacy and as a determinant of legitimacy**. Ineffectiveness in delivery may be a sign that practices (such as corruption) are being adopted which are illegitimate but an ineffective governance arrangement is not necessarily illegitimate simply because it is ineffective. In collective action, the relevant norms may be argued to be:

- Functionally, the purpose of justice may be argued to allow conflicts to be resolved in a way which critically, as a minimum, avoids the resort to violence but which practically also allows the gains

from collective action to be continued. Hence, non-stakeholders have a key interest in ensuring that assemblages follow the principles of justice;

- The way to resolve conflicts is then to appeal to some higher principles than self-interest since it is self-interest that creates many of the conflicts and where short-term self-interest militates against the longer term benefits of collective action;
- Therefore, justice involves some concept of the collective interest somehow specified (in most of Europe this has been deemed to be 'social solidarity' – specifically so in the case of natural hazards in France. In the UK, it was traditionally specified as the 'national interest'). Acting solely in the interests of decision makers has been interpreted as a sign of illegitimacy. So too is to take account of some interests whilst ignoring others. A consequence is that interested parties are compelled to present arguments for or against some action in terms of the collective interest when their real concern is the effect on their own interests;
- The use of reason, in the sense that Toulmin (1958) defined as argument, is seen, post the Enlightenment, as a norm. Reason here is understood as a logical chain of argument supported by evidence and in which counter-evidence is considered. A failure to demonstrate a reasoned approach is generally seen as demonstrating illegitimacy. Appeals to religious faith are not now given weight and the Greek and Roman practice of seeking auguries would not now be seen as legitimating a decision but was in the Graeco-Roman period;
- Lloyd (1991) argued that justice required that 'like cases be treated alike'. Inconsistency is regarded as strong evidence of illegitimacy since reason is expected to result in consistent outcomes. The argument is then what differences should be ignored and what conversely are essential similarities that constitute 'like' instances. For example, in the UK, until recently partner killings were treated as manslaughter unless there was sign of premeditation in which the case was considered to be murder. However, women's advocates argued that whereas a man might act on the spur of the moment, a woman might have to show some foresight and preparation in order to survive a violent confrontation with a man simply in consequence of the differences in physical strength. Therefore, it was unjust to rely upon apparent foresight as determining whether it should be deemed to be manslaughter or murder;
- Related to consistency is that inappropriate differences should not be taken into account. In particular, individuals (but not therefore interests) are entitled to equal consideration. Wenzel (2001) argued that justice is important because it maintains social identity.

Neither justice nor legitimacy would be useful concepts unless legitimacy or justice in the individual case could be tested. Equally, since humans are fallible, there is a need for a means to challenge any judgement. The obvious exception here in English law is that it is not possible to challenge the decision of a jury and the deliberations of a jury are required to be kept secret. Here, the principle that an individual has a right to judgement by their peers is too important to be challenged by the occasional case whether a jury reached an ill-considered decision. Hence legal systems are characterised by multiple levels of appeal. In the UK, the principle that guilt must be found "beyond all reasonable doubt" in criminal cases, is so important that review extends to cases where the person found guilty is dead. Hence, the person in question will not gain the retrospective finding of not guilty, but the principle is sufficiently critical to the living, for it to be necessary to apply to the dead.

Testing the legitimacy of an individual decision is only possible if there is a clear audit trail: the chain of reasoning that led to the decision can be inspected. This is the practical meaning of the concept of **transparency**. An audit trail requires the chain of argument and both the supporting evidence and the evidence against to be open to scrutiny.

How can we evaluate FRGAs and FRMSs from a legitimacy perspective?

Legitimacy is the most contentious of the selected sub-criteria for appropriateness. Moreover, what may be considered to be legitimate may vary considerably in different contexts. With this in mind, we propose that STAR-FLOOD partners develop a clearer understanding of how legitimacy is defined and the expectations of different actors in relation to FRM.

As shown in the above discussion, legitimacy has been an important issue at least throughout recorded human history and is complex. Fortunately in practice we can adopt the Aarhus Convention’s conceptualisation of legitimacy on the grounds that it constitutes an agreed European norm of legitimacy. Divergence from the principles of the Aarhus Convention would then require special justification and so we do not need to seek to derive criteria and indicators of legitimacy from first principles. The three Aarhus principles also reflect the principles of procedural legitimacy discussed above:

- The right to environmental information;
- The right to be heard;
- The right of appeal.

Hence, **in testing the legitimacy in the case studies, we may ask whether the decisions made, and the processes used, complied with the three Aarhus requirements.** However, it may be noted that both the WFD and the FD were specifically exempted from the provisions of the Aarhus Convention. In particular, in the case of the WFD, the requirement to achieve good ecological quality appears to trump the provisions of the Aarhus Convention. Thus, if the decision process adopted the Aarhus Convention principles and resulted in a decision that good ecological quality should be deferred until some unspecified future date, then the requirement for good ecological quality would appear to overrule the Aarhus requirements. In accordance with these three principles, FRGAs and FRMSs should be evaluated in terms of:

- Access to flood risk information;
- Public participation in decision-making;
- Access to justice and appeals process.

These points are further discussed in Chapter 5.

4.4.7 Potential indicators of legitimacy

There is a literature both on public participation in different forms and on procedural justice as well as on legitimacy. Whilst some of this is based upon social survey responses specifically concerning legitimacy or procedural justice, other material is more conjectural. Public participation has been the target of much study. There are many ways in which public participation can be viewed and evaluated (e.g. Fiorino 1990; Webler 1995; Rasche et al. 2006). Rowe and Frewer (2000) highlight the requirement of two different types of criteria: acceptance criteria and process criteria. They define the former as relating to the “effective construction and implementation of a procedure” (ibid. p11) whereas the latter relates to the public acceptance of any participatory procedure”. They provide a number of sub-criteria from which they define those aspects which need to be considered when evaluating a participatory process (Table 4.11).

Table 4.11 Potential indicators and sub-indicators for evaluating public participatory processes

Acceptance indicator	Process indicator
Criterion of representativeness: The public participants should comprise a broadly representative samples of the populations of	Criterion of resource accessibility: Public participants should have access to the appropriate resources to enable them to successfully fulfil their

the affected publics.	brief.
Criterion of independence: The participation process should be conducted in an independent, unbiased way.	Criterion of task definition: The nature and scope of the participation task should be clearly defined.
Criterion of early involvement: The public should be involved as early as possible in the process as soon as value judgments become salient.	Criterion of structured decision making: The participation exercise should use/provide appropriate mechanisms for structuring and displaying the decision-making process.
Criterion of influence: The output of the procedure should have a genuine impact on policy.	Criterion of cost-effectiveness: The procedure should in some sense be cost-effective.
Criterion of transparency: The process should be transparent so that the public can see what is going on and how decisions are being made.	

After Rowe and Frewer (2000)

These criteria have the advantage of being closely related to those identified by Leventhal (1980) and others (e.g. Lawrence et al. 1997; Thibaut and Walker 1975; Tyler and Lind 1992; Wendorf and Alexander, nd) as being significant in experimental studies of procedural justice. These criteria are:

- Bias suppression/neutrality – applied in manner which is both unprejudiced and without self-interest;
- Accuracy – the procedures succeed in their own terms and are based upon accurate information;
- Correctability – the opportunity to appeal;
- Consistency – in application across like instances;
- Representativeness – all affected should be considered in the decision;
- Ethicality – the decision should be made according to prevailing ethical standards;
- Voice/Process control – are the interested parties given a full voice?
- Standing – are the interested parties respected as people?
- Trust;
- Decision control – do the interested parties have any influence on the decision?

Other literature stresses the importance that the procedure protects the worth and dignity of those involved in the adjudication (Lind and Taylor 1988). But other experimental studies have found that the relative significant of these different criteria vary from case to case (Mitchell et al. 1993).

Rowe and Frewer (2000) have used the criteria in Table 4.11 to evaluate different types of participatory processes at the generic level. They might also provide a good starting framework for investigation in STAR-FLOOD. Researchers therefore need to ensure that within the analytical phase of the research they collect data about how a participation strategy was developed and implemented, any public or professional feedback produced about that procedure and finally the outcomes (in terms of changes to decision-making) of any participation. These criteria can be implemented at the FPRD level (e.g. policy or legislative level – to see the degree to which different actors are able to input and legitimise decisions) but is perhaps most relevant at the case study level where there is often more scope for public involvement in decision-making.

An alternative classification presented by Rasche et al. (2006) also provides an effective framework for assessing the different components of participation and what they call the ‘intensity’ of the process. They have identified six different dimensions which are intended to represent good governance principles: activity, equality, reach, flexibility, power sharing and transparency. Rasche et al. (2006) highlight that the intensity of participation can vary along these dimensions depending on

the participatory approach and each of these might be investigated to obtain a good overall picture of a participatory process. These dimensions are discussed in more detail in Table 4.12.

Table 4.12 Main questions covered by the six intensity dimensions

Intensity dimension	Main question covered
Activity	Do stakeholders have the opportunity to take actively part in the participatory process by uttering opinions and ideas and discussing planning options?
Equality	Do all stakeholders have equal chances to influence the output of a planning procedure?
Transparency	Are stakeholders informed about the project as well as the procedure of the planning process and decision making in a way that enables them to advocate their interests in a competent way?
Power sharing	Does the authority share power with the stakeholders by giving their opinion a formal status in decision –making?
Flexibility	Does participation take place at a time when major aspects of the project design are still open for discussion?
Reach	Is participation limited to a small group of representatives or does it involve major parts of the whole group of stakeholders?

After Rasche et al. (2006; p3).

Rasche et al. (2006) have utilised these intensity dimensions as ‘axes’ from which they have identified different types of participation (e.g. horizontal participation, vertical participation, symbolic participation, focussed consultation) and argue that the approach is one way in which it is possible to “qualitatively discriminate between different intensities of participation in a consistent and comparable way” and therefore might be a suitable tool for use in case-study comparison. This method is discussed further in the forthcoming section.

4.5 Operationalising the concept of Appropriateness: Combining the sub-criteria of legitimacy, effectiveness and efficiency

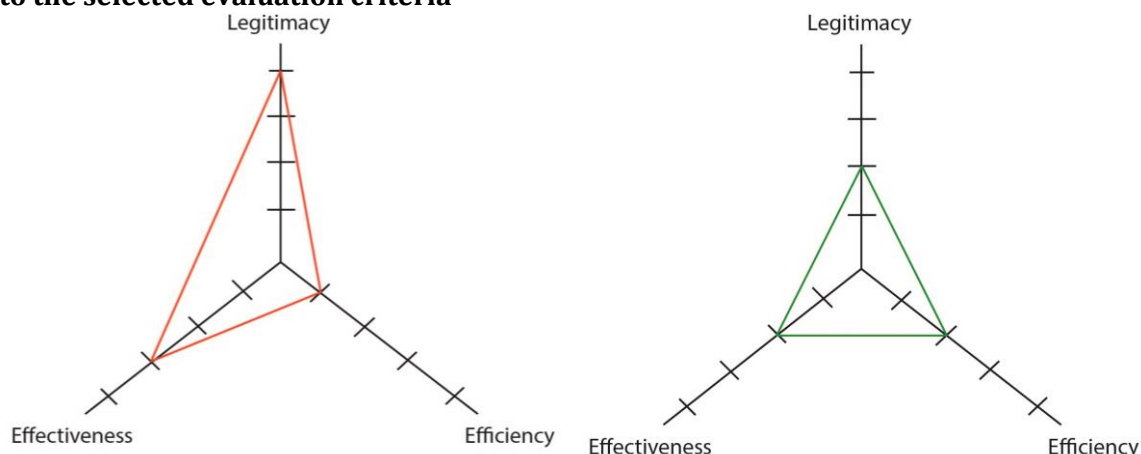
Operationalising an evaluation framework of *Appropriateness* presents the challenge of how to combine the three different sub-criteria. For instance, we could take the position that unless a FRGA completely fulfils each of the criteria then they cannot be considered to be fully “*Appropriate*”. In reality however, a flood risk governance arrangement is likely to fall along a spectrum between not meeting a criteria and full satisfying it: it is unlikely that an optimum scenario of full *Appropriateness* will occur as there is highly likely to be a trade-off between the three criteria. For instance, an approach may end up being more legitimate than efficient, or more effective than legitimate. Therefore, within the evaluation framework it is applicable to best describe the criterion of *Appropriateness* as a **relative concept**, rather than an **absolute one**, with an FRGA being described as more or less appropriate than an alternative FRGA. By combining the sub-criteria together we are able to use these notions to more fully describe the evaluation concepts on which a judgement of *Appropriateness*.

Much of the evaluation of *Appropriateness* can be achieved within a narrative approach and a discussion about the outcomes of analysis of each of the separate sub-criteria. Additionally, as comparative hypothesis are developed these might also consider different measures for testing and ways of combining the different criteria. However, the work by Rasche et al. (2006) provides some inspiration for how we might visualise different combinations of *Effectiveness*, *Efficiency* and *Legitimacy* (Figure 4.5). In their study, these authors propose a strategy to measure the intensity of

participatory approaches across six dimensions, scored according to level of attainment (high, medium or low). The resulting scores are visualised as a radial diagram, referred to by the authors as an “intensity diagram”. Such diagrams serve as tools to support the comparison and evaluation of different forms of participatory processes in water management. Importantly, the authors stress the importance of context in defining appropriate levels of participation and avoid claims of supremacy of one approach over another.

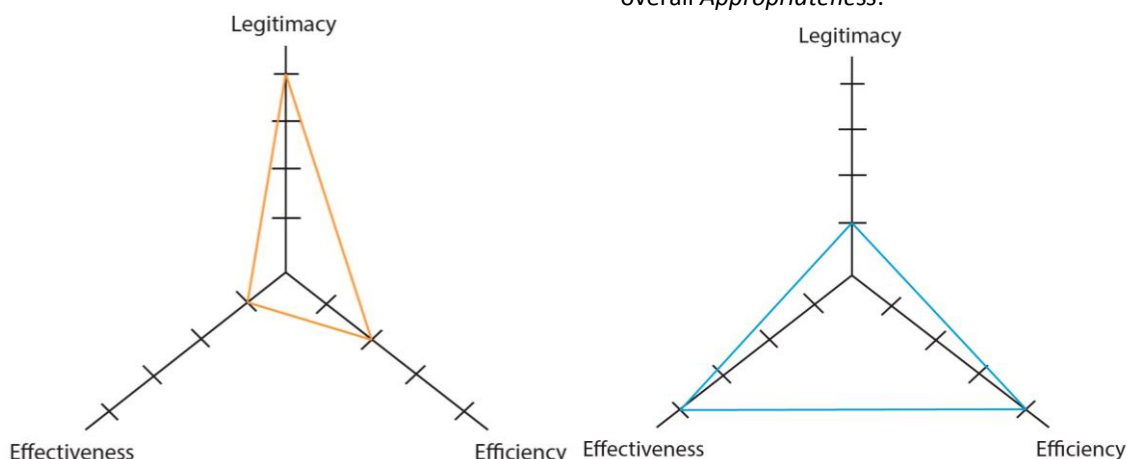
It might be fitting to also designate a minimum level to each of the sub-criteria whereby if any of the criteria are designated below this level then it might be stated that the *Appropriateness* is not positively correlated. For instance, it might be argued that a FRGA that is not effective or has low effectiveness cannot be described as *Appropriate* as it is not fulfilling the goal of managing flood risk. Whether *Appropriateness* is fulfilled in any way when the other two criteria are not met is more of a subjective decision and should be taken based on the norms and context of flood policy and regulations domain being studied. For instance, in terms of the concept of legitimacy, the norm may be to have minimum standard of public participation or there may be few guidelines which need to be followed. These should be investigated through analysis of the legal and policy documentation and also through interviews with flood risk managers.

Figure 4.5: Using radial diagrams to visualise³ the “shapes” of different FRGAs according to the selected evaluation criteria



An arrangement which shows a very high degree of *Legitimacy*, has a high degree of *Effectiveness*, but lacks *Efficiency*. This arrangement might be described as displaying *Legitimately-Effective Appropriateness* characteristics.

A governance arrangement which might be considered to be quite balanced in terms of the outcomes of the sub-criteria; however only displays moderate levels of *Effectiveness*, *Efficiency* and *Legitimacy*. Due to the low levels of attainment in any of the characteristics this approach has low overall *Appropriateness*.



An arrangement which displays elements of being highly legitimate, but has moderate levels of *Efficiency* and low levels of *Effectiveness*. Due to the low levels of *Effectiveness* in achieving the goals of flood risk management it might be questioned whether this governance arrangement is in any way *Appropriate*.

A highly effective and efficient governance arrangement is displayed in this example. However, its *Appropriateness* is reduced by the low levels of *Legitimacy*.

The principles of sound/good/proper administration present a complementary framework for evaluating and drawing together the notion of *Appropriateness*. It has the advantage that more legal perspectives may be demonstrated by examining whether a governance arrangement is acting in accordance with the principles relevant to the country of study. In this respect, STAR-FLOOD researchers are determining whether the norms established within the jurisdiction of interest (i.e. usually at the level of the FRPD) have been complied with and the decisions-made and the actions taken are in accordance with the legal system. It is important to reinforce that these principles need

³ This figure only illustrates one way of presenting the radial diagrams – other approaches might also be equally effective such as representing the different criteria on the different

to be viewed comprehensively as they jointly combine to ensure an outcome that is considered to be ‘proper’ however there are many ways in which the same result can be achieved. An example of the type of principles that might be considered is provided when reviewing the Dutch General Principles of Proper Administration (GPPA) (see Langbroek 2003) (Table 4.13).

Table 4.13: Dominant GPPA depending on the stage of administrative decision making.⁴

Stage of decision-making process	General Principles of Proper Administration		
Preparation	Fair Play	Carefulness	Legality
Balancing of concerned interests	Legitimate expectations	Arbitrariness	Misuse of Power
Content of the ruling	Equality	Legal certainty	Proportionality
Communication	Justification		

Information taken from Langbroek (2003)

In some respects, the GPPA can be considered to be an operationalisation of March and Olsen’s *Logic of Appropriateness* and the principles in Table 4.13 map onto the three sub-criteria for *Appropriateness*. However, with some of the principles there is overlap between sub-criteria. For instance, the principle of Proportionality can relate to both *Efficiency* and *Effectiveness*. Additionally, within the principles there is a strong emphasis on the *Legitimacy* of the process. For use in STAR-FLOOD, researchers might choose to investigate how these concepts have been adopted or implemented in practice and used in the case study locations. Additionally, as part of the evaluation STAR-FLOOD researchers can investigate how the interpretation of the principles may have changed due to a broadening of FRMS? (Including for instance the potential of more interests to balance) and therefore what are the consequences of a broadening of FRMSs on the appropriateness of governance approaches. For countries which share similar principles of sound/proper/good administration it is possible to compare and contrast the appropriateness of FGRAs and the normative position they adopt within administrative law.

4.6 Operational challenges for evaluation

Building upon the previous discussion, this section considers the operational challenges of combining the selected criteria to provide an overall evaluation of FRGAs. Two key challenges are discerned, to which we must remain mindful.

Challenge 1: Aggregating evaluation criteria to inform an overall assessment of FRGAs. Here, we must question how we weight the relative importance of each criterion; i.e. are *Resilience* and *Appropriateness* equally important (equal-weighted)? How should the sub-criteria of *Appropriateness* be weighted? Who should make this decision? Is it necessary for us to be consistent across STAR-FLOOD partners?

⁴Elaboration of the principles:

- *Carefulness*: (1) conduct sufficient research in order to find the information which need to be taken into account for a decision, (2) give interested parties procedural opportunities to offer additional information, (3) identify adequately the relevant facts and interests concerning a decision
- *Fair play*: administrative body needs to perform duties without prejudice; personal interest of administrator should not influence decision
- *Legality*: need to act legal, identification of relevant legislation
- *Legitimate expectations*: a competent legal entity should come up to the legitimate expectations it created
- *Arbitrariness*: authority needs to balance interests involved
- *Misuse of power*: administrative authority shall not use power for other purpose than it was meant for
- *Equality*: all cases should be treated equally
- *Legal certainty*: burdensome ruling should not have retroactive force
- *Proportionality*: content of an administrative order is proportionate to its aim

Challenge 2: Working across different spatial and temporal boundaries of evaluation. Not only must evaluation occur at the Flood Policy and Regulation Domain (FPRD) and case study scale, but we must also negotiate assessments of the selected criteria over different spatial and temporal scales.

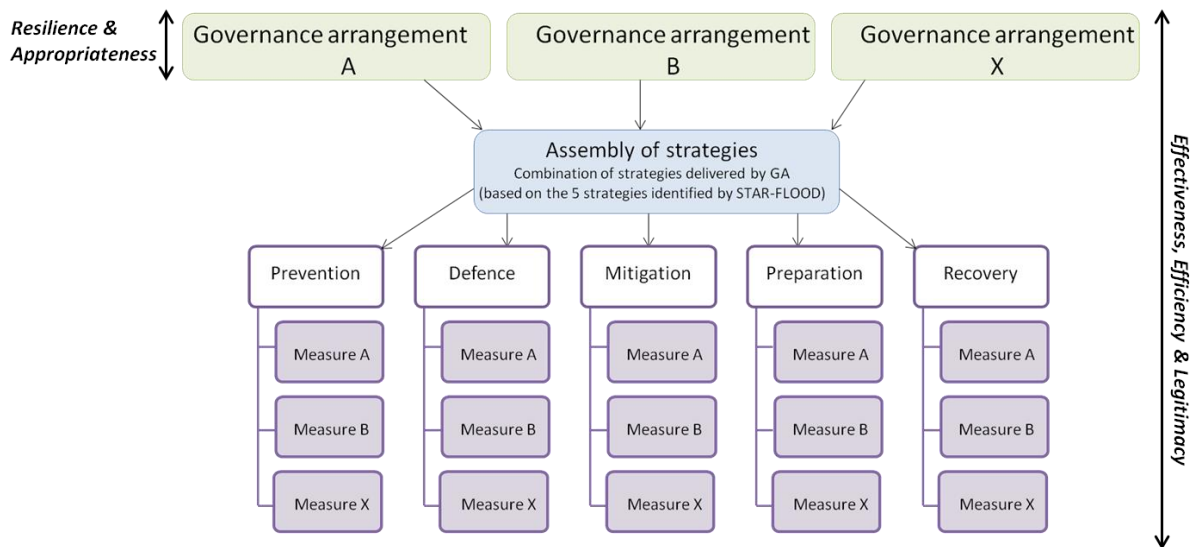
The evaluation framework in Chapter 5 presents a number of criteria upon which FRGAs should be evaluated. The first challenge is how to draw these different criteria together to assess the strengths and limitations of different arrangements. With this goal in mind, it is important to recognise that the criteria and sub-criteria are not able to be used as success criteria in isolation. The ‘findings’ from two or more of the criteria when viewed together may interact in either a positive or negative way. For example, the development of a particular flood risk governance approach may encourage the adoption of measures which may increase the perception of safety from flood risk, which may in turn lead to development which may overall increase the flood risk within an area and reduce an area’s overall resilience. Another example may be the introduction of a flood risk approach which involves a minimal amount of consultation but may therefore arguably have a better economic input to output ratio and be considered to be more efficient. *Resilience* might also be in conflict with *Efficiency*; using resources most efficiently in normal conditions means that there is no reserve to redeploy when a shock occurs. Therefore, within the evaluation framework there is the need to recognise, analyse and understand these complex interactions.

Deciding whether and how we should aggregate criteria is a decision that could be made by relevant stakeholders. Certainly, within the fieldwork stage of our research this is something we could examine and observe what trade-offs and weighted importance is placed upon certain criteria. These questions can be addressed through future iterations of this report. At this stage, it is suggested that STAR-FLOOD researchers approach each criterion and sub-criterion firstly in isolation, before examining how they may be positively or negatively related. Additionally, researchers may be more able to tailor the evaluation approach following the development of comparative hypotheses.

With regards to the second challenge, evaluation of the FRGAs is further elaborated by the different scales of the analysis of approach adopted within the STAR-FLOOD project. Outputs from using the evaluation framework should in WP4-6 enable a comparison and recommendations to be developed at each of the levels: European, Flood Policy and Regulations Domain (FPRD) and case study level. It is expected that within the STAR-FLOOD researchers will undertake a multi-level evaluation of governance and that the analysis and evaluation of FRGAs will be conducted from both a top-down and bottom-up perspective. Therefore, data collection will need to match this multi-level approach. It is also important for evaluation to consider the success of governance arrangements over short as well as long-term temporal scales.

For the purpose of STAR-FLOOD, we have argued that whilst *Resilience* and *Appropriateness* may guide evaluation of FRGAs; the sub-criteria for *Appropriateness* (i.e. *Effectiveness*, *Efficiency* and *Legitimacy*) may transcend arrangements and be applied to the evaluation of measures, strategies and assembly of strategies (as indicated in Figure 4.6).

Figure 4.6: Clarifying the object of evaluation and applicable criteria



The evaluation approach should be flexible to enable, where possible and practical, evaluations of *Resilience* and some of the other concepts beyond the flood framing of the project. For instance, does the adoption of one FRGA and the associated instruments and policies mean that society is more resilient to flooding but less resilient to other hazards or in other ways (e.g. lead to a less cohesive community or more susceptible to water scarcity)? It may be possible for researchers to attempt to evaluate some of these criteria in a more comprehensive and holistic way to see whether attempts to be resilient in one sphere positively or negatively impact upon other notions of resilience. From a European perspective the most attractive and appropriate governance arrangements are likely to be those which maximise resilience across a range of different areas.

A final consideration is that there are two questions which we have not so far addressed. Firstly, discussions on governance arrangements are somewhat abstract. Notably, in making the important distinction between institutions and organisations, between the rules of the game and the players, it can appear that we simply have to get the rules right and the organisational forms will automatically follow. A real question is instead: is the problem the rules or the players? Is it possible to develop organisations which match the rule arrangement that conceptually follows from the nature of the problem? Or are there limitations on organisational design? Given the very large literature on management science and organisational theory, this will also have to be reviewed in parallel to WP3 in order to determine what the constraints upon organisational design are. In turn, in WP3, it is necessary to look for organisational problems. In addition, it is necessary to remember that governance is always produced by people interacting. The quality of governance is consequently ultimately the product of the people involved and the nature of their interaction. The quality of that interaction, and its productivity, depend upon the interaction skills and techniques available to the people involved. So both as an *ex post* and an *ex ante* evaluation criteria, a potential indicator is the emphasis within each organisation on developing and delivering training in these skills.

Secondly, the development of criteria for *ex ante* evaluation of governance arrangements requires the development of causal hypotheses. These hypotheses should be developed in parallel to WP3, both drawing upon and feeding back into the case study analyses.

4.7 Evaluation framework: Summary

The evaluation framework developed here, presents a number of criteria upon which FRGAs and delivered FRMSs should be evaluated. The challenge is how to draw these different criteria together to assess the strengths and limitations of different arrangements. From this, there is a question of how we can facilitate comparisons between different configurations of governance arrangements.

Additional ideas will emerge as we begin to develop comparative hypotheses however this as a starting point to address these challenges, we have looked to existing literature and in particular to the work conducted by Rasche et al. (2006) (see Section 5.8).

In addition, to move from *ex post* to *ex ante* evaluation, we have to develop and test hypotheses as to why particular forms of FRGAs either do or not work and help. WP4 requires not only defining what the appropriate FRGAs are in particular circumstances but stating how this appropriate FRGA can be achieved.

The criteria themselves are often inherently antagonistic rather than complementary or synergistic. It is ultimately for the stakeholders to decide what trade-offs between the criteria in terms of the performance of alternative FRMGAs they choose to make. Perhaps more implicit in the arguments developed here is that we are dealing with complex dynamic systems and rather than criteria that refer to a characteristic or state, relational criteria are needed. Equality is obviously a relationship and so less obviously are legitimacy and justice. That the systems are dynamic also means that procedural issues are the critical issue rather than substantive outcomes. But from a legal perspective, procedural requirements provide a less secure test of compliance than very precise definitions of substantive requirements

Although we focus on two principal criteria, it is recognised that additional criteria may be required and added in the learning process of evaluation (Box 5.1). In addition, that there is a significant difference between *ex post* and *ex ante* evaluation; *ex ante* evaluation will require the development and testing of causal hypotheses as why some FRGAs are more appropriate than others. STAR-FLOOD researchers need to ensure that these issues are considered when tailoring and finalising how they operationalise this evaluation framework.

References

- Aarhus Convention 1998, *Convention on access to information, public participation in decision-making and access to justice in environment matters*, <http://www.unece.org/fileadmin/DAM/env/pp/documents/cep43e.pdf>, 25 June 1998, accessed 2 September 2013
- Abbot, J and Guijt, I 1998, *Changing views on change: Participatory approaches to monitoring the environment*, The International Institute for Environment and Development (IIED), London, UK
- Abrams, P, Borrini-Feyerabend G, Gardner J and Heylings P 2003, *Evaluating Governance: A Handbook to Accompany a Participatory Process for a Protected Area*, Parks Canada and CMWG/TILCEPA, Ottawa
- Adger, N.W and Kelly, M.P 1999, Social vulnerability to climate change and the architecture of entitlements, in *Mitigation and Adaptation Strategies for Global Change* vol. 4, pp. 253-266
- Adger, N.W, 2006, Vulnerability, in *Global Environmental Change* vol. 16 no. 3, pp. 268–281
- Adler, R and Straube, M 2000, Lessons from Large Watershed Programs, in *National Academy of Public Administration Learning from Innovations in Environmental Protection*, National Academy of Public Administration, Washington DC
- Arts, B, 1998, *The political influence of global NGOs. Case Studies on the climate and biodiversity conventions*, International Books, Utrecht
- Arts, B 2000, Global environmental policies: between ‘interstatist’ and ‘transnational’ arrangements’, in: Van Tatenhove, J, Arts, B and Leroy, P (eds), *Political modernization and the environment. The renewal of environmental policy arrangements*, Kluwer, Dordrecht, pp. 117-44
- Arts, B. and Leroy, P (2010) (ed) Institutional dynamics in environmental governance, *Environment and Policy*, vol 47, Springer, Dordrecht, The Netherlands
- Arts, B, Leroy, P and Van Tatenhove, J 2006, Political modernisation and policy arrangements: a framework for understanding environmental policy change, in *Public organization review*, vol. 6, no. 2, pp. 93-106
- Arts, B *et al.*, 2001, *Biologisch groeien?! Een multi-actor analyse van het beleidsarrangement biologische landbouw (2000-2020)*. KU Nijmegen, Milieu en Beleid, Nijmegen
- Aubry, A and Elliot, M 2006, The use of environmental integrative indicators to assess sea-bed disturbance in estuaries and coasts: Application to the Humber Estuary, in *Marine Pollution Bulletin*, vol 53, pp175-185
- Backes, Ch.W, 1999, Towards international water boards and cross-border spatial plans? In Van Hall, A., Drupsteen, Th.G and Havekes, H.J.M, *The State of Water*, Koninklijke Vermande, Lelystad: 179- 194 (in Dutch)
- Bader, V 2008, Complex Citizenship and Legitimacy in Compound Polities (MLPs and MLG): The EU as example, Eurosphere Online Working Paper 05, Amsterdam: Departments of Sociology and Philosophy, University of Amsterdam
- Bakker, M.H.N, Green, C, Driessen, P, Hegger, D, Delvaux, B, Van Rijswick, M, Suykens, C, Beyers, J.C, Deketelaere, K, Van Doorn-Hoekveld, W and Dieperink, C 2013, *Flood Risk Management in Europe: European flood regulation*, STAR-FLOOD Consortium, Utrecht, The Netherlands
- Beck, U 1994, The Reinvention of Politics: Towards a Theory of Reflexive Modernization, in: Beck, U, Giddens, A and Lash, S, *Reflexive Modernization. Politics, tradition and aesthetics in the modern social order*, Polity Press, Cambridge, pp. 1-55
- Bell, H.M. and Tobin, G.A. 2007, Efficient and effective? The 100-year flood in the communication and perception of flood risk, in *Environmental Hazards*, vol. 7, pp. 302-311
- Bentham, J 1970, *An Introduction to the principles of Morals and Legislation*, Methuen, London

- Berkes, F, Colding, J and Folke, C, 2003 (eds), *Navigating social-ecological systems: Building resilience for complexity and change*, Cambridge University Press, Cambridge
- Berkhout, F, Smith, A and Stirling, A 2003, Socio-technological regimes and transition contexts. SPRU Electronic Working Paper. Available online at: www.sussex.ac.uk/Units/spru/publications/imprint/sewps/sewp106/sewp106.pdf
- Bingham, T 2011, *The Rule of Law*, Penguin, London
- Biologica 2004, *Eko-monitor Cijfers en Trends. Jaarrapport 2003*, Biologica, Utrecht
- Blomberg, A.B, and Michiels, F.C.M.A 1997, *Handhaven met effect*, The Hague
- Bohle, H.G, 2001, Vulnerability and criticality: Perspectives from social geography. IHDP Update. 2. 3-5
- Briggs, D, Tantram, D and Scott, P 1996, *Improving information for management and decision making in national parks: The report of the PIMS development project*, Nene Centre for Research, Nene College of Higher Education, Northampton
- Briscoe, J 1995, *The German water and sewerage sector: how well it works and what this means for developing countries*, TWU21, The World Bank: Transportation, Water, and Urban Development Department, Washington DC
- Brock, W.A, Karl-Göran, M. and Perrings, C 2002, Resilience and sustainability: The economic analysis of nonlinear dynamic systems, in Gunderson, L and Holling, C.S (eds) *Panarchy: Understanding transformations in human and natural systems*, Island Press, Washington DC
- Brouwer, S and Biermann, F 2011, Towards adaptive management: examining the strategies of policy entrepreneurs in Dutch water management, *Ecology and Society*, vol. 16, no. 4, pp. 5
- Buckle, P, Marsh, G and Smale, S 2001, Assessing resilience and vulnerability: Principles, strategies and actions, Australian Capital Territory, Emergency Management, Australia
- Buijze, A 2008, On the justification and necessity of legal effectiveness norms, Unpublished Master's Thesis, Utrecht University, Utrecht
- Cadman, T 2012 Evaluating the quality of global governance: A theoretical and analytical approach, Earth System Governance Working Paper no 20, Lund and Amsterdam: Earth System Governance Project, Amsterdam
- Caldwell, R 2003, Models of Change Agency: A Fourfold Classification, in *British Journal of Management*, vol. 14, pp. 131-142
- Capano, G and Howlett, M 2009, Introduction: The Determinants of Policy Change: Advancing the Debate, in *Journal of Comparative Policy Analysis*, vol. 11, no. 1, pp. 1-5
- Capano, G 2009, Understanding policy change as an epistemological and theoretical problem, *Journal of Comparative Policy Analysis*, vol. 11, no 1, pp. 7-31
- Cardona, O.M 2003, The need for rethinking the concepts of vulnerability and risk from a holistic perspective: A necessary review and criticism for effective risk management. [In] Bankoff, G, Frerks, G and Hilhorst, D (ed) *Mapping vulnerability: Disasters, development, development and people*, Earthscan Publishers, London
- Carmine, E.G. and Zeller, R.A 1979, *Reliability and validity assessment*, Sage, London
- Carpenter, S, Walker, B, Anderies, M.J and Abel, N. 2001, From metaphor to measurement: Resilience of what to what?, in *Ecosystems*, vol 4, pp. 265-781
- Carpenter, S.R, 2003, Regime Shifts in Lake Ecosystems: Pattern and Variation. Excellence in Ecology Series 15. Ecology Institute, Oldendorf/Luhe, Germany
- Centre for Coastal Management 1993, *Monitoring the efficiency and effectiveness of the Wet Tropics Management Plan*, The University of New England-Northern Rivers, Lismore
- Centre for Disease Control and Prevention CDC (CDC) 1999, Framework for Program Evaluation in Public Health, in *Morbidity and Mortality Weekly Report (MMWR)*, vol 48, no RR-11, pp. 1-40
- Chambers, W.B 2004, Towards an improved understanding of legal effectiveness of international environmental treaties, in *Georgetown International Environmental Law Review*, vol. 16, 2003-2004, pp. 501-532
- Clark, W.C, Jaeger, J, Corell, R, K-asperson, R, McCarthy, J.J, Cash, D, Cohen, S.J, Desanker, P, Dickson, N.M, Epstein, P, Guston, P.H, Hall, J.M, Jaeger, C, Janetos, A, Leary, N, Levy, M.A, Luers, A,

- MacCracken, M, Melillo, J, Moss, R, Nigg, J.M, Parry, M.L, Parson, E.A, Ribot, J.C, Schnellhuber, H, Schrag, D.P, Seielstad, G.A, Shea, E, Vogel, C, Wilbanks, T.J 2000, *Assessing Vulnerability to Global Environmental Risks, Report of the Workshop on Vulnerability to Global Environmental Change: Challenges for Research, Assessment and Decision making*, Warrenton, Virginia
- Cleaver, F 2001, Institutional bricolage, conflict and cooperation in Usangu, Tanzania, *IDS Bulletin*, vol 32, no 4, pp. 26-35
- Cojocaru, S. 2009, Clarifying the theory-based evaluation, *Revista de cercetare si interventie sociala*, vol. 26, pp. 76-86
- Conway, G 2007, Monitoring the state of the Solent, *Marine Policy*, vol 31, pp. 632-637
- De Poorter, J.C.A 2013, Constitutional Review in the Netherlands: A Joint Responsibility, *Utrecht Law Review*, vol. 9, no. 2, pp. 89-105
- Dicey, A.V, 1915 (reprinted 1982), *Introduction to the study of the law of the constitution*, Indianapolis, Liberty Fund
- Dieperink, C, Green, C, Hegger, D.L.T, Driessen, P.P.J, Bakker, M, Van Rijswijk, M, Crabbé, A, Ek, K 2013, Flood Risk Management in Europe: governance challenges related to flood risk management (report no D1.1.2), STAR-FLOOD Consortium, Utrecht
- Dietz, T., Ostrom, E. and Stern, P.C, 2003, The struggle to govern the commons, in: *Science*, vol. 302, no 5652, pp. 1907-1912
- Disco, C, 2002, Remaking Nature: The ecological turn in Dutch water management, in: *Science, Technology and Human Values*, vol. 27, pp. 206-235
- Downing, T.E and Franklin, S 2004, *Resilience and vulnerability, SEI poverty and vulnerability programme / GECAFS Briefing Paper*, Stockholm Environmental Institute, Stockholm
- Dryzek, J.S 1997, *The politics of the Earth, Environmental Discourses*, Oxford University Press, Oxford, UK
- Dupuis, J and Biesbroek, R, 2013, Comparing apples and oranges: the dependent variable problem in comparing and evaluating climate change adaptation policies, *Global Environmental Change*, in press, <http://dx.doi.org/10.1016/j.gloenvcha.2013.07.022>, accessed 5 November 2013
- Dworkin, R 1986, *Law's Empire*, Fontana, London
- Dyzenhaus, D 2001, *Legality and Legitimacy*, Oxford University Press, Oxford
- European Parliament and the Council (eds) 2007, Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks, *Official Journal of the European Union*, vol L288
- Everitt, B.S 1984, *An Introduction to Latent Variables Models*, Chapman & Hall, London
- Fabienne, P 2010, Political Legitimacy, *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/legitimacy/>, 29 April 20110, accessed 11 September 2013
- Falaleeva, M and Rauschmayer, F 2013, Evaluating environmental governance in a Belarusian World Bank Biodiversity project, in: *Environmental Conservation*, vol. 40, pp. 147-156
- Farrington, J, Turton, C and James, A.J (eds) 1999, *Participatory Watershed Development: Challenges for the Twenty-First Century*, Overseas Development Institute, London
- Fiorino, D.J, 1990, Citizen participation and environmental risk: A survey of institutional mechanisms, in: *Science, Technology, and Human Values*, vol. 15, no. 2, pp. 226-43
- Folke, C. 2006, Resilience: The emergence of a perspective for social-ecological systems analyses, in: *Global Environmental Change*, vol. 16, no. 3, pp. 253-267
- Folke, C, Berkes, F and Colding, J 1998, Ecological practices and social mechanisms for building resilience and sustainability, in: Berks, F. and Folke, C. (eds) *Linking social and ecological systems: Management practices and social mechanisms for building resilience*, Cambridge University Press, Cambridge
- Foster, N and Sule, S 2002, *German Legal System and Laws, 3rd ed.*, Oxford University Press, Oxford
- Friedman, M 1957 *A Theory of the Consumption Function*, Princeton University Press, New Jersey
- Friedman, M 1980, *Free to Choose*, Penguin, Harmondsworth

- Gallopín, G.C 2006, Linkages between vulnerability, resilience, and adaptive capacity, in: *Global Environmental Change*, vol 16, pp. 293-303
- Garcia, M 2011, *Micro-Methods in Evaluating Governance Interventions*, BMZ Evaluation Division: Evaluation Working Papers, Federal Ministry for Economic Cooperation and Development, Bonn
- Garmestani, A.S, Allen, C.R and Cabezas, H 2008, Panarchy, adaptive management and governance: Policy options for building resilience, in: *Nebraska Law Review*, vol. 87, no. 4, pp. 1036-1054
- Geels, F.W, 2002, Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case study, *Research Policy*, vol. 31, pp. 257-1273
- Geels, F.W 2005, The dynamics of transitions in socio-technical systems: A multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860-1930), in: *Technology Analysis & Strategic Management*, vol. 17, no. 4, pp. 445-476
- Giddens, A 1984, *The constitution of society*, Polity Press, Cambridge
- Glasbergen, P 1989, *Beleidsnetwerken rond milieuproblemen*, VUGA Uitgeverij, Den Haag
- Godfroid, A.J.A and Nelissen, N.J.M (eds), 1993, *Verschuivingen in de besturing van de samenleving*, Coutinho, Bussum (in Dutch)
- Green, C, Dieperink, C, Ek, K, Hegger, D.L.T, Pettersson, M, Priest, S, Tapsell, S, 2013, *Flood Risk Management in Europe: the flood problem and interventions (report no D1.1.1)*, STAR-FLOOD Consortium, Utrecht, the Netherlands
- Green, C.H, 2003, *Handbook of Water Economics*, John Wiley, Chichester
- Gunderson, L and Holling, C.S 2002, *Panarchy: Understanding transformations in human and natural systems*, Island Press, Washington DC
- Hajer, M 1995, *The politics of environmental discourse. Ecological modernization and the policy process*, Oxford University Press, Oxford
- Hajer, M 2003, Policy without polity? Policy Analysis and the Institutional Void, in: *Policy Sciences*, vol. 36, pp. 175-195
- Hanley, N and Spash, C 1993, *Cost-benefit analysis and the environment*, Edward Elgar, Cheltenham
- Hanley, N and Black, A.R 2006, Cost benefit analysis and the Water Framework Directive in Scotland, in: *Integrated Environmental Assessment and Management*, vol. 2, no. 2, pp. 156-165
- Hart, H.L.A and Honore, T 1990, *Causation in the Law*, 2nd edition, Oxford: Clarendon Press
- Haverland, M, 1999. *National autonomy, European integration and the politics of packaging waste*, Thela-thesis, Amsterdam
- Hegger, D.L.T, Green, C, Driessen, P, Bakker, M, Dieperink, C, Crabbé, A, Deketelaere, K, Delvaux, B, Suykens, C, Beyers, J.C, Fournier, M, Larrue, C, Manson, C, Van Doorn-Hoekveld, W, Van Rijswijk, M, Kundzewicz, Z.W and Goytia Casermeiro, S, 2013, *Flood Risk Management in Europe: Similarities and Differences between the STAR-FLOOD consortium countries*, STAR-FLOOD Consortium, Utrecht
- Hegtvædt, K.A and Johnson, C 2000, Justice Beyond the Individual: A Future with Legitimation, in: *Social Psychology Quarterly*, vol. 63, no. 4, pp. 298-311
- Herweijer, M. 2007, Effectieve bestuursbesluiten, in: *Ars Aequi*, vol. 11, pp.895-903
- Hofer, K 2000, Labelling of organic food products, in: Mol, A.P.J, Lauber, V and Liefferink, D (eds), *The voluntary approach to environmental policy. Joint environmental policy-making in Europe*, Oxford University Press, Oxford, pp. 156-91
- Holling, C.S 1973, Resilience and stability of ecological systems, in: *Annual Review of Ecology and Systematics*, vol. 4, pp. 1-23
- Holling, C.S 2001, Understanding the complexity of economic, ecological and social systems, *Ecosystems*, vol. 4, pp. 390-405
- Holling, C.S and Meffe, G.K 1996, Command and control and the pathology of natural resource management, in *Conservation Biology*, vol. 10, pp. 328-337
- Holling, C.S 1996, Engineering resilience versus ecological resilience. In: Schulze, P. (Ed.), *Engineering Within Ecological Constraints*, National Academy Press, Washington DC, pp. 31-44

- Holling, C.S, Gunderson, L.H and Peterson, G.D 2002, Sustainability and panarchies, in: Gunderson, L and Holling, C.S (eds), *Panarchy: Understanding transformations in human and natural systems*, Island Press, Washington DC
- Hooper, B 2005, *Integrated River Basin Governance: Learning from International Experience*, IWA Publishing, London
- Hough M, Jackson J and Bradford, B (in press), Trust in justice and the legitimacy of legal authorities: topline findings from a comparative European study, in: *Policing: A Journal of Policy and Practice*
- Howlett, M and Cashore, B 2009, The dependent variable problem in the study of policy change: understanding policy change as a methodological problem, *Journal of Comparative Policy Analysis*, vol. 11, no. 1, pp. 33-46
- Huberts, L and Kleinnijenhuis, J (eds) 1994, *Methoden van invloedsanalyse*, Boom, Meppel (in Dutch)
- Hughes, T.P 1987, The evolution of large technological systems, in: *The social construction of technological systems: new directions in the sociology and history of technology*, Bijker, W.E, Hughes, T.P, Pinch, T (Eds), MIT Press, Cambridge
- Huitema, D, Lebel, L and Meijerink, S 2011, The strategies of policy entrepreneurs in water transitions around the world, *Water Policy*, vol. 13, pp. 717-733
- Imperial, M.T and Hennessey, T 2000, *Environmental Governance in Watersheds*, National Academy of Public Administration, Washington DC
- Johnson, C, Dowd, T.J and Ridgeway, C.L 2006, Legitimacy as a Social Process, *Annual Review of Sociology*, vol. 32, pp. 53-78
- Jost, J.T and Major, B (eds) 2001, *The Psychology of Legitimacy: Emerging Perspectives on Ideology, Justice, and Intergroup Relations*, Cambridge University Press, Cambridge
- Kasperson, J.X, Kasperson, R.E, Turner, I.I.B.L., Schiller, A and Hsiel, W.H 2005, Vulnerability to global environmental change, in: Kasperson, J.X and Kasperson, R.E (eds) *Social contours of risk*, Earthscan, London, pp. 245-285
- Keesen, A.M, Hamer, J.M, Van Rijswijk, H.F.M.W and Wiering, M 2013, The concept of resilience from a normative perspective: examples from Dutch adaptation strategies, *Ecology and Society*, vol. 18, no 2, pp. 45
- Keohane, Robert O 2007, Governance and Legitimacy. Keynote Speech Held at the Opening Conference of the Research Center (SFB) 700 (with comments by Fritz W. Scharpf), SFB-Governance Lecture Series, No. 1, DFG Research Center(SFB) 700, Berlin, February 23, 2007
- Kickert, W, Klijn, E and Koppejan, J (eds), 1997, *Managing Complex Networks*, Strategies for the Public Sector, SAGE, London
- Kingdon, J 1984, *Agendas, Alternatives, and Public Policies*, Little, Brown, Boston
- Lake, D.A 2006, Relational Authority in the Modern World: Towards a Positive Theory of Legitimacy, paper prepared for the workshop on Legitimacy in the Modern World, University of California, San Diego
- Langbroek, P.M 2003, General principles of proper administration and the General Administrative Law Act in The Netherlands, for the World Bank Workshop on Regulating Citizen-State Interactions: Administrative Law in the United Kingdom and the Netherlands, Washington, US, 23 January 2003
- Lawrence, R.L, Daniels, S.E and Stankey, G.H 1997, Procedural Justice and Public Involvement in Natural Resource Decision Making, in: *Society and Natural Resources*, vol. 10, pp. 577-589
- Leventhal, G.S 1980, What should be done with equity theory? New approaches to the study of fairness in social relationships, in: Gergen, K, Greenberg, M and Willis, R (eds), *Social exchanges: Advances in theory and research*, Plenum, New York
- Levi, M, Sacks, A and Tyler, T 2010, Conceptualizing Legitimacy, Measuring Legitimizing Beliefs, in: *American Behavioral Scientist*, vol. 53, no. 3, pp. 354-375
- Lieverink, D 2006, The dynamics of policy arrangements: turning round the tetrahedron, in: Arts, B, and Leroy, P (eds.), *Institutional dynamics in environmental governance*, Springer, Dordrecht, pp. 45-68

- Lind, E.A and Tyler, T.R 1988, *The social psychology of procedural justice*, Plenum, New York
- Lloyd, D 1991, *The Idea of Law*, Penguin, Harmondsworth
- Lloyd, R 2005, *The Role of NGO Self-Regulation in Increasing Stakeholder Accountability*, One World Trust, London
- LNv (Ministerie van Landbouw, Natuurbeheer en Visserij) 2000, *Een biologische markt te winnen. Beleidsnota biologische landbouw*, Ministerie van LNv, Den Haag
- Lockwood, M., Davidson, J., Curtis, A., Stratford, E. and Griffith, R. (2010) 'Governance principles for natural resource management', *Society and Natural Resources*, vol 23, no 10, pp986-1001
- Lord Steyn 1999, *The Constitutionalisation of Public Law*, www.ucl.ac.uk/spp/publications/unit-publications/38.pdf, May 1999, accessed 11 September 2013
- Lorenz, D.F 2013, The diversity of resilience: Contributions from a social science perspective, in: *Natural Hazards*, volume 67, no, pp. 7-24
- Mackintosh, J.P 1977, The Politician's View of Planning, in: Sewell W R D and Coppock J T (eds.) *Public Participation in Planning*, John Wiley, Chichester
- March, J and Olsen, J 1989, *Rediscovering institutions*, Free Press, New York
- March, J.G and Olsen, J.P 2008, The logic of appropriateness, in: Moren, M, Rein, M and Goodin, R.E (eds) *The Oxford Handbook of Public Policy*, Oxford University Press, Oxford, pp. 689-708
- March, J.G and Olsen, J.P 2008, *The logic of appropriateness*, ARENA working papers, WP 04/09
- Marsh, D and Rhodes, R.A.W 1992, *Policy networks in British government*, Clarendon Press, Oxford
- Margoluis, R and Salafsky, N 1998, *Measures of Success: Designing, Managing and Monitoring Conservation and Development Projects*, Island Press, Washington, DC
- Mark, M 2005, Evaluation Theory or What Are Evaluation Methods for?, in: *Evaluation Exchange*, vol. XI, no 2, <http://www.hfrp.org/evaluation/the-evaluation-exchange>, accessed 11 September 2013
- McFadden, L, Green, C and Priest, S 2008, *Social science indicators for Integrated Coastal Zone Management (ICZM)*, Spicosa Project Report, London, Flood Hazard Research Centre, Middlesex University, http://www.coastal-saf.eu/design-step/support/Social_science_indicators_for_ICZM.pdf, accessed 11 September 2013
- McLaughlin, P and Dietz, T 2008, Structure, agency and the environment: Toward an integrated perspective on vulnerability, in: *Global Environmental Change*, vol. 18, no. 1, pp. 99-111
- Mees, H.L, Driessen, P.P.J and Runhaar, H.A.C (in press), Legitimate adaptive flood risk governance beyond the dikes: the cases of Hamburg, Helsinki and Rotterdam, in: *Regional Environmental Change*, Springer-Verlag, Berlin Heidelberg
- Ménard, C and Saleth, R.M 2011, The effectiveness of alternative water governance Arrangements, in: Young, M (ed) *Towards a Green Economy*, United Nations Environment Programme, http://hal.archives-ouvertes.fr/docs/00/62/42/50/PDF/Menard-Saleth--Governance_in_Water_Sector-2011.pdf, 23 September 2011, accessed 2 September 2013
- Michiels, F.C.M.A (ed) 1998, *Recht op het doel af; Opstellen over doelmatigheid en doeltreffendheid van het overheidshandelen in de democratische rechtsstaat*, Tjeenk Willink, Deventer
- Ministry of Agriculture, Nature Conservation and Food Quality, 2003, *Lonkend Rivierenland*, National Forest Service (in Dutch)
- Ministry of the Interior and Kingdom Relations 2009, *Netherlands Code for Good Public Governance*, Ministry of the Interior and Kingdom Relations, The Hague
- Ministry of Transport, Public Works and Water Management, 2000, *A Different Approach to Water, Water Management Policy in the 21st Century*. [English version 2001]
- Mitchell, G, Tetlock, P.E, Mellers, B.A and Ordóñez, L.D 1993, Judgments of Social Justice: Compromises Between Equality and Efficiency, in: *Journal of Personality and Social Psychology* vol. 65, no. 4, pp. 629-639
- Modigliani, F 1966, The life cycle hypothesis of savings, the demand for wealth and the supply of capital, in: *Social Research*, vol. 33, no. 2, pp. 160-217

- Moench, M, Dixit, A, Janakarajan, S, Rathore, M.S and Mudrakartha, S 2003, *The Fluid Mosaic: Water Governance in the Context of Variability, Uncertainty and Change*, Institute for Social and Environmental Transition, Boulder CO
- Myhill, A and Quinton, P 2011, *It's a fair cop? Legitimacy, public cooperation, and crime reduction*, National Policing Improvement Agency, London
- Newig, J and Fritsch, O 2009, Environmental Governance: Participatory, Multi-Level – and Effective?, in: *Environmental Policy and Governance*, vol. 19, pp. 197–214
- North, D.C 1990, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge
- Nozick, R 1974, *Anarchy, State and Utopia*, Blackwell, Oxford
- O'Connor, M 2006, Building knowledge partnerships with ICT? Social and technical conditions for conviviality, in: Guimarães Pereira, Â, Guedes Vaz, S and Togentti, S (eds) *Interfaces Between Science and Society*, Greenleaf Publishing
- OECD 1991, *Principles for evaluation of development assistance*, Development Assistance Committee, OECD, Paris
- OECD 1999, Evaluating country programmes, Evaluation and Aid Effectiveness, Vienna Workshop, 11-12 March 1999, OECD, New York
- OECD 2001, *Evaluation feedback for effective learning and accountability*, Development Assistance Committee, Evaluation and Aid Effectiveness 5, OECD, Paris
- OECD 2010, *Evaluating development co-operation: Summary of key norms and standards*, Second edition, OECD DAC Network on Development Evaluation, <http://www.oecd.org/development/evaluation/dcdndep/41612905.pdf>, accessed 11 September 2013
- OECD 2011, *Quality standards for Development Evaluation*, DAC Guidelines and Reference Series, <http://www.oecd.org/dac/evaluation/dcdndep/49334307.pdf>, accessed 11 September 2013
- OECD 2013, *Evaluating Development Activities: 12 Lessons from the OECD DAC*, <http://www.oecd.org/dac/peer-reviews/12%20Less%20eval%20web%20pdf.pdf>, accessed 11 September 2013
- Official Journal of the European Communities 2000b, *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy*, L 327/1
- Oliver-Smith, A 2004, Theorising vulnerability in a globalised world: A political ecological perspective, in: Bankoff, G, Frerks, G and Hillhorst, D (eds), *Mapping vulnerability: Disasters, development and people*, Earthscan, London
- Olsson, J and Hysing, E 2012, Theorizing inside activism: understanding policymaking and policy change from below, *Planning Theory & Practice*, iFirst article, pp. 1-17
- Ostrom, E 2006, The institutional analysis and development framework in historical perspective, prepared for delivery at the panel on *Starting from Here: Understanding the Context of Development and Democratization – From Nineteenth-Century Theory to Twenty First-Century Practice*, 2006 Annual Meeting of the American Political Science Association, 30 August - 3 September 2006
- Paavola, J 2008, Science and social justice in the governance of adaptation to climate change, *Environmental Politics*, vol. 17, pp.644-659
- Paredis, E 2009, Socio-technische systeeminnovaties en transitie: van theoretische inzichten naar beleidsvertaling. Working paper n° 10. CDO/UGent: Steunpunt Duurzame Ontwikkeling
- Transboundary Freshwater Dispute Database (TFDD), 2013. Online database available at <http://www.transboundarywaters.orst.edu>, last accessed May 2013
- Pelling, M 1998, Participation, social capital and vulnerability to urban flooding in Guyana, in *Journal of International Development*, vol. 10, pp. 469-486
- Pena, H and Solanes, M 2003, *Effective Water Governance in the Americas: A Key Issue*, Global Water Partnership, Stockholm
- Pimm, S.L 1984, The complexity and stability of ecosystems, in: *Nature*, vol. 307, pp. 321-326

- Posner, R.A 1995, *Overcoming Law*, Harvard University Press, London
- Rasche, K, Krywkow, J, Newig, J and Hare, M 2006, Assessing the intensity of participation along six dimensions, proceedings of the PATH Conference, 4-7 June 2006, Edinburgh, Scotland, http://www.macaulay.ac.uk/PATHconference/PATHconference_proceeding_ps3.html, accessed 2 September 2013
- Rayner, J 2009, Understanding Policy Change as a Historical Problem, in *Journal of Comparative Policy Analysis: Research and Practice*, vol. 11, no. 1, pp. 83-96
- Real-Dato, J 2009, Mechanisms of policy change: a proposal for a synthetic explanatory framework, *Journal of Comparative Policy Analysis*, vol. 11, no. 1, pp. 117-143
- Remmelzwaal, A and Vroon, J 2000, *Werken met water: Veerkracht als strategie*, RIZA Lelystad
- Rhodes, R.A.W 1986, *The national world of local government*, Allen and Unwin, London
- Rittberger, V (eds), 1993, *Regime theory and international relations*, Clarendon Press, Oxford
- Rogers, P and Hall, A.W 2002, *Effective Water Governance*, TEC Background Paper No 7, Global Water Partnership
- Rossi, P, Freeman, H and Lipsey, M 1999, *Evaluation: A systematic approach*, Sage, Thousand Oaks CA
- Rowe, G and Frewer, L 2000, Public Participation Methods: A Framework for Evaluation, in: *Science, Technology and Human Values*, vol. 25, no. 3, pp. 3-29
- Ruhl, J.B 1999, The co-evolution of sustainable development and environmental justice: Cooperation, then competition, then conflict, in: *Duke University Law and Policy Forum*, vol. 9, 1998-1999, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1354736, accessed 11 September 2013
- Sabatier, P. and Jenkins-Smith, H.C 1988, An advocacy Coalition Model of Policy Change and the Role of Policy Oriented Learning Therein, in: *Policy Sciences*, vol. 21, pp. 129-168
- Sabatier, P.A and Jenkins-Smith, H.C 1993, *Policy change and learning: an advocacy coalition approach*, Westview Press, Boulder (CO)
- Sabatier, P and Weible, C.M 2007, The Advocacy Coalition Framework: Innovations and Clarifications, in: Sabatier, P.A (ed) *Theories of the policy process*, Westview Press, Davis CA
- Sabatier, P.A, Focht, W, Lubell, M, Trachtenberg, Z, Vedlitz, A and Matlock, M (eds) 2005, *Swimming Upstream: Collaborative Approaches to Water Management*, MIT Press, London
- Scharpf, F.W 1999, *Governing in Europe. Effective and Democratic?*, Oxford: Oxford University Press
- Smit, B and Wandel, J 2006, Adaptation, adaptive capacity and vulnerability, in: *Global Environmental Change*, vol. 16, pp. 282- 292
- Smits, A.J.M, Nienhuis, P.H and Leuven, R.S.E.W (eds), 2000, *New approaches to river management*, Backhuys Publishers, Leiden
- Steinführer, A, Kuhlicke, C, De Marchi, B, Scolobig, A, Tapsell, S and Tunstall, S 2009, *Local communities at risk from flooding. Social vulnerability, resilience and recommendations for flood risk management in Europe*, final report for FLOODsite, Winterwork, Grimma
- Stufflebeam, D.L and Shinkfield, A. J 2011, *Evaluation Theory, Models and Applications*, John Wiley, London
- Talwar, S, Wiek, A and Robinson, J 2011, User engagement in sustainability research, in: *Science & Public Policy*, vol. 38, no. 5, pp. 379-390
- Therborn, G, 1982, *The ideology of power and the power of ideology*, London, NLB
- Thibaut, J and Walker, L 1975, *Procedural Justice: A Psychological Analysis*, Erlbaum, Hillsdale NJ
- Thomas, C.W, 2008, Evaluating the Performance of Collaborative Environmental Governance, paper prepared for presentation at the *Consortium on Collaborative Governance Mini-Conference*, Santa Monica, April 10-12, 2008, <http://stuff.mit.edu/afs/athena.mit.edu/dept/cron/project/urban-sustainability/USA%20Resources/Logic%20Model-%20RECO%20example/Thomas.pdf>, accessed 11 September 2013
- Toulmin, S 1958, *The Uses of Argument*, Cambridge University Press, Cambridge
- True, J.L, Jones, B.D and Baumgartner, F.R 2007, Punctuated-equilibrium theory: explaining stability and change in public policymaking, in: Sabatier, P.A. (ed) *Theories of the policy process*, Westview Press, Davis CA

- Tyler, T.R and Lind, E.A 1992, A Relational Model of Authority in Groups, in: *Advances in Experimental Social Psychology*, vol. 115, pp. 150-162
- Tyler, T.R 2006, Psychological Perspectives on Legitimacy and Legitimation, in: *Annual Review of Psychology*, vol. 57, pp. 375-400
- Underdal, A 2002, One Question, Two Answers, in: Miles, E.L, Underdal, A, Andresen, S, Wettestad, J, Birger Skaereth, J and Carlin, E.M (eds), *Environmental Regime Effectiveness: Confronting Theory with Evidence*, The MIT Press, Cambridge, Massachusetts and London
- UNDP 2004, Glossary of Terms, in United Nations Development Programme Global Report, *Reducing Disaster Risk: A Challenge for Development*, http://www.undp.org/content/dam/undp/library/crisis%20prevention/disaster/asia_pacific/Reducing%20Disaster%20risk%20a%20Challenge%20for%20development.pdf, accessed 2 September 2013
- United Nations Development Programme (UNDP) 1997, *Governance for Sustainable Human Development*, UNDP, New York
- United Nations Development Programme (UNDP) 2004, *Water Governance for Poverty Reduction*, UNDP, New York University: Research, Counsel, Education", Publication Series of UNU-EHS, n.4/2006
- Valman, M 2012, Institutional stability and change in the Baltic Sea: 30 years of issues, crises and solutions, in: *Marine Policy*, <http://dx.doi.org/10.1016/j.marpol.2012.05.019>
- Van de Bilt, S, 2004, *Eén terrein, waar twee beheren*, MSc-thesis Political Sciences of the Environment, Radboud University Nijmegen (in Dutch)
- Van de Ven, G.P, 2004, *Man-made Lowlands. History of water management and land reclamation in the Netherlands*, Matrijs, Utrecht
- Van Eeten, M, 1999, *Dialogues of the deaf, defining new agendas for environmental deadlocks*, Eburon, Delft
- Van Hemert, M, 1999, Ruimte voor de ingenieur, rivierbeheer in Nederland eind jaren negentig, in: *Kennis en Methode, Tijdschrift voor Empirische Filosofie*, vol. 23, pp. 361-387
- Van Leussen, W, 2002, *Leven met water, vermaatschappelijking van het waterbeheer – consequenties voor de civiel ingenieur en voor waterbeheerorganisaties*, Inaugural adres Universiteit Twente, Enschede (in Dutch)
- Van Rijswijk, M and Salet, W 2012, Enabling the contextualization of legal rules in responsive strategies to climate change, *Ecology and Society*, vol. 17, no 2, p. 18
- Van Steen, P.J.M and Pellenbarg, P.H, 2004, Water management challenges in the Netherlands, in: *Tijdschrift voor Economische en Sociale Geografie*, vol. 95, pp. 590-598
- Van Stokkom, H.T.C, Smits, A.J.M and Leuven, R.S.E.W 2005, Flood defense in the Netherlands: a new era, a new approach, in: *Water International*, vol. 30, pp. 76-87
- Van Tatenhove, J, Arts, B and Leroy, P (eds) 2000, *Political modernization and the environment. The renewal of environmental policy arrangements*, Kluwer, Dordrecht
- Van Wijk, H.D, Konijnenbelt, W and Van Male, R.M 2005, *Hoofdstukken van bestuursrecht*, 13th ed. Elsevier, Den Haag
- Walker, B, Holling, C.S, Carpenter, S.R and Kinzig, A 2004, Resilience, adaptability and transformability in social- ecological systems, in: *Ecology and Society*, vol. 9, no. 2
- Walker, B.H, Meyers, J.A, 2004, Thresholds in ecological and social-ecological systems: a developing database, in: *Ecology and Society*, vol. 9, no. 2, p. 3, Available online <http://www.ecologyandsociety.org/vol9/iss2/art3/>, accessed 18 October 2013
- Warren, P 1998, *Developing Participatory and Integrated Watershed Management*, Food and Agricultural Organisation (FAO), Rome
- Weber, M 1947, *The Theory of Social and Economic Organisation*, Oxford University Press, New York
- Webler, T 1995, "Right" discourse in citizen participation: an evaluative yardstick, in: Renn, O, Webler, T and Wiedemann, P (eds) *Fairness and competence in citizen participation: Evaluating models for environmental discourse*, Kluwer Academic, Dordrecht, pp35-86

- Wendorf, C.A and Alexander, S (nd), *Distributive and Procedural Justice in a Moral Reasoning Context*, Wayne State University
- Wenzel, M 2001, The impact of outcome orientation and justice concerns on tax compliance: the role of taxpayers' identity, Working Paper No 6, Canberra: Centre for Tax System Integrity, Australian National University
- Westley, F, Carpenter, S.R, Brock, W.A, Holling, C.S and Gunderson, L.H 2002, Why systems of people and nature are not just social and ecological systems, in: Gunderson, L and Holling, C.S (eds), *Panarchy, understanding transformations in natural and human systems*, Island Press, Washington DC
- Wiering M, Arts, B 2006, Discursive shifts in Dutch river management: 'deep' institutional change or adaptation strategy? In: *Hydrobiologia*, vol. 565, pp. 327-338
- Wiering, M.A and Driessen, P.P.J 2001, Beyond the art of diking. Interactive policy on river management in the Netherlands, in: *Water Policy*, vol. 3, pp. 283-296
- Wiering, M.A and Immink, I.F.G, 2003, Nieuwe beleidsarrangementen voor waterbeheer en ruimtelijke ordening? In: Van der Vlist, M and Hidding, M (eds), *Water en Ruimte*, The Hague, SDU uitgevers bv (in Dutch)
- Wilson, T 2008, *Defining and Mapping Societal Vulnerability and Resilience: A Literature Review. Deliverable 3.7a for FRMRC2*, from FHRC at Middlesex University
- Winch, D.M 1971, *Analytical Welfare Economics*, Penguin, Harmondsworth
- Wisner, B, Blaikie, P, Cannon, T and Davis, I 2004, *At Risk: Natural Hazards, People's Vulnerability and Disasters*, Routledge, New York
- Wisserhof, J, 2000, *Inventarisatie kennisinfrastructuur*, The Hague, Nationale Raad voor Landbouwkundig Onderzoek, Adviesraad voor het Wetenschaps- en Technologiebeleid, Raad voor het Milieu- en Natuuronderzoek. Report NRLO/2000/4
- Woolf, the Rt Hon Lord, Jowell, J and Le Sueur, A.P 1999, *de Smith, Woolf and Jowell's Principles of Judicial Review*, Sweet and Maxwell, London
- Worthington, S 2003, *Equity*, Oxford University Press, Oxford
- Young, O.R 1994, *International governance*, Cornell University Press, Ithaca
- Young, O.R, Berkhout, F, Gallopin, G.C, Janssen, M.A, Ostrom, E and van de Leeuw, S 2006, The globalisation of socio-ecological systems: An agenda for scientific research, in: *Global Environmental Change*, vol. 16, pp. 304-316
- Zahariadis, N 2007, The Multiple Streams Framework; Structure, Limitations, Prospects, in: Sabatier, P.A (ed), *Theories of the policy process*, Westview Press, Davis CA
- Zittoun, P 2009, Understanding policy change as a discursive problem, in: *Journal of Comparative Policy Analysis*, vol. 11, no. 1, pp. 65-82
- Zohlnhöfer, R 2009, How Politics Matter When Policies Change: Understanding Policy Change as a Political Problem, in: *Journal of Comparative Policy Analysis*, vol. 11, no. 1, pp. 97-115

