Infrastructure for Sustainability

Infrastructure for Sustainability

MARCO A. JANSSEN AND JOHN M. ANDERIES

Contents

Copyright

Published by Center for Behavior, Institutions and the Environment Arizona State University, ECA 307, 1031 S. Palm Walk, Tempe, AZ 85281
https://cisi.info/

Dedication

In memory of Crawford Stanley Holling and Elinor Ostrom

Preface

In October 2009 the Nobel committee announced that political scientist Elinor Ostrom would receive the Nobel Prize in Economic Sciences, together with economist Oliver Williamson, "for her analysis of economic governance, especially the commons." Ostrom's work showed that a traditional approach in economics to the study of the management of shared resources (e.g., public infrastructure, common-pool resources) was incomplete. The conventional approach assumed that when people share a resource, such as groundwater, fish or a forest, everyone acts in their own self-interest leading to overharvesting of the shared resource. The only way to avoid this so-called "tragedy of the commons" is to privatize or nationalize the resource.

In a series of studies over several decades with many colleagues around the world, Ostrom showed that people are able to self-organize and successfully govern their shared resources. Her analysis provides insights into the conditions under which self-governance is possible. These findings have major implications for policy and can help explain the ineffectiveness of many policies and governance regimes.

The theoretical framework she developed over her career is applicable to the study of the governance of shared resources in many different contexts. We worked with "Lin," as she preferred to be called, from 2000 until her death in 2012. We collaborated on various projects focused on the governance of the commons, especially on questions related to robustness. Lin Ostrom had been a professor at Indiana University for her entire career starting in 1965, and beginning in 2006 she held a part-time appointment as a research professor at Arizona State University in order to collaborate in what is now called the Center for Behavior, Institutions and the Environment, which we direct.

The book is also reflecting the increasing understanding of what makes systems resilient. The world we live in is experiencing many changes, with climate change as the main challenges. The concept of resilience helps us to understand how we may manage resilience. We had the pleasure to work with "Buzz" Holling as part of the Resilience Alliance. The framework we present in this book connects the understanding of governing shared resources with the understanding of resilience of complex systems.

This book extends our previous book Sustaining the Commons, to governance of shared infrastructure. Building on the work of Ostrom, we see the governance of shared infrastructure as a natural extension of governance of shared resources. With shared infrastructure we emphasize the importance of creation, design, and maintenance of the infrastructure as well as the distribution of the outcomes of the infrastructure. With infrastructure, we not only focus on traditional hard infrastructure like bridges and roads, but also environmental, human, social and soft infrastructure, who together create the societal dynamics we observe.

While "Sustaining the Commons" book focuses on extraction from natural resources, we now extend the governance discussion to physical infrastructure, social relations, educational systems, etc. It enables us to address questions that are prominent in current sustainability debates, such as a just transition to a sustainable future. We need to know from case study analysis what worked well in the past, but we also need to be able to extend our insights to the big challenges of today.

Another reason for this book is that the COVID-19 pandemic revealed many vulnerabilities of existing coupled infrastructure systems. We were surprised about the inability of many countries to have an effective response to a disease that disrupted society, and the reluctance of many people to comply with governmental guidelines to avoid getting infected by a disease that could handicap or kill them. It is especially surprising since the outbreak of a pandemic was expected and governments could have been prepared. If it is so difficult for individuals and governments to change behavior if a disease is spreading, how do we expect a societal transformation to a sustainable future is possible?

There are no simple solutions to the big challenges we are facing, but we hope that this book provides insights on how to analyze interdisciplinary governance questions. We also provide some suggestions on how to apply insights from the book to practical solutions. As Lin would say, “there are no panaceas”. But we can improve your human infrastructure to address sustainability questions from an infrastructure perspective.

We would like to thank our students and teaching assistants who participated in classes where we used earlier versions of this manuscript for providing helpful feedback.

PART I

PART 1: SETTING THE STAGE

CHAPTER 1

The Challenges We Face

Key Concepts

In this chapter we will:

- Discuss the broad challenges humanity is facing
- Define the meaning of the term "commons" as it relates to modern social challenges
- Provide examples and a critique of the "tragedy of the commons" concept

1.1 Living in the Anthropocene

We live in interesting times. Humanity is causing such an impact on the functioning of Planet Earth that we have even started referring to recent times a new geological era, the Anthropocene. Human activities have caused such a wide and deep impact on the ecological and biochemical processes on Earth that the climate is changing, coral reefs are bleaching, the forests are on fire, the oceans turn to acid, the permafrost is melting, and groundwater levels are declining. These changes are caused by extracting huge amounts of minerals and biomass for our own energy, food, and material consumption and dumping it after use in concentrated forms in the wrong places in the Earth's biogeochemical cycles.

Humanity has experienced an unusually stable climate during the last 10,000 years which has enabled us to create complex societies. From hunter-gatherers, we transitioned to sedentary agricultural societies. We have major social and technological achievements to show for our skill at extracting resources from the biosphere through agriculture. We put humans on the Moon, controlled nuclear reactions to generate power, decoded DNA, created art, iPhones, beautiful architecture, global supply chains, and relatively peaceful life together at very high density among genetically unrelated individuals in large cities. But, the accumulation of disturbances caused by those complex societies will make the environment more unpredictable, impacting the way we can produce food, use energy, and find shelter. We can expect more frequent major storms, forest fires, heat waves, landslides, and more salinity of dry lands. Besides the environmental challenges, we are also experiencing increasing inequality in wealth within and between countries. The increased population densities in an urbanized world makes us more prone to infectious diseases, such as the COVID-19 pandemic, and increased resource scarcity to maintain our highly materialistic way of life is causing conflicts within and between countries.

How did we get here and how can we cope with these changes? In this textbook, we will discuss some of the fundamental dynamics behind these changes we observe. Building on principles from the social and life sciences as well as engineering, we present an integrated framework, the coupled infrastructure systems framework, that can be used to analyze complex problems. Because we use a common framework to look at many systems around the world, we also have developed an understanding of what leads to sustainable outcomes, and what is more likely to fail.

The coupled infrastructure systems framework originated in the study of the commons, and provides a broader perspective on applying the lessons learned from the extensive literature on governing the ecological/biological commons to other similar resources.

In the rest of this chapter, we will focus on 'the commons dilemma', one of the fundamental challenges that must be overcome to achieve the sustainable use of natural resources. The next chapter will translate this to a broader set of collective action problems, namely around any type of shared infrastructure, and in chapter 3 we continue with a discussion on the governance challenges we face.

1.2 What are the commons?

The original meaning of the term "commons" comes from the way that communities managed shared land in Medieval Europe. This shared land was not owned by any single individual but, rather, was "held in common," by a community, usually a village, thus the term "commons." Along with this shared land was a clear set of rules developed by the community about how it was to be used. Technically, the term "commons" thus refers to the land and the rules that go with it to govern its use. Over time, the term commons has taken on several meanings. Most generally it can be used to refer to a broad set of resources, natural and cultural, that are shared by many people. Examples of resources that are referred to as "commons" include forests, fisheries, or groundwater resources that are accessible to members of the community. The key term here is "shared." Forests, for example, need not be shared—there are many examples of private forests. Thus, implicit in the term "commons" as it is frequently used today is that there are no property rights established over the resource. That is, the resource is "open access." This departs somewhat from the original meaning and has, unfortunately, caused some confusion as we shall see later. Other examples of commons that the reader will encounter in everyday life include open source software, Wikipedia, public roads, and public education. Throughout this book, we will use the term "commons" to refer to a resource, or collection of resources over which private property rights have not been established.

Regardless of how they are managed, these examples show that the types of resources that can be defined as "commons" are essential for our societies. We share them, inherit them from previous generations, and create them for future generations. The commons are therefore crucial for our wealth and happiness. Those commons are an example of collective action problems, which are situations where there is a conflict between the interest of the individual and the interest of a group. Collective action, such as coordination and cooperation, is needed to achieve collective outcomes.

Why would we care to study the commons? In this chapter, we will explain that there is a big challenge associated with sustaining the commons. Because of the lack of clear rules of use and mechanisms to monitor and enforce those rules, many commons are overharvested. Examples include fishers fishing the oceans in international waters, farmers pumping up groundwater, or movie watchers using the limited bandwidth of the community internet connection, reducing data availability for other users. How can we make sure that the commons are used wisely and fairly? Who should regulate the use of the commons? Who should make the rules? In the original commons in Medieval Europe, the answer to these questions was clear: the community that held the land in common made the rules and enforced them to regulate the use of the commons. In modern commons, where the resources in question are typically much more complex, answering these questions is much more difficult.

When we, the authors of the book, were teaching a course in Beijing, we had to walk on the streets wearing masks to protect ourselves from air pollution. This experience is a powerful reminder that the air we breathe is part of a commons. As individuals, we have no control over the pollution in the air and, as a result, of the quality of the air we breathe because there are no comprehensive property rights governing access to the atmosphere. In some cities, the air quality is dangerously bad, while in others the sky is blue and there are no measurable pollutants. What underlies these differences? Is this due to regulation, population density, or the geography of the landscape? What are the costs and benefits of improving air quality and who will lose and who will gain from such changes? Who is making the decisions on activities that affect air quality? So the type of question that is of interest to people who study the commons is “what enables some groups to successfully resolve commons problems and what prevents others from doing so?”

There are many successes and failures regarding governing the commons. We will introduce a framework that can be used to help us analyze the various types of commons that are so important to our well being and illustrate how it can be used to provide a better understanding of how to better govern our shared resources. There is no silver bullet solution that will always lead to the outcomes we desire, but we can learn about mechanisms that increase the likelihood of achieving desirable outcomes.

How to effectively govern the commons has been a long debate in academia. Over the last 50 years, the traditional approaches to solving the commons problem through privatization or state regulation have been challenged. The next section will introduce the basic elements of the debate, the controversy that has arisen around it, and some alternative solutions.

1.3 The tragedy

In 1968, biologist Garrett Hardin (Figure 1.1) wrote a famous essay in the journal Science titled “The Tragedy of the Commons.” Garrett Hardin was an American ecologist who warned of the dangers that the increasing human population would impose on the environment. He argued that when people share a resource they will overharvest it because it is in their individual interest to take as much as possible.

Figure 1.1: Garrett Hardin

Hardin used the metaphor of sheepherders sharing an open-access pasture. He erroneously referred to this open-access shared resource as a "commons" (if it were really a commons the community would use a common-property governance regime to regulate access—more on this later). The title of his paper should have been "The Tragedy of Open-Access." Unfortunately, this use of the term "commons" stuck and, in fact, has had unfortunate consequences, as we will see shortly. Because there are no restrictions on the use of the pasture, each herder can benefit as an individual by adding extra sheep. Unfortunately, if all the herders add sheep, as a group they will eventually bear the costs of the additional grazing, especially when it creates a situation in which the total number of grazing animals consumes grass faster than the pasture can regenerate new grass. The effect of overgrazing is shared by all herders, but the benefit of adding extra sheep goes to the sole owner of the sheep (as long as other herders do not add too many sheep).

Based on the reasoning that people are rational selfish actors, any time the benefits of using a shared resource are private and the costs are shared, we can expect the commons will be overgrazed. Hardin formulates this as follows:

Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit – in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in a commons brings ruin to all (Hardin, 1968 p. 1245).

The observation that people cause problems for the common good when they follow their self-interest is not new. The Greek philosopher Aristotle noted more than 2000 years ago that “what is common to the greatest number has the least care bestowed upon it.” The reason that Hardin’s argument got so much attention was due to his recognition that the concept can be applied to many modern environmental problems. With the emerging interest in environmental conservation in the 1960s, he provided an explanation for why we were causing so much damage to the environment.

Hardin concluded that there were only two options to avoid the depletion of the commons. One option was to give the herders private property rights. If each herder owned a piece of the common land and the herder’s sheep caused overgrazing and erosion, the costs would be felt by the individual herder only. For this reason, the rational herder would choose to put an appropriate number of sheep to graze on the land in order to maximize her long-term earnings. The other possible option is for a government body to restrict the amount of grass that can be consumed. However, in order to enforce the restriction, the government would have to monitor the amount of grass consumed by each herder—a costly exercise. An alternative would be for the government to require that herders pay a tax per head of sheep, which the government would use to hire a guard to monitor whether the herders follow the rules.

The importance of Hardin's argument is its conclusion that people are not able to self-govern common resources. That is why he calls it a tragedy. The fact that Hardin focused on this inevitable tragedy is perhaps related to his use of the term "commons." In fact, in traditional contexts there was no "freedom in a commons"—a commons always had a set of rules associated with its use, and these rules did not necessarily include either of Hardin's two options. Unfortunately, Hardin's judgment has been widely accepted due to its consistency with predictions from traditional economic sciences and increasing numbers of examples of depletion of environmental resources. What this judgment fails to take into account are the many cases of successfully managed commons in which the shared resource is used sustainably. That is, there are many cases where a "tragedy of the commons" has been averted without privatization or state control.

The consequences of this work were significant. Hardin and others did distinguish three types of property rights: communal, private, and state. However, they equate communal property with the absence of exclusive and effective rights and thus with an inability to govern the commons. Experience does not bear this definition out: communal property, or common-property governance regimes do provide exclusive and effective rights, which are often used to govern the commons. From Hardin's perspective, which neglected this third governance regime, sustainable use of shared resources without the state or private property was only possible when there was little demand or a low population density.

Garrett Hardin provided a compelling explanation for the emerging environmental movement in the 1960s. There was an increasing awareness of the decline of natural resources due to human activities, including the perceived scarcity of raw material; deforestation; overfishing; as well as increasing levels of water and air pollution, leading to smog and acid rain as well as health problems for human populations.

A few years after the publication of Hardin's article, the first oil crisis took place which led to a rapid increase in oil prices. This shock generated the perception that oil was becoming scarce and that we were overusing our shared resources. Hardin's paper provided a simple analysis and a simple solution. Assuming people make rational decisions, the implications for policy were clear. To avoid overexploitation of resources shared in common it was critical for the state to either 1) establish, monitor, and enforce private property rights or 2) directly regulate the use of the commons either by taxing or directly restricting (e.g., licensing) its use.

Figure 1.2 shows the decline of the stock of predatory species in the world's oceans over a 40–50 year period during the second half of the 20th century (Myers & Worm, 2003). Since the 1968 essay, policies have changed, yet we haven't seen a reversal of the overall trends. The fish stocks in Figure 1.3 still have not started to recover even after the institution of many new fishing policies since the early 1970s. Moreover, we are now beginning to experience new environmental commons problems, like the loss of biodiversity and climate change, despite efforts by nations to draft international treaties to regulate these "global commons."

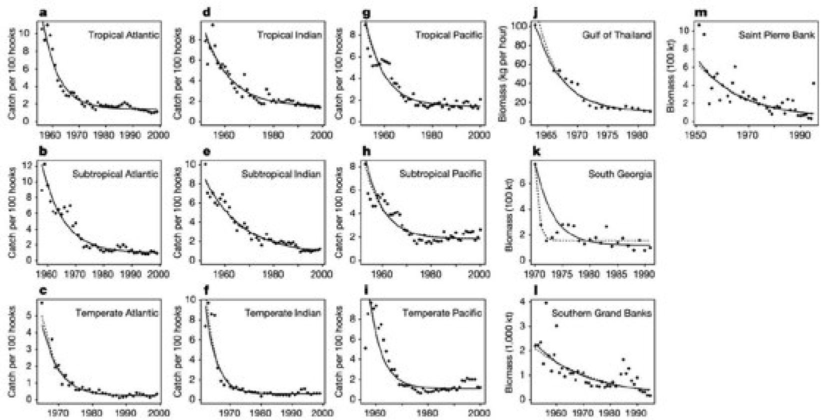

Figure 1.2: Relative biomass estimates from the beginning of industrialized fishing (Myers & Worm 2003)

As we have hinted above, we will show in this book why Hardin's analysis was limited. Although we see resource collapses around the world (tragedies of open access), we also see many success stories of long-lasting governance of shared resources (triumphs of the commons). Open access situations are not always tragedies. Many times common-property management regimes fail, as do private property and state-centric regulatory governance regimes. There are no panaceas. The goal of this book is to illustrate a set of tools that can be used to determine what conditions make overexploitation more likely and what conditions are more likely to lead to the sustainable use of shared resources.

1.4 The common pasture of Hardin

As we mentioned above, in his description of the "commons," Garrett Hardin implicitly assumed open access to the pasture. The example Hardin gave was grazing on common land in Medieval

Europe. Let's look at the actual situation of the medieval open-field system in Europe, especially in England, in more detail (Figure 1.3).

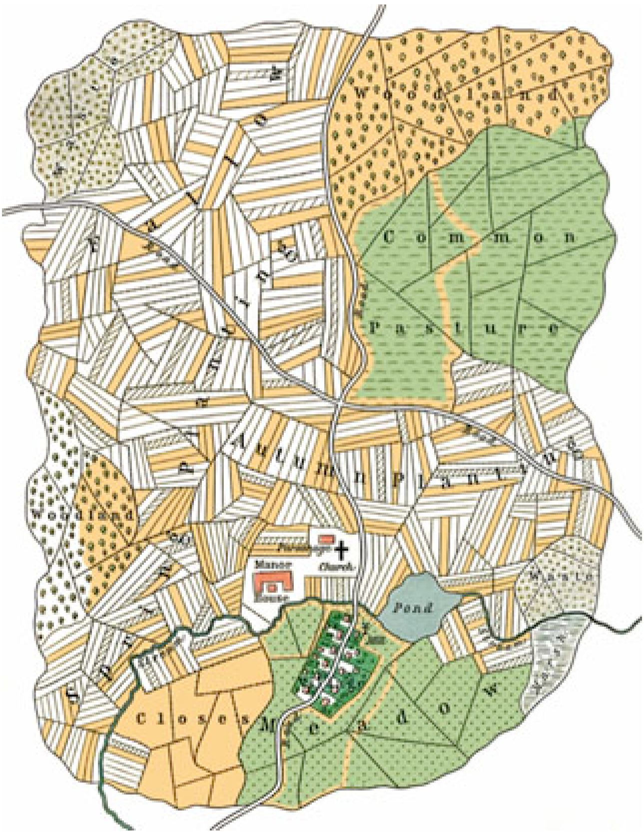

Figure 1.3: Open field system

In the open-field system, peasants had private property rights to the grain they grew on multiple small strips of land that were scattered around a central village. However, during particular seasons, peasants were obligated to throw the land open to all the landowners in a particular village so that they could all graze their sheep on the common land under the supervision of one herdsman. The decision to convert the strips of privately used land into shared land for a period during each year was made by a village council. This enabled people to take advantage of economies-of-scale in grazing (as well as providing manure for their land) and private incentives in grain-growing (which lacks important economics of scale and suffers from free-riding when communal groups try to share labor inputs. This is an example of a social dilemma, a topic we will discuss in Chapter 3.

The purpose for scattering small strips of land has been debated among scholars, as the benefits of the two scales could be achieved with or without the scattering of the agricultural land. Further, the scattering of land appears to have been an inefficient system, given that a single farmer had to divide his time between multiple, small agricultural strips rather than being able to economize on his own time and focus on one piece of land. Some scholars argue that the need to share risk due to different soil and precipitation patterns may have been a contributing factor. Others argue that by not allowing anyone farmer to gain a large amount of contiguous land, the village avoided creating a situation of asymmetric bargaining power. No farmer-owned enough land to be able to “hold out” from the commons and graze his own animals on his own land. Nor did an individual have a right to exclude others once the village decided the land should be converted from agriculture to pasture. If all of the farmers had owned sizable chunks of agricultural land in “fee simple” (a form of private ownership in England), rather than the village being responsible for land-allocation decisions, transaction costs would be very high.

If the argument that the commons were managed effectively in the open-field system has some validity, why did the open-field system disappear? And why did it take such a long time for it to disappear across most of Northern Europe? If private property alone was a very efficient solution to the production of food, once a particular location discovered this efficient solution, one would expect to see a change occur rapidly throughout Europe. The explanation might relate to transportation costs. Due to high transportation costs, local communities needed to produce both meat and grain in a small local area for their own consumption. This was only feasible if they could convert agricultural land to a common pasture when the crops had been harvested. When transportation networks improved and communities gained access to markets in grain and meat, there was no longer a need to continue with this complicated adaptation. Communities could specialize in meat or grain. Interestingly, this shift was facilitated by the development of a new “commons,” i.e., the shared resource of the public transportation system.

Thus, as we mentioned above, the medieval commons used by Hardin in his metaphor were, in reality, not open access. The commoners had crafted effective norms and rules to govern their shared pasture and to avoid overexploitation. Moreover, there are many implicit rules involved in the use of the commons. For example, a herd of livestock is the private property of the farmers, but the grass they consume does not become private until the animal swallows it. Could farmers directly harvest the grass for their livestock? A farmer who does so will likely get in trouble as there might be informal rules that grass can only be harvested via the livestock.

This simple example of a shared pasture with grazing sheep illustrates how common-property governance typically involves many rules and norms. Often, the intentions of these rules and norms and the way they function are not at all obvious from casual observation. We will see that there are always many such norms and rules involved in the use of the commons—some obvious, some very subtle.

In summary, at the time Hardin wrote his now-classic article, the work on collective action was rooted in rational choice theory. A key assumption of this theory was that actors made rational (calculated) decisions based on selfish motives (weighing individual costs and benefits). The implications for policy were clear: to avoid overharvesting of shared resources it was critical to establish private property rights or tax the use of the commons. Much work since has shown that this simply isn't the case.

1.5 The tragedy is not inevitable

Since Hardin's essay, an increasing awareness has emerged that tragedy is not the only possible outcome when people share a common resource. There are many examples of long-lasting communities that have maintained their shared resources effectively. Since the 1980s there has been a steady increase in interdisciplinary efforts to debunk the simplistic view of the tragedy of the commons. Elinor Ostrom (Figure 1.4) and others showed through comparative analysis of many case studies that communities can self-govern their shared resources.

Elinor Ostrom was a political scientist who developed a theoretical framework to study the ability of communities to overcome the tragedy of the commons. This research earned her the 2009 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel (better known as the Nobel Prize in Economics). Her Ph.D. thesis, which she finished in 1965, focused on the management of shared groundwater resources in Southern California. In her first fifteen years on the political science faculty at Indiana University she studied police forces in U.S. cities, seeking to discover which types of organizations led to the most effective policing.

Figure 1.4: Elinor Ostrom.

Because she worked on various types of projects related to the governance of shared resources, she started to see commonalities. Since the early 1980s, Ostrom developed a more theoretical understanding of the institutions, rules and norms that communities use to organize themselves. This led her to develop the Institutional Analysis and Development (IAD) framework, which is a core framework in this book.

During the mid-1980s Ostrom returned to the study of problems related to the governance of environmental commons. An increasing number of scholars at the time were realizing that reality clashed with the conventional view that the use of a shared resource would end in environmental disaster. Ostrom proved instrumental to this revolution in thinking by leading an effort to compile hundreds of case studies—successes and failures—from the lobster fisheries of Maine to the irrigation systems of Nepal.

The comparative analysis of these case studies allowed her to identify features that were more common in successful cases. In her 1990 book Governing the Commons, she identified eight design principles that characterized successful self-governance strategies, including having monitors who are accountable to the users of a resource and cheap mechanisms for conflict resolution. Those principles are discussed in Chapter X and have held up to the test of time.

Since the early 1980s, an increasing number of anthropologists, sociologists, political scientists, ecologists, and many other scholars have been documenting examples of resources shared in common that have been managed sustainably for a long time without private property rights or governmental interventions. This led to the development of a community of scholars who came together to create the International Association for the Study of the Commons of which Elinor Ostrom was the first president.

The work coming out of this community has provided an alternative framework for studying the use of shared resources, i.e., resources held in common. The material discussed in the coming chapters is largely based on this alternative framework which has been widely recognized. Besides a Nobel Prize in Economics (which was seen by Ostrom as a recognition of the whole research community in this area, not an individual accomplishment), insights derived from this research are increasingly applied to governance and policy issues. We worked with Ostrom from 2000 till her death in 2012, and started developing a broader perspective of her framework, which is the focus of this book.

Applications of Ostrom's work can be found in organizations that manage development projects in developing countries, advance agricultural practices to improve food security, and protect biodiversity. Moreover, the insights on how to sustain the commons are increasingly applied to non-traditional commons such as in the areas of knowledge, culture, education, and health. For example, the communication revolution driven by the internet has generated all kinds of new challenges related to governing the digital commons. Creations consisting mainly of information (movies, books, music) are so easy to copy, that many get distributed without any payment to the owners of the intellectual property rights. Strangers can post improper comments to websites. Emails are sent around in order to gain access to your private information.

1.6 Outline of the book

The book consists of 7 parts. The first part of the book introduces concepts like the commons, collective action, shared resource and shared infrastructure, and the related governance challenges at different levels of scale. The second part, introduces basic concepts and frameworks such as **institutions, the institutional analysis and development framework, and action arenas**. These concepts will provide the key theoretical foundation for analyzing problems related to the shared resources. We define institutions, the rules and norms that structure human interactions. This is a very broad concept but we will see that understanding the rules and the norms related to the use of the shared resources helps us understand how to sustain them. We will use the general terminology of "institutions" rather than of private property or markets, since those two examples are vague and imprecise definitions of clusters of possible institutional arrangements. The **institutional analysis and development (IAD) framework** that we will discuss in this book provides a more general and accurate way of studying institutions and their performance. When we focus on action arenas, a key component of the IAD framework it enables us to dissect what are the **incentives**, the possible **actions**, and the **positions** of people who are using the shared resources.

Part 3 will introduce concepts from system science and apply them to collective action and problems of shared resources. We will discuss **feedback** loops (positive and negative), resilience and **tipping points**.

Part 4 introduces an extension of the IAD framework by introducing resilience and robustness concepts with infrastructure in **coupled infrastructure systems**. Discuss different type of infrastructures. Discuss infrastructure related to water management to illustrate the framework in more detail.

Part 5 will focus on current challenges of **different types of infrastructure,** from maintaining roads and bridges, to the provision of schooling and health care

Part 6 explores the need for a **societal transition** to a new configuration of our society to reduce the pressure on the environment, and to cope with the consequences of the irreversible changes we already have made. What transitions are possible, and what is needed to make those changes.

Part 7 list some of the **practical lessons** from this book. We do not have a solution to the problems humanity is facing, but building on the transdisciplinary knowledge discussed in this book we provide some guidelines on how to bring those insights into practice.

1.7 Critical reflections

Commons are natural and cultural resources that are shared by many people. People can affect the commons by harvesting from them and making contributions to their construction and/or preservation. The core question this book attempts to address is how we can sustain the commons. Garret Hardin introduced the notion of the tragedy of the commons which can occur if people share a resource. The opportunistic behavior of individuals can lead to overharvesting of the shared resource. The only way to avoid the tragedy, according to Hardin, is to establish private property rights or tax the use of the commons. Elinor Ostrom and her colleagues show from case study analysis that overharvesting is not inevitable and that successful self-governance of the commons is possible.

1.8 Make yourself think

1. Come up with commons you experience yourself.
2. Are these commons functioning well?
3. Did your grandparents use different commons than you do?
4. Now that you know about the commons, can you relate the idea of the commons to the budget discussions in Washington D.C.?

1.9 References

Hardin, G. (1968) Tragedy of the commons. *Science, 162*, 1243–1248.

Myers, R. A. & Worm, B. (2003) Rapid worldwide depletion of predatory fish communities.*Nature, 423*, 280–283.

Ostrom, E. (1990) *Governing the commons: the evolution of institutions for collective action.* Cambridge University Press.

Ostrom, E. (2005) *Understanding institutional diversity.* Princeton, NJ: Princeton University Press.

CHAPTER 2

Infrastructure Everywhere

Key Concepts

In this chapter we will:

- Define infrastructure
- Learn that shared infrastructure depends on collective action
- Become familiar with the different types of infrastructure

2.1 Introduction

In this book, we focus on infrastructure. With infrastructure, we refer to structures that enable systems to produce certain outcomes. “Infra” refers to “below” and literally infrastructure means the “underlying structure”. The colloquial use of infrastructure refers to roads, bridges, dams, water and sewage systems, but in this book we have a broader interpretation of structures underlying and supporting our society.

Some infrastructure is private, e.g. a factory or your house, others are shared, e.g. a sewage system. In this book, we focus mainly on shared infrastructures. Common examples include roads, bridges, and electricity distribution systems, as well as internet communication protocols and computer software. We will discuss different types of infrastructure throughout the book in more detail. A basic aspect of infrastructure is that it requires investment to create and maintain. In the case of **public infrastructure** (roads, dams, electrical grid, electromagnetic spectrum) that is shared, society must invest **collectively** to create and maintain the infrastructure. Who pays and who can use the infrastructure are critical collective choice questions. Are new roads paid for from a tax on gasoline or by all taxpayers? Is the road open to everyone, or only for those who can pay the toll?

Humans are not the only species who have shared infrastructure. In fact, shared infrastructure is an essential feature of many social organisms who live together in large groups and often cooperate on certain tasks. For bees, it is the hive. For termites, it is the mound. In these cases, shared infrastructure provides controlled environments where resources can be concentrated to increase fitness. There are many types of infrastructure among the diversity of ant species. Some create highways using pheromone trails, grow fungus, remove dead individuals and waste as a sanitary system, and have armies to protect the nest. In human societies, we share infrastructure such as roads, water management, sewage, and telecommunications systems which help concentrate resources, including food, water, energy, information, and people, in space and time. Although humans have more complex infrastructure systems than social insects, the basic concept remains the same.

We focus on shared infrastructure in this book since it is a broader concept than shared resources used in the study of governing the commons. In the next section we discuss the collective action problem of creating and maintaining shared infrastructure.

2.2 Shared Infrastructure as a collective action problem

In the study of the commons, the focus was typically on extraction from a pool of shared resources of some type, as if those resources are coming from natural pristine resources. Productive land shared in common depends on long term investments to maintain the productivity of the land. Sheep grazing on Hardin's pastures were moved around by herders to spread the grazing pressure. The manure from sheep fertilized the land, and grazing killed small saplings from trees and shrubs, stimulating the regeneration of grass. In fact, by careful herding of the sheep, a productive pasture can be created. Hence, the pasture Garrett Hardin referred to in his essay on the commons was a shared infrastructure, not an open access pristine resource.

In creating shared infrastructure, one needs to invest time in the creation and maintenance of the infrastructure. Herein lies the collective action problem, since one could freeride on the efforts of others. Not all infrastructure is shared. Infrastructure can be private and the responsibility of creation and maintenance lies with an individual. For example, houses can be owned by individuals, and those individual home owners are responsible for the maintenance of the house. There might be regulations related to the quality of upkeep you are expected to have about your house (e.g. by a homeowner association (HOA)), but it is the responsibility of the homeowner to do or pay for the maintenance. A household could go off the grid by installing solar panels and a water tank. This is not uncommon in rural areas where the cost of providing shared infrastructure could become too high for the community to bear.

Another infrastructure challenge is the distribution of the affordances created by infrastructure. Who can access the road, every car, or does one need to pay a fee? Who can participate in a conference, only members of an organization, or is it open for anyone? Defining who shares in the outcomes of shared infrastructure impacts the incentives for people to contribute to the construction and maintenance of it.

2.3 Different types of infrastructure

Infrastructure is a general concept and can be described in physical terms or in economic terms. The most natural definition is the economic one: infrastructure is a collection of materials (e.g., machines) and information (e.g., knowledge about how to use machines) that can produce a stream of materials (food, cars, houses) and information (music, movies) that society values. The second key feature of infrastructure is that it requires investment (a so-called opportunity cost) to produce and maintain (machines must be built, and maintained; without practice, skills decline) and infrastructure typically has little value in its own right (farm machinery isn't exciting in its own right—its main value comes through its capacity to produce food). There are different forms of infrastructure that come together to produce output (you can't operate farm machinery without knowledge) and we discuss these different forms of infrastructure in the following sections. They can all experience collective action challenges in order to be created and/or maintained.

2.3.1 Hard infrastructure

With "hard infrastructure" we are mainly referring to human-made infrastructure such as roads, irrigation systems, and nuclear power stations. Hard infrastructure enables the production and distribution of clear freshwater, waste, energy, products, people, and information. We have infrastructure to move people in cars, trains, on water, in the air, via rail and road, and by foot and on bicycles. To facilitate the movement of people, we need energy which can be produced by various types of power generation plants or processing of fossil fuels. This energy needs to be distributed through the electrical grid, or in the case of gasoline, through a combination of trucks and roads, in order to be useful.

As you can see, hard infrastructure provides the key underlying structure for society and, as such, is too large to be produced by individuals (e.g., few individuals are wealthy enough to build the Golden Gate bridge on their own—it cost around $340 million in today's dollars). It must be collectively produced and, as a result, various types of collective action problems must be solved in order to produce functional infrastructure. Who will create the infrastructure, and where will it be located? After it is created, who is responsible for maintenance? The difficulty these problems pose is evidenced by the fact that infrastructure in the U.S. is failing by some accounts. Examples include falling bridges, blackouts and flooded neighborhoods. According to the website the U.S. needs to invest 3.6 trillion dollars (15.7% of 2021 GDP) over a five-year period to maintain the function of its infrastructure, which is more than $2,000 a year for each person living in the U.S. over that period. We discuss in the next section some of the collective action problems related to provisioning shared infrastructure. First, we want to discuss some other types of infrastructure.

2.3.2 Soft infrastructure

With soft infrastructure we refer to the human-made "instructions" (think software for your computer) for using other types of infrastructure. Instructions require knowledge of how systems work and thus involve significant investment. They are also absolutely essential to generate valuable outputs. The computer (hard infrastructure) on which this book was written is useless without software (soft infrastructure). In this sense, soft infrastructure can be thought of in general as the instructions by which society is run (across all levels of organization from the individual, to neighborhoods, to counties, to nations, and the United Nations). One type of soft infrastructure essential to the topic of this book is that of “institutions.” Recall that institutions are rules (instructions, mostly in the form of if-then statements) that structure repeated interactions between people. To be effective, these institutional arrangements must be combined with other types of infrastructure (i.e., all types of infrastructure necessary for organizations to function such as buildings, communication, transportation) which create, implement and monitor the rules. Examples include the rules by which local government functions, the protocols by which crime labs and emergency services are run, the constitutional law upon which the supreme court bases its decisions, and the tax law by which the tax collector functions. Soft infrastructure enables societies to solve collective action problems and coordinate their activities.

2.3.3 Natural infrastructure

This is the hard infrastructure that is not man-made but still is critical for society. Wetlands absorb and filtrate water, trees capture water and reduce erosion, and bees pollinate flowers. Some people may use the term “ecosystem services” to refer to natural infrastructure but those services only exist within an anthropomorphic context. In our view, ecosystems are forms of infrastructure and humans can limit or enhance the capacity and performance of those infrastructures through the use of other types of infrastructures. That is, “services” only flow in coupled infrastructure systems! Without knowledge of how to hunt or which plants are useful (human infrastructure—see below), etc. ecosystems do not produce services.

2.3.4 Human infrastructure

To throw a spear, solve a differential equation or drive a car, you need to train your brain and muscles. The first time you sit behind a steering wheel might be scary and each action is done deliberately.

But when you practice enough, it becomes a routine. Human infrastructure relates to the build up of human capacity to do a certain activity. This could be the knowledge of how ecology works for the hunter-gatherer, knowledge of seed varieties and soil characteristics for the agriculturalists, knowledge of 2-D projection for the painter, knowledge of stone for the sculptor, knowledge of machinery for the industrialist, or knowledge of kinesiology for the athlete. But it can also relate to the capacity of your muscles, and having muscle memory, to play an instrument, lift 100 kg, run a marathon, or drive a car.

Human capacity is itself infrastructure because it requires investment and can produce valued outputs (when combined with other types of infrastructure). Right now, you are investing in developing your human infrastructure. The physical manifestation of human infrastructure is the neural network in each of our brains and the muscle fibers in our body. These neural networks and muscle fibers require great effort (investment) to train to do specific tasks. Your age, your health, your diet, your gut flora, all impact your capacity to do specific tasks. One might argue that we are also part of natural infrastructure. This is correct, but for analytical purposes it is useful to define this as a special class of natural infrastructure called human infrastructure.

2.3.5 Social infrastructure

Social infrastructure refers to the relationships we have with others. These relationships (e.g., trust) are essential for our economy to function. Imagine the number of times you needed help from a friend or relative to get something done. What would you have done without friendship? Hire someone? Think of the trouble of hiring someone to do something you would ask a friend to do. The "trouble" of relying on markets (to hire someone, you would need a labor market for "miscellaneous friend tasks") in economic jargon is "transaction cost." Social infrastructure reduces transaction costs. Metaphorically, social infrastructure is the grease that reduces the friction of human interaction and allows the machinery of society to function. Further, building trust (either via friendships or professional relationships) is extremely time-intensive (i.e., requires significant investment). Thus, because social infrastructure produces benefits and requires investment, it is infrastructure.

2.4 Beyond the commons

By focusing on shared infrastructure, we aim to extend the study of the commons to a broader set of problems. We see the traditional study of the commons as the study of maintaining the use of shared natural infrastructure and evaluate which soft infrastructure configurations lead to desirable outcomes. Increasingly, the commons perspective is being applied to new topic areas like knowledge commons, health care, outer space and urban environments, where we see a prominent role of hard infrastructure.

Scholars have started to talk about social-ecological systems, and social-ecological-technical systems. To us those systems are all variations of infrastructure and it makes sense to have a more unifying framework instead of defining all kinds of applications (social, ecological, technical). By deriving understanding of the fundamental governance problems with infrastructure, we can apply those insights to many different applications, including those which we may not yet have anticipated.

2.5 Critical reflections

Infrastructure is a collective structure that enables systems to produce certain outcomes. The creation and maintenance of shared infrastructure is a collective action problem. There are different types of infrastructure such as hard infrastructure, soft infrastructure, natural infrastructure, human infrastructure, and social infrastructure.

2.6 Make yourself think

1. What shared infrastructures have you used today?

2. In what ways do you contribute to the maintenance of infrastructure? Give three examples.

3. If you find a mistake in the textbook, please let us know. This is an example of contributing to the maintenance of the shared soft infrastructure: the textbook!

CHAPTER 3

The Governance Problem

Key Concepts

In this chapter we will:

- Learn how to define governance
- Understand the three basic types of governance
- Become familiar with the four types of goods
- explore different levels of governance

3.1 Types of Governance?

Governance refers to the norms, institutions, and processes that determine how power and responsibilities over infrastructure systems are exercised, how decisions are made, and how different people participate in these processes. Three iconic types of governance can be broadly defined. They differ in how power is distributed (e.g. who makes the decisions):

- The Government. The decisions are made by a small group of individuals who are elected representatives or other actors who have obtained power to make decisions for the jurisdiction of the government.

- The Market. Individuals make decisions based on price signals. Property rights need to be clearly defined in order to establish market mechanisms.
- The Community. Members of a community come together to mediate conflicts and coordinate the creation and maintenance of shared infrastructure.

There is limited understanding regarding which governance system is preferred in which situation. In this textbook a lot of our focus will be on providing tools to identify what governance options are preferred in which circumstances. One of the challenges is that there can be different perspectives on how to evaluate a governance system. What are the relevant outcomes on which governance should be judged? This could be the provision and state of the shared infrastructure, how costs and benefits from the shared infrastructure are distributed, and how sustainable the governance system might be.

3.2 A typology of goods

In many everyday situations, there exists a dilemma, between what is best for the individual and what is best for the group. For example, it would be beneficial for an individual to be able to use a siren so that all the other cars on the road will pull over, allowing the one with the siren to get from point A to point B as fast as possible. It would also be beneficial to an individual not to pay taxes (at least in the short term). Nevertheless, society would not function if everybody used a siren when they drove or did not pay their taxes. In general, society will not function if individuals do only what is best for themselves alone. This is called a **social dilemma**. It has often been argued that what is best for the individual is best for society. The key to this claim is that it is true for perfectly functioning markets which only exist in theory– they do not, and cannot exist in the real world.

The commons dilemma discussed in Chapter 1 is an example of a social dilemma. We now discuss the broader scope of social dilemmas. We will see that social dilemmas lead to the prediction that people will not contribute to the common good or will overharvest from shared resources. These predictions are based on a very narrow notion of human behavior, namely that everyone behaves as selfish rational beings. Empirical studies provide a more nuanced perspective and find that many people are conditional cooperators. Nevertheless, it helps to think about social dilemmas using the naïve model of decision-making. It will point to some potential problems related to social dilemmas in which we have to make choices.

The very existence of Wikipedia demonstrates that we do not all behave as selfish rational beings. Many people now use Wikipedia because it is a very powerful resource with high-quality information on many topics. There are far fewer people who write articles for Wikipedia than there are people reading Wikipedia articles. Those people who only consume Wikipedia "free ride" on the contributions of others. Those who write articles spend time (experience a cost) to do so. Luckily there are enough people willing to make a contribution voluntarily in order to have a very useful product. If people really are fundamentally selfish, why do you think these individuals contribute to Wikipedia?

Many social dilemmas are related to the production or consumption of goods. In this section, we describe different categories of goods and how rules relate to them. We can use two attributes, exclusion and subtractability, to distinguish four basic categories of goods and services (Table 3.1).

Exclusion relates to the difficulty of restricting the use of those who benefit from the resource or service. Subtractability refers to the extent to which one individual's use subtracts from the availability of a good or service for consumption by others.

Different levels of exclusion and subtractability define different categories of goods. So what do these dimensions actually mean for goods and services? Let's discuss the different dimensions in more detail.

- **Subtractable**: The use of a good or service by one participant in an action arena reduces the availability of the good or service to another participant.Examples: A fossil fuel, like oil for example, is a non-renewable resource that is used for many energy sources. The gallons of gasoline that you put in your car cannot be used by somebody else after you have burned the gasoline during your trip. A more direct example is a cake you have made for your friends. Every piece of cake eaten by one person is not available for somebody else. This is related to the old adage "you can't have your cake and eat it too."
- **Non-subtractable**: The consumption or use of the good or service by one participant in the action arena does NOT reduce availability or utility of the good or service to another participant.Examples: Reading an article on Wikipedia does not reduce the availability of the article for somebody else. Many information goods like movies, photos, books, and scientific knowledge have this property.
- **Excludable**: Any excludable good or service is one that a participant can be prevented from accessing if they do not pay for it, or have passed another form of entry/access barrier.Examples: Going to the movies requires you to buy a ticket to get access. The movie theater is constructed in such a way that the formal requirement of having a ticket can be translated into the real-world outcome of being prevented from entering, i.e. being excluded from the movie theater. Some websites require

you to sign up, like Facebook. You can be excluded if you are too young or have misbehaved in the past. As with the movie theater, excludability on a website requires that the infrastructure be constructed in such a way that a formal measure like requiring a password (a ticket, or more generally a 'token') can be translated into the desired practical outcome.

- **Non-excludable**: Any good or service that a participant cannot be prevented from accessing or it is extremely expensive to exclude.Examples: Public roads are available to all cars, even though not all participants pay taxes for the maintenance. It is very costly to deny fishing boats access to the oceans, especially outside the control zones of countries. This makes an ocean fishery very difficult to regulate. This relates to the excludability examples above. We may define a formal requirement of a permit (a ticket, password, or token) to fish, but this has little practical effect without the infrastructure required to translate it into action. Imagine trying to wall in a fishery like a theater, or patrol a large open air theater with no fence like a fishery.

When we combine these two dimensions we can define four different theoretical categories of goods (Table 3.1): private goods, club goods, public goods and common-pool resources. We will discuss below examples of each of these categories.

	Excludable	**Non-Excludable**
Subtractable	Private goods	Common-pool resources
Non-Subtractable	Club goods	Public goods

Table 3.1: Four basic types of goods.

Let's start with private goods. One can restrict the use of the good easily and when that good is in use, someone else cannot use it. For example, a mobile phone or a car is a private good (Figure 3.1). You can restrict the use of these goods by having password protection or a key. If somebody else takes the private good you can go to the police to report a theft. If you are driving your car, no one else can drive it. If you are making a call on your mobile phone, others cannot use it during your call. Hence private goods are the typical products we own as an individual. Can you think of examples of private services?

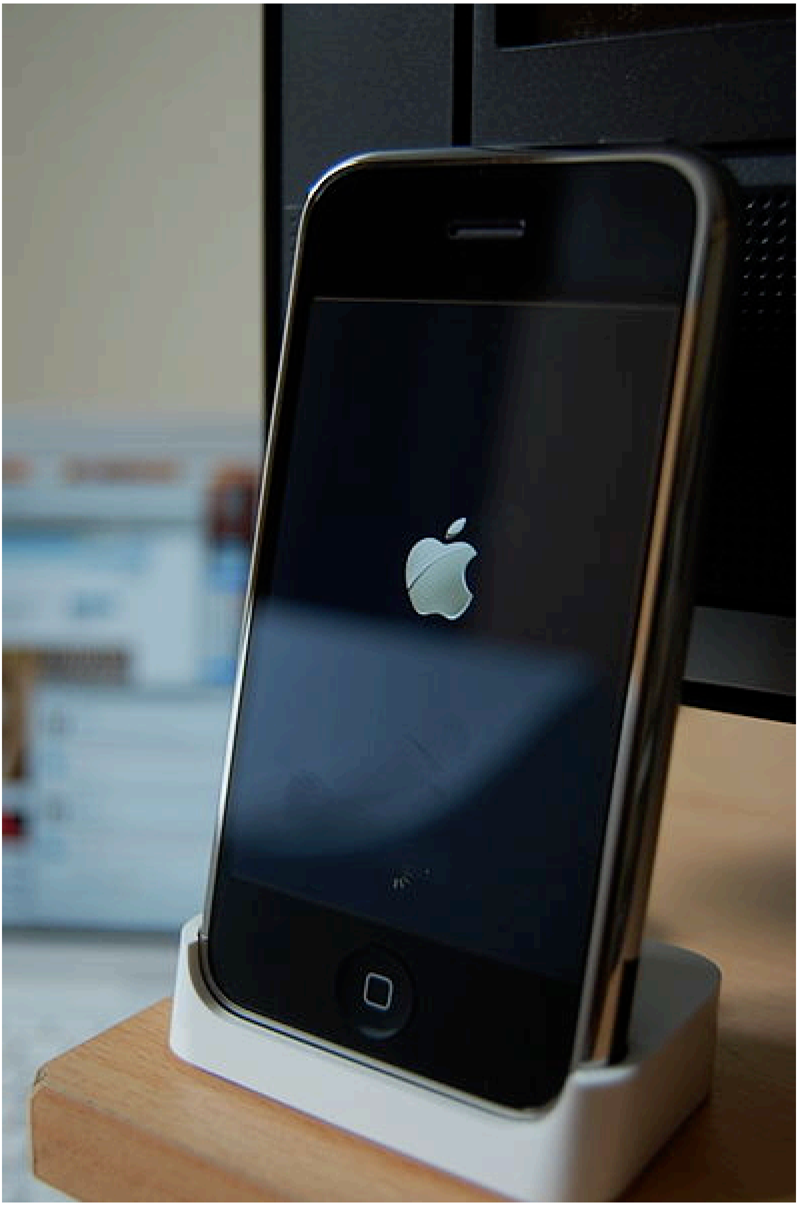

Figure 3.1: Smartphone.

The second type is club goods. Compared to private goods, the use of the club good by one person does not affect the use by others. Examples of these so-called club goods are streaming services like Netflix, cinemas, and toll roads (Figure 3.2). Access is restricted to "members of the club," but those members can consume the same product. For toll roads, it means that you pay a fee to use the road, moviegoers pay a price for a ticket to watch a movie during a showing, and Netflix users need to pay a subscription fee (and not share their account information with other households).

The provision of goods is more difficult for those types of goods for which it is challenging to exclude consumers. Public goods—open-source software, Wikipedia, clean air—can be used by everybody, and use by some does not reduce the ability of others to use it. The challenge associated with providing public goods is having a sufficient number of people to invest in their provision. There is a temptation to "free ride" on the contributions of others. Due to the potential for free riding, there might be an underinvestment in public goods. Consider for a moment whether the roads in your city in good condition and are there enough of them? Are there enough parks and open spaces?

A public road is open to all, but not everybody has necessarily provided a significant contribution to its construction, which is financed by various types of local and federal taxes. On a toll road, each user must pay to get access to the road, but access to a public road is not restricted to those who pay for it's use upfront. The same comparison is true in pay-per-view versus public television. The same physical product can be offered as different types of goods, by changing the rules governing who has access to the good. A public library is a public good for those who want to read a book or newspaper in the library. If you want to take a book home for a limited amount of time, you need to become a member of the library. But if you check out a book from the library no other patron can use that book. So how then is the library a public good? Although the book is available, some coordination is still needed to make the physical objects available for a limited time for those who ask for it.

Figure 3.2: Access to toll roads is restricted to those who are willing to pay a fee.

The final example is a common pool resource, such as a lake, an ocean, an irrigation system, a fishing ground, a forest, the Internet, or the stratosphere. These are natural or constructed resources from which it is difficult to exclude or limit users once the resource is provided by nature or produced by humans. One person's consumption of resource units, such as water, fish, or trees, removes those units from what is available to others. Thus, the trees or fish harvested by one user are no longer available for others. The Internet has a limited capacity to move all the information around. Bandwidth that is used by some to watch movies may cause delays in the sending of email by others. The many satellites required for communication (along with space debris from the past) is causing problems for new operations in space.

When the resource units produced by a common-pool resource have a high value and institutional rules do not restrict the way resource units are appropriated (an open-access situation), individuals face strong incentives to appropriate more and more resource units eventually leading to congestion, overuse, and even the destruction of the resource itself. Because of the difficulty of excluding beneficiaries, the free-rider problem is a potential threat to efforts to reduce appropriation and improve the long-term outcomes achieved from the use of a common-pool resource. If some individuals reduce their appropriation levels, the benefits they generate are shared with others whether the others also cut back on their appropriation or not. Some individuals may free ride on the costly actions of others unless ways are found to reduce free-riding as an attractive strategy. When free riding is a major problem, those who would be willing to reduce their own appropriations for the benefit of all, provided others would reduce as well, become unwilling to make such a sacrifice for the benefit of a large number of free riders.

Space on Earth is a resource too (Figure 3.3). You have, no doubt, experienced the heavy appropriation of space on the road during rush hour. If enough people drive to work earlier or later than rush hour, this would free up space for others motorists. But who wants to arrive an hour early to work? Or take the bus—an option that might not be convenient for everybody as many people prefer the privacy and control of driving their own car. Hence, the next time you experience a rush hour traffic jam, think about the options everybody has. Due to the free-riding behavior, all of the drivers experience lower performance of the road.

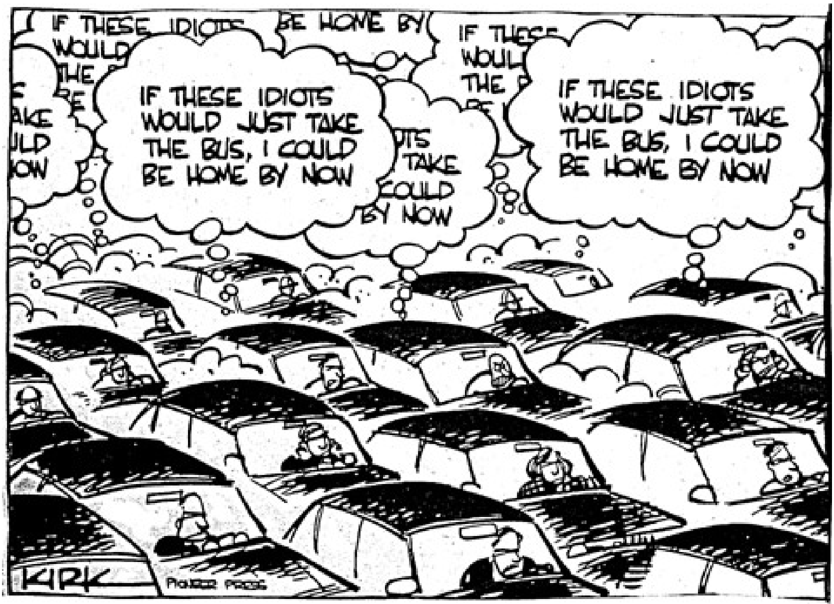

Figure 3.3: Rush hour in traffic.

Let us return to the question of how (or whether) the library is a public good. We have also said that roads are a public good. But we have also mentioned the fact that roads and libraries can become congested. Doesn't that violate the non-subtractability of the resource? It turns out that there are very few examples of **pure** public goods or any of the other goods for that matter. Typically public goods provide multiple streams of services with different characteristics. Further, they must typically be combined with other types of goods to produce a final service. For example, a road isn't much good without a private car, public bus, private scooter, etc. Thus, it is better to think of a road or a library as **shared infrastructure** that makes available common-pool resources. In the case of the road, the common-pool resource is transportation space (or capacity). In the case of the library, the common-pool resource is book contact time. The reasonable use of these common-pool resources typically involves solving coordination dilemmas in space and time. Sometimes, three or more types of infrastructure (goods) are combined. Take, for example, the iPhone.

The iPhone itself is a private good. It is useless, however, without a network. The contract you have with the telephone company enables you to use their infrastructure (a club good) to make calls. Many apps can be downloaded for free and used on an iPhone. The ecology of free apps really are pure public goods since they are available without restriction, at least if you have the proper technology to run it. Finally, if many people use their iPhone to watch movies, the wireless bandwidth gets crowded. This is exactly like cars crowding a road: information bits crowd fiber optic and copper cables. The so-called "bandwidth" is a common-pool resource that is made available by telecommunication infrastructure and if too many people start downloading movies the available bandwidth is completely consumed and the wireless signal is not available for anyone else.

As we conclude this section, we encourage the reader to reflect on how different types of goods (which we can also refer to as infrastructure) must be combined to produce any final service. Yosemite is a "public good." But unless it is combined with roads, cars, or planes, what services can it provide? For some, it does provide a truly pure public good simply in the idea that such a beautiful place exists. For most, Yosemite is, in fact, a piece of shared infrastructure that provides a common-pool resource consisting of "nature viewing time" exactly as the library is a piece of shared infrastructure that provides a common-pool resource consisting of "book viewing time." Understanding the subtle nature of how almost all final services are provided by a complex combination of the different types of goods we have discussed in this chapter is a critical prerequisite for designing institutions to wisely govern their use.

3.3 Critical reflections

During the past several decades, delegates from all over the world have come together for conferences on climate change to reach agreements on emissions reduction (Figure 3.4). Every few years there is a critical conference on climate change, and every time, thus far, the agreements made are limited and do not result in actual changes that bring about emission reductions.

Figure 3.4: International Negotiations.

Climate change is a very challenging problem since we cannot exclude anyone from using the atmosphere to dump their carbon waste (we all do it every time we take a car, bus, or plane ride). Nor can any individual or group really shield themselves from the consequences of climate change. Thus, solutions will require that a large number of people change their behavior at a cost to themselves, for the benefit of everyone—a classic social dilemma. Moreover, the benefits will only be experienced decades later by future generations. If we cannot solve the climate change problem through negotiations at the international level, at which level of organization might we attempt to address the problem?

Social dilemmas and governance challenges can happen at different levels. As we see later in this book, we have a good understanding how communities can be successful in governing their shared resources. We don't have that same level of understanding with larger scales, where there are more actors, and many of those actors have overlapping responsibilities and authorities.

For example, the Colorado River passes through five states in the U.S. and two states in Mexico. Actions of the upstream states affect the downstream states. At the time of the writing of this book, the southwestern region of the U.S. had been experiencing a drought for many years and this led to tensions over water availability for the different states. The U.S. Supreme Court has ruled on the allocation of Colorado River water, but that allocation was based on a period of higher-than-average water supply and much lower population densities. What will happen when Lake Mead dries up and less water is allocated to Arizona? In 2022 and 2023 the water allocation to Arizona has been reduced, which have to be absorbed by the agricultural sector (in line with the rules). At the time of writing this book there are tough discussions going on how to reduce use in the agricultural sector. With the consequences of the decades-long drought in the US south-west becoming visible, the government of Arizona is investing in other water sources? Will they also aim to reduce water demand? Will upstream states be requested or required to reduce their water use? Who will enforce this?

The allocation of Colorado River water is a traditional case of the problem of scale. The governance of rivers that cross state and/or national borders can be problematic as a result of the fact that there are many actors from different jurisdictions with different regulations that don't align well trying to work together to solve many problems. For examples, check out the website. It is not uncommon that organizations are created to focus specifically on the governance problems of a particular river, as happened with the Rhine River in Europe. In the 1980s the downstream country, the Netherlands, experienced major problems with the pollution caused by upstream nations like Germany and France. In order to solve the problem, regulations had to be implemented in different countries. The main costs of the regulations fell to the upstream countries and the main benefits were enjoyed by the downstream countries. Such a policy involving a complete mismatch between costs and benefits of the individual stakeholders could not be implemented unless there is coordination, monitoring, and enforcement at the river basin level.

3.4 Polycentric Governance

As we have seen in previous examples, the governance of shared resources in larger systems can become complicated. A bottom-up approach might be problematic due to a lack of coordination among different actor groups and an inability to constrain actions of particular actors that negatively affect the welfare of others. A top-down approach might solve the coordination and enforcement problem, but may involve a big separation between the resource users and public infrastructure providers, making the transmission of information across levels of organization difficult. As such, top-down approaches may be less effective in monitoring and enforcement and may lack knowledge in creating institutional arrangements that fit the local conditions. A third approach is **polycentric governance**.

Polycentricity refers to a social system composed of many decision centers having power to make decisions. Each decision center has a limited and autonomous set of prerogatives (rights and privileges) over certain sets of potential actions, and operate under an overarching set of rules. The concept was developed in the 1950s and 1960s by scholars focusing on governance of metropolitan areas (i.e., to provide policing, road maintenance, etc.). For a metropolitan area it might be effective to have just one crime lab rather than one in each neighborhood. But for a typical police officer, it is important to have a police station in each neighborhood instead of one centralized police station. For the performance of the police it is important that officers have access to local knowledge and connections with the neighborhoods both of which cannot be achieved if they are housed in one central police station. For different functions there might be infrastructure at different scales, and each function might have its own rules and regulations on how the infrastructure is supported. For example, schools in the U.S. are supported by tax revenue at the school district level, and thus richer neighborhoods often have schools that have better infrastructure. Further, if infrastructure is more desirable in other neighborhoods, we may observe families moving between neighborhoods setting up a vicious cycle. People who can afford it vote with their feet. Those who can't, are stuck. To avoid major inequalities between neighborhoods, tax revenue at the metropolitan level might be distributed to help schools in poor neighborhoods.

The United States of America, for example, is organized as a polycentric system. The 50 states have their own rules and a certain level of autonomy to make specific types of decisions. This can be observed by differences in regulations, from death penalties to recreational marijuana use. For many types of problems, states have authority to implement laws and regulations that are consistent with the attributes of the local population. Certain problems, such as national security, national currency, and trade agreements, cannot be addressed at the state level and this will be addressed at the federal level.

Which level is the proper level of authority is not without controversies. For example, due to the decision of the federal supreme court, same-sex marriage is now allowed in each state. Before the supreme court's decision same-sex marriage was allowed in a select number of states. Some people argue that the meaning of marriage should be regulated at the state level (i.e., is a matter of states rights), and do not agree that a federal mandate should be imposed on a state. On the other hand, before the federal ruling a married couple who moved to another state may have experienced different regulations, which could have had major consequences for that family.

3.5 Critical reflections

Social dilemmas are situations in which two or more participants can benefit collectively from cooperation, but an individual who is **selfish** and **rational** can also benefit from **free-riding** on the cooperation of others. Four types of goods can be distinguished based on the extent to which it is possible to exclude others and the subtractability of the resource: private goods, toll (or club) goods, **public goods**, and **common-pool resources**. In this book, we will mainly look at problems of public goods (underinvestment) and common-pool resources (overuse).

In this chapter, we have also considered the challenges of governing systems at larger scales. Who has the authority and the knowledge to create effective policies? Governance structures at different levels influence each other. To address national level problems, does not mean we should only focus on national level policies. In fact, bottom up initiates could create a platform to solve problems at a higher level.

3.6 Make yourself think

1. What are examples of public goods and common-pool resources you have experienced lately?

3.7 References

Aligica, P. D. & Tarko, V. (2012). Polycentricity: From Polanyi to Ostrom, and beyond *Governance, 25*(2), 237–262.

PART II

PART 2: INSTITUTIONAL ANALYSIS

CHAPTER 4

Defining Institutions

Key Concepts

In this chapter we will:

- Learn how to define institutions and recognize them in everyday situations
- Learn to analyze institutions using action situations
- Become familiar with an overview of the institutional analysis and development framework, which will be used throughout this book
- Understand that incentives that impact decision-making can be studied using rigorous scientific methods
- Recognize the wide diversity of institutions in use around the world

4.1 Overview

There are many ways to induce people to contribute to provision and maintenance of shared infrastructure. Nudges and incentives such as rewards and punishment, shame and prestige might work in some localities but not in others. If village members in the hills of Nepal do not contribute to the maintenance of the shared irrigation system, the family cow can be confiscated and put on display in the center of the village. Since everyone recognizes the cow in these small communities, it is known to everyone that you are cheating the community. Other village members could milk the cow, until the offender paid a penalty. The cow jail works in rural Nepal but is unlikely to be effective in urban areas.

We will see that there are many different possible mechanisms that groups use to solve problems related to sustaining the commons. These different mechanisms, however, all rely on the same principles. Therefore, in this book, we will try to understand how people solve these problems by studying the basic principles contained in the institutions they use.

Broadly defined, **institutions** are the prescriptions that humans use to organize all forms of repetitive and structured interactions. This includes prescriptions used in households, schools, hospitals, companies, courts of law, etc. These prescriptions can function at different scales, from households to international treaties. These prescriptions can be one of two broad types: rules or norms. Because rules and norms are essentially human constructs, agreed-upon or recognized by a group of people, they are not immutable. That is, individuals can make choices whether or not to follow the rules or norms and these rules and norms can be changed. Importantly, their choices and actions have consequences for themselves and for others.

In the following chapters we will see that rules and norms are everywhere and define —sometimes literally, sometimes indirectly—how we live our lives. For example, rules and norms can affect who we marry, which schools we go to, which countries we enter, where we may sit on a bus, where we may park, who leads a discussion in a group, etc. Where do these rules and norms come from, and why do they differ in different countries and contexts? In this book, we are especially interested in answering this question for different types of shared infrastructure.

We will see that all of us can play a big role in defining rules and norms if we take the initiative to do so. Crafting rules and norms is not something that is undertaken exclusively by those in business suits in Washington, D.C. We ourselves create rules and norms too. For example, when you undertake a group project during a course, you will have to rely on some rules and norms. Some rules might come from the syllabus while others are created by you and the members of your group during your meetings.

In more abstract terms, the rules (or the absence of rules) in a particular situation affect who gets what benefits, who bears what costs, who is allowed to participate, and who gets what information. Further, the rules affecting one situation are themselves crafted by individuals interacting at higher levels. For example, the rules we use when playing basketball at lunchtime were themselves crafted by officials who have to follow such rules and norms to structure their deliberations and decisions.

This chapter provides a brief overview of the framework we will use in this book to study institutions.

In the following sections we will discuss these core questions:

- Why are there so many different types of institutions?
- How do we analyze institutions?
- What is the appropriate unit of analysis for studying institutions in general, and the commons in particular?
- How do we use one choice of an analytical unit, the action arena, to study institutions?
- What are the core components of an action arena?

2.2 Institutional diversity

During a typical day, we experience many situations in which we interact with others in a structured way where rules and norms may apply. This can be at work, in the classroom, on the sports field, in the supermarket, during commuting, when we bring our kids to daycare, when we watch a movie online, when we go to church, when we eat at the dinner table, etc. In all these different settings different types of norms and rules hold. At work you may have a formal contract regarding the duties that are expected from you and the compensation you are given for undertaking those duties. At the dinner table you may adopt some manners (which are the equivalent of norms) taught to you by your parents and relatives. In traffic you follow the norms and rules of the road. For example, one rule of the road is a speed limit. A norm is you don't cut other drivers off when changing lanes. Can you tell the difference between a rule and a norm just based on these examples? Finally, we interact with many strangers every day whom we expect to follow the same rules.

When you start realizing the number of rules and norms we implicitly deal with on a daily basis, it might become overwhelming. But most of us are easily able to participate in all these diverse sets of situations without thinking too much about the rules and norms that structure them or specific decisions we make in those situations. Several scholars have explored the question of what enables us to do this. Not only are we faced with many different situations each day, the situations we can experience change over the generations. It is likely that today we experience more different types of situations at different levels of social organization as compared to previous generations. People living in a small village in Europe in 1200 AD were not thinking about the implications that political developments in China might have on their lives.

We now expect to communicate with our relatives or be able to check on the latest news wherever we are in the world. Our meals are not restricted by the seasonal availability of foods produced by local farmers. We transport ingredients for our meals from all over the world (e.g., tropical fruits and vegetables in the winter in New

England) at considerable environmental costs. Such changes are not just caused by technological developments, but also through changes in institutions. To make sure fruits and vegetables are transported reliably from location A to B, we have to create institutions to structure repetitive interactions between all the individuals involved. Without institutions, the transaction costs for exchanges between farmers, transporters, and retailers would make long distance transport of food extremely costly.

It is obvious to us what to do when we are shopping in a supermarket. We take the items we prefer from the shelves. We then "arrange a meeting with the cashier," which is made easy by check-out lines and the norms for standing in lines (standing in lines is not a norm everywhere!). The cashier knows we wish to have a meeting by virtue of the fact that we are standing in line – they do not have to ask us why we have come to meet them. We then engage in an exchange with the cashier. What exactly do we exchange with the cashier? Do we exchange food? No—the cashier does not own the food. We exchange information. We may give the cashier a piece of plastic with information on it (a credit card) or we may use cash—which is also a form of information about value and obligation. But this strategy does not work everywhere. When we are shopping in an open bazaar in Asia or Africa, we may bargain over the price of the fruit that is left on the stand at the end of the day. Such bargaining to get a lower price is also happening for other goods in a bazaar. In fact, not bargaining (i.e., not adopting a local norm) for a lower price would be a clear indication that you were a stranger and that you do not know what to do in this situation. This may drastically affect the price of the goods. In this case, in contrast to the grocery store example, the seller may actually own the fruit and thus will be exchanging goods with you. Further, more information is exchanged. In the grocery store example, we believe, or at least accept if we are not willing to shop around at different grocery stores, that the "market" provides information about the correct value of a given item. In the bazaar, haggling over price “reveals” supply and demand prices which, through a very local market interaction (between just two people) drives the price to the ‘market clearing equilibrium price’. Anyone who has haggled over prices can attest extracting this information can be a quite costly in terms of your time. Finally, once the supply and demand prices converge, what will you be exchanging for the goods? Probably not a number and an expiration date on a piece of plastic!

Can you use U.S. dollars, say in Africa? Maybe, maybe not. These examples illustrate that there are many (subtle) changes from one situation to another even though many variables are the same. These subtle changes can have major consequences for the interactions between people.

Figure 4.1: Differences in technology affect the type of institutions that are used. Keeping fruits and vegetables refrigerated increases the shelf life of the product and allow sellers to ask for a better price for a longer period of time.

The types of institutional and cultural factors we have been discussing affect our expectations regarding the behavior of others and their expectations regarding our behavior. For example, once we learn the technical skills associated with driving a car, driving in Phoenix (Arizona) or Bloomington (Indiana)—where everyone drives fast but generally follows traffic rules—is quite a different experience from driving in Rome, Rio de Janeiro, and even in Washington, D.C., where drivers appear to be playing a game of chicken with one another at intersections rather than following traffic rules. Driving in India can seem like a life-threatening experience. Nobody seems to follow traffic rules but there are clear norms such as "the cows are free to go wherever they want, including highways," or "honk when you drive behind somebody so they know," and "expect the unexpected." When playing racquetball with a colleague, it is usually okay to be aggressive to try to win by using all of one's skills. On the other hand, when teaching a young family member how to play racquetball, the challenge is how to help them have fun while they learn a new skill. Being too aggressive in this setting—or in many other seemingly competitive situations—may be counterproductive. A "well-adjusted and productive" adult adjusts their expectations and ways of interacting with others to "fit" a wide range of different situations. Such adjustments are often second nature.

Although we may not explicitly realize it, we have a lot of implicit knowledge of expected dos and don'ts in a variety of situations. Frequently, we are not even conscious of all of the rules, norms, and strategies we follow. Nor have the social sciences developed adequate tools to help us translate our implicit knowledge into a consistently explicit theory of human behavior. In most university courses students learn the language of a particular discipline, from anthropology to economics, from psychology to political science, etc. This disciplinary narrowing of language may hinder our understanding of how to analyze the diverse sets of situations we encounter is social life. The framework we discuss in this book may provide a common language to study these different situations.

Figure 4.2: If you want to buy a ticket for a concert and you see people standing in line, you would automatically join the back of the line. What do you suppose would happen if you bypassed the line to buy your tickets? Although there might not be any formal signs that say you need to wait for your turn, it is generally assumed you understand you have to do this.

4.3 How to analyze institutions?

There are millions of different species on our planet that interact in complex ways at different spatial and temporal scales. How does one study such complexity? One of the breakthroughs in biology is the concept of genes and the discovery of DNA, the building blocks of the diversity of life forms on Earth. Can we develop an equivalent set of concepts for building blocks that create institutions?

If the situations in which people experience different norms and rules are so diverse, how can we study them? How can we make sense out of such social complexity? Given that there is such a large variety of regularized social interactions in markets, hierarchies, families, legislatures, elections, and other situations, is it even possible to find a common terminology to study them? If so, what framework could we use to analyze these different situations across different cultures? Can we learn from one type of institutional arrangement for a particular context and apply the lessons to another context?

Can we identify attributes of the context in which people carry out their repeated interactions in order to find communalities that distinguish success stories from failures? If we are successful with this, we may be able to explain behavior in a diversity of situations varying from markets and universities to religious groups and urban governance. This analysis of interactions among people may take place at a range of levels from the local to the global, and we may analyze whether processes occurring at the local level may explain some of the challenges at the global level.

These are all very ambitious goals. However, as you will see from the material in this book and associated coursework, the framework that we will discuss will help to provide us with a much better understanding of key features that appear throughout a diverse set of situations. The framework is an outcome of many studies conducted at the Vincent and Elinor Ostrom Workshop of Political Theory and Policy Analysis at Indiana University, which was created in 1973 by Vincent and Elinor Ostrom (Figure 4.3). Many of their colleagues all over the world have contributed to this framework by testing it on diverse sets of problems.

Figure 4.3: Elinor and Vincent Ostrom

In the rest of this chapter we will provide a brief overview of the basics of the framework. The framework is called the Institutional Analysis and Development (IAD) framework. One of the aspects of social systems that makes the IAD framework complex is the existence of different types of regularized social behaviors that occur at multiple levels of organization. There is no simple theory that predicts everything, and therefore we need to understand what kind of behavior is to be expected in each type of context.

4.4 Action arenas and institutional analysis

An **action arena** occurs whenever individuals interact, exchange goods, or solve problems. Some examples are teaching a class, playing a baseball game, and having dinner.

When two people exchange a product on eBay, they are in an **action arena**. This is an example of the focal level of analysis we use throughout this book. In an action arena, participants, rules and norms, and attributes of the physical world come together. The latter two elements, the rules and norms and the physical world are said to define an **action situation**. Action situations remain stable over time relative to the participants who may take part. For example, the eBay action situation does not change over the course of a day during which millions of participants can enter the action situation and generate an action arena. As participants interact in the action arena, they are affected by exogenous variables and produce outcomes that, in turn, affect the participants and the action situation. Exogenous variables are those whose characteristics (values or probabilistic distributions) change much more slowly than the relevant time-scale of the action arena. Specifically, actions may change in an action arena in the timespan of minutes or seconds whereas norms may take generations to change. Action situations exist in homes, neighborhoods, regional councils, national congress, community forests, city parks, international assemblies, and in firms and markets as well as in the **interactions** among all of these situations. The simplest and most aggregated way of representing any of these arenas when they are the focal level of analysis is shown in Figure 4.4., where exogenous variables affect the structure of an action arena, generating interactions that produce outcomes. **Evaluative criteria** are used to judge the performance of the system by examining the patterns of interactions and outcomes.

Let's discuss some examples. Consider two participants: John and Alice. When John and Alice play a game of chess, the action situation is composed of (a) the physical game of chess including the board with 64 squares and the pieces: 8 black and 8 white pawns, 2 black and 2 white rooks, 2 black and 2 white knights, 2 black and 2 white bishops, 1 white and 1 black queen and 1 white and 1 black king along with the context in which the chess game is located, i.e. is it at a picnic table outside, a dim room, a cold room, etc.; and (b) the rules of chess—how each piece can be moved, how pieces interact, and what constitutes a victory. When John and Alice sit down at the chessboard to play, this forms an action arena. The interactions between the players may lead to either John or Alice winning the game or a tie. Hence the outcome is whether the game is won by one of them or whether it was a tie. The same persons may also be in an action arena involving money lending. In this action arena, the action situation may be less structured than the chess game. Consider the action arena in which Alice lends money to John. Suppose Alice and John are good friends and the amount of money is small. Alice gives the money to John who agrees to return the money at some specified date (often rather vague in such situations). In this case, the action situation is simple: it is defined by the shared norms of informal money lending and shared understandings of trust and trustworthiness in Alice's and John's culture. Suppose, on the other hand, that this exchange is performed in a formal way. Another participant enters the action arena, a notary public, who formulates a contract that is signed by Alice and John. In this case, the action situation is slightly more complex as it involves a formal contract legitimized by the legal infrastructure in the jurisdiction where the contract is drawn up, the notary's presence, and the signatures of Alice and John. Now the formal rules of contract law, the testimony of a third party recognized by the state (notary) who will testify to the identity of the signatories of a contract, and an entity that will archive the contract form the action situation. The outcome of this transaction is that John receives the money and pays it back according to the conditions as stated in the contract. A third party enforcer ensures that the conditions of the contract are met. A third possible action arena would be an election. Alice and John are both candidates for president of the student association. Within the

action arena participants include all the students of the association who are allowed to vote for one of the candidates. The interactions include debates, a campaign, and finally the election day in which a winner is decided. The evaluation criteria stipulates that the winner is determined based on which candidate has a simple majority (i.e., more than 50%) of the votes. In the last example, Alice and John are neighbors who have a conflict about the barking of Alice's dog. The action situation is a conflict. Within the action arena we have Alice, the dog, John, and the local authorities whom John calls to intervene. Alice and John may both hire lawyers to represent themselves when the action situation (conflict) is played out in court. The interactions include the daily occurrences of the dog barking, the initial friendly requests of John to silence the dog, and the escalation of the conflict into a court case. There are various possible outcomes: either John or Alice moves out of the neighborhood, the dog gets training to stop barking, the dog is sold, John gets a financial compensation for the inconvenience, etc. Each outcome is evaluated differently by each of the participants, including the lawyers. For example, if John's lawyer gets a certain percentage of the financial compensation, she may focus on winning a case to get that financial compensation, although this may lead to long-term bad relations between Alice and John.

Outcomes feedback into the participants in the action arena (the dashed arrow from outcomes to the action arena in Figure 4.4. For example, the fact that a player loses a chess game affects her next decision regarding the action situation of playing chess (play another game or not). The dog continuing to bark after one interaction (John asks Alice to quiet the dog) will undoubtedly affect John's next decision. This changed view by one or several participants may induce the action situation to transform over time as well. Over time, outcomes may also slowly affect some of the exogenous variables. For example, decisions people make regarding energy use creates outcomes including emissions of CO2 which in the long term affect the climate system. In a world with a changed climate the costs and benefits of various human activities are affected, which will affect action arenas. In undertaking an analysis, however, one treats the exogenous variables as fixed—at least for the purpose of the analysis.

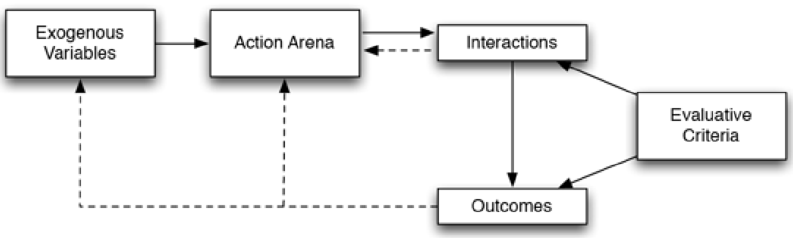

Figure 4.4: The focal level of analysis—an action arena (adapted from Ostrom, 2005).

When the interactions yielding outcomes are productive for those involved, the participants may increase their commitment to maintaining the structure of the situation as it is, so as to continue to experience positive outcomes. For example, wealthy people who may have benefited from low taxes in the past may support tax cuts that the Bush administration introduced. However, if participants view interactions as unfair or otherwise inappropriate, they may change their strategies even when they are receiving positive outcomes from the situation. For example, a group of millionaires requested that President Obama raise taxes for wealthy people.

When current outcomes are perceived by those involved (or others) as less desirable than other outcomes that might be obtained, some participants will raise questions about particular action situations and attempt to change them. But rather than trying to change the structure of those action situations directly, they may move to a different level and attempt to change the exogenous variables. The Occupy Wallstreet movement of 2011 was a protest against the perceived unfairness in society due to a culture of greed by bankers and other participants who control the financial system. The protesters requested a change of the financial system (the exogenous variable) in order to move toward a more equitable society in which they may also succeed (a different outcome). But they didn't try to change the banking system directly. They tried to affect the exogenous variables, for example, the perception of the general public toward the actions of banks.

Figure 4.4 is the simplest schematic representation of an action arena. As you see from the example, there are many important layers to each action arena. We unpack this simple representation in Figure 4.5 in order to make these layers more apparent. An action arena refers to the social space where participants with diverse preferences interact, exchange goods and services, play a game, solve problems, have an argument, receive and deliver health care, etc. We make a distinction between an action situation and an action arena to emphasize that the same participants can fill different roles in different action arenas as we saw with John and Alice. The action situation refers to the positions, actions, outcomes, information and control that provide the structure by which participants interact. Thus the action situation provides the institutional context with which the participants in an action arena are confronted. In Chapter 5, we will zoom in and unpack the action arena. Let's look at a broader overview of the IAD conceptual map.

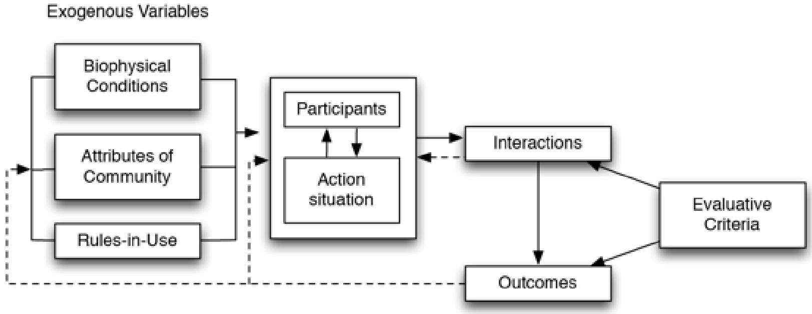

Figure 4.5: A framework for institutional analysis (adapted from Ostrom, 2005).

Let’s apply the framework to a concrete example, namely the course you are taking for which you are reading this book (Figure 4.5). The action situation is defined by the general rules about taking a course at your university (grades, credits, conduct) further specified by the syllabus for this particular course and the characteristics of the space in which the participants meet. Taking this course (along with all the other students) then constitutes an action arena. In the action arena there are a number of different participants, namely the students, the professor and the teaching assistant. The participants interact via lectures, taking exams and writing essays. The syllabus of the course specifies what is needed to receive a good grade in the course. It specifies the weight of the different types of interactions, from participation in class, giving a talk, writing an essay and taking an exam. For each of these activities there are more detailed evaluation criteria on how to receive a good exam grade or writing the essay. The final outcome of the course is a grade.

The exogenous variables in which these interactions take place are the facilities of the university campus (the quality of the classrooms, computer commons, etc.), the attributes of the students (what criteria is required to be admitted to the university, quality of other courses, etc.), and the university regulations. These are specific examples of the general categories of exogenous variables in Figure 4.5: The biophysical conditions, the attributes of the community, and the rules in use, respectively.

Although the final grade is mentioned here as the outcome of the course, this can be debated. If this were truly the only outcome we cared about, the participants could agree (e.g., all vote to give each other an A) that the students could all get an A without putting in the effort of taking the course. Obviously, this is not the purpose of a course and is a violation of university regulations. Although the focus of many participants in the action situation might be on the grade, there are other outcomes that we may include. Does the course material lead to new insights and useful experiences for the students? Do the students comprehend the material and can the students apply this to other topics or problems they may encounter in life? Is the atmosphere in the classroom pleasant and productive? These kinds of outcomes are more difficult and costly to quantify, but are nonetheless very important. However, the difficulty with measuring such outcomes might be a reason that officials may choose to focus on grades to measure course outcomes.

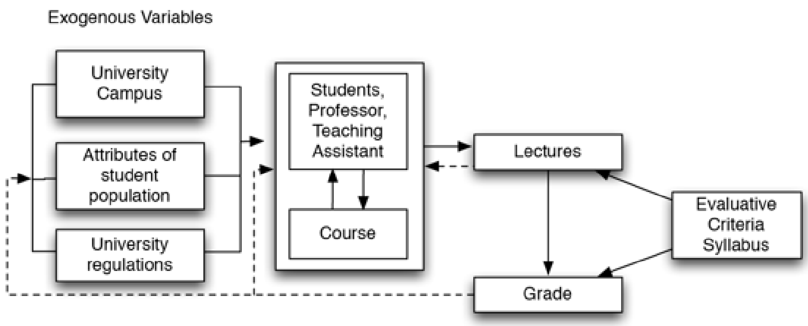

Figure 4.6: Framework applied to taking a course.

4.5 Context of the action arena

The action arena does not occur in a vacuum. Participants are interacting in an action situation which is affected by a broader context. As mentioned above, this broader context is defined by three clusters: (1) the rules used by participants to order their relationships, (2) the biophysical world that are acted upon in these arenas, and (3) the structure of the more general community within which any particular arena is placed.

Different scholarly disciplines focus on different clusters. Anthropologists and sociologists may focus more on the role of the community and culture while economists focus more on how rules affect the incentives of the participants. Environmental scientists may focus more on the biophysical attributes of the action arena. In this book, we focus on the rules, but take into account the role of the community and the biophysical environment.

Rules

Many of the readers of this book are used to an open and democratic governance system where there are many ways in which rules are created. Under these conditions, it is not illegal or improper for individuals to self-organize and craft their own rules for many activities. This may be in stark contrast to more dictatorial states in the world. At work, in a family, or in a community organization there are many ways we experience the crafting of rules to improve the outcomes we can expect in the future. Some of these rules are written down on paper, others are verbal and may be confirmed by a handshake.

In our analysis of case studies in this book, we make a distinction between rules-on-paper (de jure) and rules-in-use (de facto). It is not uncommon that in practice, somewhat different rules are used at the work floor, in the classroom, or on the sports field than those officially written down on paper. For example, a referee in a soccer match may not stop the game for each possible rule infraction, but judge whether the infraction is severe enough to stop the flow of the game and enforce penalties.

Human behavior, including the tendency of humans to comply with rules, relies on the extraordinarily complex structure of neural networks in our brains and, as a result, is not as predictable as most other biological or physical phenomena. Humans are reflexive and have opinions and moral values. They may not necessarily obey instructions from others. All rules are formulated in human language. As such, rules might not always be crisp and clear, and there is a potential for misunderstanding that typifies any language-based phenomenon. Words are always simpler than the phenomenon to which they refer. In many office jobs, for example, the rules require an employee to work a specified number of hours per week. How accurately do we need to specify what the employee will be doing? If the employee is physically at her desk for the required number of hours, is daydreaming about a future vacation or preparing a grocery list for a shopping trip on the way home within the rules? Written rules are always incomplete and therefore the very act of interpreting the rules may lead to different outcomes. Monitoring rule compliance is a challenging activity if rules are not always clear and fully understood. Thus, when we study an action arena, we will look not only at the official rules on paper, but also the rules in use. Misinterpretations may lead to differences between the two. For good performance of institutional arrangements, it is important that the rules are mutually understood.

The effectiveness of a set of rules depends on the shared meaning assigned to words used to formulate them. If no shared meaning exists when a rule is formulated, confusion will exist about what actions are required, permitted, or forbidden. The effectiveness of rules is also dependent upon enforcement. If rules are perfectly enforced then rules simply say what individuals, must, must not, or may do. Participants in an action arena always have the option to break rules, but there is a risk of being caught and penalized. Has the reader ever driven faster than the official speed limit? If the risk is low, rule breaking might be common. Further, because of the feedbacks in action arenas, the likelihood of rule breaking can grow over time. If one person cheated without being caught, others may follow and the level of cheating will increase. This will increase the detection of cheating behavior and more rigorous rule enforcement might be implemented. If the risk of exposure and sanctioning is high, participants can expect that others will make choices from within the set of permitted and required actions.

One of the main benefits that accrue to participants when the majority of people follow the rules is the increased predictability of interactions. Virtually all drivers in the U.S. use the right side of the road to drive almost all the time. If such a rule were not obeyed frequently, imagine how difficult it would be to drive and how ineffective it would be to use the road. Knowing what to expect in interactions with others vastly improves the performance of many social systems.

Biophysical conditions

As we will see throughout the book, the rules affect all the different aspects of the action arena. The biophysical world also has an important impact on the action arena. What actions are physically possible, what outcomes can be produced, how actions are linked to outcomes, and what actors can observe are all strongly affected by the environment around any given action situation. For example, water can't run up hill. Once you say something, you can't retract it. The same set of rules may yield entirely different types of action arenas depending upon the context. For example in New York city there is regulation that residents of buildings are responsible for removing the snow on sidewalks in front of those buildings within four hours after the snow ceased to fall. Why does Phoenix not have such a regulation? We will discuss many case studies in different application domains in this book that will help recognize how context affects decision making and the effectiveness of rule configurations.

The recent COVID-19 pandemic impacted the biophysical conditions in many action arenas. Teaching via Zoom is not the same as in person, and it is expected that the learning outcomes for many students is lower than it would have been in an in-person classroom. However, for some students, online learning works well, and as such it is important to evaluate when what kind of intervention leads to desirable outcomes.

Attributes of the community

A third set of variables that affects the structure of an action arena relate to the attributes of the community of which the participants are members. Examples of attributes that might be important are the shared values within the community, the common understanding and mental models that the community members hold about the world in which they live, the heterogeneity of positions within the community such as class and caste systems, the size of the community, and the distribution of basic assets within the community.

The term culture is frequently applied to the values shared within a community. Culture affects the mental models and understanding that participants in an action arena may share. Differences in mental models affect the capacity of groups to solve problems. For example, when all participants share a common set of values and interact with one another frequently, it is more likely that the participants will be able to craft adequate rules and norms for an action arena. If the participants have different mental models, come from different cultures, speak different languages, have different religions, it will become much harder to craft effective institutional arrangements.

4.6 Critical reflections

Institutions are rules and norms that structure human interactions. They are complex and difficult to study. The Institutional Analysis and Development (IAD) framework helps us organize our thoughts and direct our questions. The focal element of the IAD framework is the **action arena** in which **participants** interact in an **action situation**. These **interactions** lead to **outcomes** which affect decisions made in the next iteration. The interactions are affected by the social and biophysical context in which the action situation takes place.

4.7 Make yourself think

1. Come up with institutions you deal with every day, some you don't like and some that you do like.

2. Do you think banks should be regulated in their lending practices? What are the key elements necessary to address this question?

3. What is the most important outcome for you in taking this class?

4. What can explain the fact that people solve problems differently in India as compared to the U.S.?

4.8 References

Ostrom, E. (2005) *Understanding institutional diversity.* Princeton, NJ: Princeton University Press.

CHAPTER 5

Action Arenas and Action Situations

Key Concepts

In this chapter we will:

- Learn how action situations define the structure of interactions
- See that adding individuals to an action situation leads to an action arena
- Dissect the structure of an action situation

5.1 Action arenas

Whenever two or more individuals are faced with a set of potential actions that jointly produce outcomes, these individuals can be said to be "in" an action situation. Within an **action situation**, a participant occupies a certain position. The same participants can interact in another action situation where they occupy different positions.

An **action arena** combines the action situation, which focuses on the rules,norms and biophysical context, with the participants who bring with them their individual preferences, skills, and mental models, i.e. the attributes of the community. The need to distinguish between action arenas and action situations is a result of the fact that when different participants occupy positions in the same action situation, this may lead to very different outcomes. Put simply, the action situation remains the same for a given period, but a new action arena is generated every time a new set of participants enters the action situation. For example, an action situation might be the marketplace on eBay. The same product offered by different sellers might not lead to the same price since it depends on the preferences and actions of the different participants who enter the action situation and generate a new action arena. Other examples of action situations include resource users who can extract resource units (such as fish, water, or timber) from a shared resource, politicians in congress crafting new laws, and schools with educators and students.

Likewise, the same participants can have very different types of interactions in different action situations. This could be the result of the simple fact that the participants in different action situations occupy different positions. This could also be due to different rules on the information available in different action situations. A boss and their employee in one action situation might become two squash players in another arena. The boss and employee interact very differently in terms of their power relationship—they leave their professional relationship at the squash court door.

The structure of all action situations can be described and analyzed by using a common set of variables. These are: (1) the set of participants, (2) the positions to be filled by participants, (3) the potential outcomes, (4) the set of allowable actions and the function that maps actions into realized outcomes, (5) the control that an individual has in regard to this function, (6) the information available to participants about actions, outcomes, and the linkages between them, and (7) the costs and benefits—which serve as incentives and deterrents—assigned to actions and outcomes. The internal structure of an action situation can be represented as shown in Figure 5.1. In addition to the internal structure, whether a situation will occur once, a known and finite number of times, or indefinitely, affects the strategies individuals adopt. And again, with the same action situation but different individuals participating, we have a different action arena.

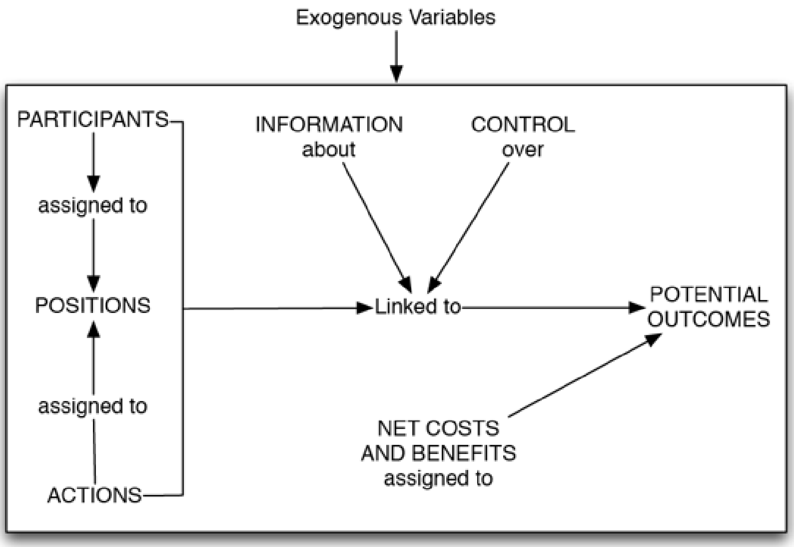

Figure 5.1: The internal structure of an action situation (adapted from Ostrom, 2005).

Within a college course, participants have different positions for which different actions are assigned. Students have different responsibilities compared to the professor and teaching assistant. For example, a professor has information regarding the scores of all the students and the authority to give the grades. Students do not have full information about the scores of individual students in the classroom. They may, however, have aggregate information about all the student scores (i.e., the average). A teaching assistant can grade essays based upon an agreed-upon evaluation criteria, but it is the responsibility of the professor to give the grades. Some of the costs and benefits for a professor include the amount of time spent in preparing the class content and lectures and grading and the wage she receives for doing so. The consequences of different allocations of time invested can be seen in the grades the students receive and the evaluations the professor receives. Also, the student has to balance the investments of time in taking the course and other activities and this choice will be materialized in the grade received.

An individual can take a class one year (be in the position of the student), and become a teaching assistant the next year. That is, the same participant can occupy many different positions. The student could attend a course in the morning as a student, then act as a teaching assistant for a different course in the afternoon. In the morning course, the student has no information about other students' grades. In the afternoon course, she will have more information about the individual students, but now also bears more responsibility for the performance of the students in the class.

The number of **participants** and **positions** in an action situation may vary, but there must be at least two participants in an action situation. Participants need to be able to make choices about the **actions** they take. The collection of available actions represents the spectrum of possibilities by which participants can produce particular **outcomes** in that situation. **Information** about the situation may vary, but all participants must have access to some common information about the situation otherwise we cannot say that the participants are **in** the same situation. The **costs and benefits** assigned to actions and outcomes create incentives for the different possible actions. How these affect the choice of participants depends on the preferences, resources and skills participants have. Who has power? Not all participants may have the same level of **control**, allowing some to have substantial power over others and the relative benefits they can achieve.

There is inequality in wealth between countries and within countries. There are differences in information access and access to decision-makers between the haves and the have-nots. Poor people typically have fewer possible actions available to them than do rich people. With wealth comes access. The rich man in Figure 5.2 can polish his own shoes but can also pay somebody else to do this. The poor man does not even have shoes that need to be polished let alone resources to pay somebody else to polish the shoes. This example shows that not every person can occupy each possible position in an action situation. Both men can occupy the positions of citizen, or legal adult, but only the wealthy man can be both polisher and “polishee.” The fact that the poor man cannot be the “polishee” is due to one of two factors (1) formal rules or norms about social roles and occupations allowed for different social roles such as the caste system in India or (2) because the poor man lacks the resources or capacity to be the “polishee.” In the first case, it is the action situation that limits the actions of the poor man. In the second case, it is an outcome of the action arena that limits his choices. Whether an individual can occupy a certain position may be affected by wealth, education, elections, inheritance, passing a test, age, gender, and many other criteria. As we will see in later chapters, the rules that affect positions play an important role in how communities can sustain their commons.

Figure 5.2: A poor man polishing the shoes of a wealthy man

When we study an action situation, we analyze the situation as given. We assume the structure of the action situation is fixed for at least the short run. Then we can analyze the action arena by exploring assumptions of the likely human behavior of the individuals leading to particular outcomes.

Within a particular situation, individuals can make choices about their own actions. However, in the longer term, individuals may—at least those who are living in an open society—take actions that may eventually affect the structure of action situations (i.e., the choices others can make). This is possible when one is able to change the rules affecting the action situation. For example, the rules regarding the marketplace at eBay have changed over time because participants have learned what works and what does not work. If action situations do not lead to good outcomes, one may attempt to change the rules. To do so, they must move to action situations at a higher level of decision making such as collective-choice or constitutional-choice action situations, where the outcomes generated are changes in the rules that structure other action situations such as who can participate, what actions are available to them, what payoffs are associated with actions, etc. In a closed society, individuals at an operational level may have little opportunity to change rules at any level and may find themselves in highly exploitative situations. Democratic countries are examples of open societies, while dictatorships are examples of closed societies (see Figure 5.3). We discuss the process of shifting to higher-level action situations in the last half of this chapter.

Figure 5.3: North Korea is an example of a closed society. Citizens have no control over the decisions they can make, have no control over the information they receive, and have no ability to change regulations.

5.2 The basic working parts of action situations

Let us now discuss the elements of an action situation so we can begin to understand what is common to all of the interactive situations we may observe or experience in our lives.

5.2.1 Participants

Participants in an action situation are assigned to a position and capable of making a choice between different possible actions. The participants in action situations can be individuals but also corporate actors such as nations, states in a federal system, private corporations, NGOs, and so forth. Whenever participants are organizations, one treats them in the situation as if they were a single individual but one that is linked to a series of additional situations within their own organization. When one is interested in the outcome of an action situation for the organization, we may ignore the linked situation and just focus on the strategy of the organization as an individual actor. However, if we notice that there are problems with the functioning of an organization within an action situation, we may look at the functioning of the organization itself, and study the action arena of that organization. As such, action arenas can be composed of action arenas of lower level actors. For example, the United Nations consists of many countries. To understand the functioning of the action arena of the United Nations, we may look at the ambassadors as participants, or we may look into the action arena of a country in which the ambassador also participates to understand the decisions made by that ambassador.

Several attributes of participants are relevant when representing and analyzing specific situations. These include (1) the number of participants, (2) their status as individuals or as a team or composite actor, (3) and various individual attributes, such as age, education, gender, beliefs, knowledge, skills, and experience.

The number of participants

The focus of this book is on those action situations associated with governing shared infrastructure. Therefore, action situations that are of interest to us require at least two participants where the actions of each affect the outcomes for both. This could be two farmers sharing a water source.

The specific number of participants is often specified in detail by formal regulations, such as for legislation (number of seats in the Senate and Congress), juries (number of jury members), and most sports (number of members on a team). Some descriptions of a situation, however, specify the number of participants in a looser fashion such as a small or a large group, or face-to-face relationships versus impersonal relationships. Since many other components of an action situation are affected by the number of participants, this is a particularly important attribute in the analysis of any action situation. Figure 5.4 shows some examples from sports illustrating the number of participants in action situations.

Figure 5.4: Different numbers of players in the action situation: a) two sumo wrestlers, b) eleven players per soccer team c) 200 cyclists in a stage of the Tour de France, and d) more than 10,000 runners in a marathon.

The individual or team status of participants

Participants in many action situations may be individual persons or they may represent a team or composite actors, such as households. A group of individuals may be considered as one participant (a team or organization) in a particular action situation. What might be the conditions in which it makes sense to treat a group of individuals as a participant?

To consider a group of individuals to be a participant, one must assume that the individuals intend to participate in collective action. One needs to assume that the individuals who are being treated as a single actor intend to achieve a common purpose. Sometimes there are groups of individuals who share many similar characteristics, such as "veterans," "urban voters," or "legal immigrants," but they have different individual preferences and do not act as a cohesive team. Corporate actors, such as firms, are not so dependent on the preferences of their members and beneficiaries, because they are legally defined as an individual entity. The activities in firms and organizations are carried out by staff members whose own private preferences are supposed to be neutralized by formal employment contracts.

A fully organized market with well-defined property rights, for example, may include buyers and sellers who are organized as firms as well as individual participants. Firms are composed of many individuals. Each firm in a market is often treated as if it were a single participant.

So when do we consider a group of individuals as a collective rather than as a bunch of individuals? This depends on the questions we have. The action arena of a basketball game, for example, maybe represented as having either ten participants or two teams composed of five individuals. If we are interested in studying a league or a tournament, we will include more teams and focus on the teams as participants rather than as individual players. If we are interested in the performance of a single team within the league we would look at individual players, coaches, trainers, and owners.

Attributes of participants

Participants differ in their characteristics such as their skills, ethnic background, education, gender, values, etc. These characteristics may influence their actions in some situations, but not in others. The educational level of participants is not likely to affect the actions of drivers passing one another on a busy highway. But when the participants meet each other in an emergency room in the positions of patient and physician, education becomes an important attribute. Whether gender or ethnic background is important varies between cultures and countries. In some cultures, female patients are not allowed to be examined and treated by male physicians. During the Apartheid regime in South Africa, Black patients did not receive the same treatment as White patients did.

The outcomes of many situations depend on the knowledge and skills of the parties. Experienced drivers will have on average a different driving style compared to younger drivers. This fact is born out in the differences between insurance policies for the two participants. Drivers who have a reputation with an insurance company for getting involved in accidents will have to pay a higher insurance premium compared to those who are accident-free.

5.2.2 Positions

Participants occupy positions in action situations. Examples of positions include students, professors, players, referees, voters, candidates, suspects, judges, buyers, sellers, legislators, guards, licensed drivers, physicians, and so forth. It is very important to understand that “positions” do not refer to people, but rather to **roles** that participants can play in an action situation. For example, in a market situation (in your local mall), the same person may be a “seller” when they are at work at the Apple Store helping customers choose their latest iPhone, and a “buyer” when they go for lunch in the food court. Thus, positions and participants are separate elements in a situation even though they may not be clearly so identified in practice.

In practice, the number of positions is frequently significantly less than the number of participants. In a class, there are typically only two positions—student and professor—while there may be hundreds of participants. Hunters who have a valid license all occupy the position of a licensed hunter; and while there are more than a billion participants at Facebook, there are only a limited number of positions (such as the person featured on the webpage, or the administrator of a page representing an organization).

Depending on the structure of the situation, a participant may simultaneously occupy more than one position. All participants will occupy whatever is the most inclusive position in a situation — member, citizen, employee, and the like. In a private firm, additional positions such as foreman, division manager, or president will be occupied by some participants while they continue to occupy the most inclusive position—that of employee. Some examples of positions are given in Figure 5.5, where some positions are filled by election, others are filled by selection after an interview.

Figure 5.5: a) Barack H. Obama in the position of President of the United States of America, b) the U.S. Supreme Court, c) judges on the television show American Idol, and d) a police officer.

Positions connect participants with potential actions that they may take in an action situation. Not all positions have the same potential actions. A surgeon can do surgery on a patient. It is likely not advisable to allow the patient this potential action, unless in the unlikely event that the patient happens to be a surgeon too. Other positions are less restrictive. Every person with a driver's license shares a large set of potential actions with every other driver. Some drivers have special positions and additional potential actions, such as drivers of ambulances or large trucks.

The President of the United States can sign a bill into new legislation, which confirms that the new legislation will be implemented. The President can only sign such a document under particular conditions (agreement in the Senate and Congress), but a signature of a regular citizen does not have the same effect. A U.S. citizen who has registered as a voter can vote, but a permanent resident (Green Card holder) cannot register as a voter, even if such a permanent resident is a professor at a prominent university.

The nature of a position assigned to participants in an action situation both defines the set of authorized actions and sets limits on those. For example, licensed drivers may operate a motor vehicle on a road or highway, but this action is also restricted by speed limits. Those who hold the position of a member of a legislative committee are authorized to debate issues and vote on them. The member who holds the position of chair can usually develop the agenda for the order of how issues will be brought before the committee or even whether a proposal will even be discussed. The order of events on this agenda may affect how the votes turn out.

If you do a group project how do you organize a group? Will different group members have different roles? Is one of the members leading the discussion?

Participants may occupy different positions, but which position one can hold is not always something a participant can choose. A defendant in a criminal trial does not control her movement into or out of this position. A candidate for the U.S. Congress can certainly influence her chances of winning an election and securing the position, but does not have full control. In the end, this decision is in the hands of voters. Holding the position of a pedestrian in traffic is available without much limitation to most people. Individuals have to compete vigorously for getting a tenured professorship at a university, but once obtained, they may hold their positions for life, subject only to legal actions. This might be true for universities in the U.S., but in most European countries professors are required to retire at the age of 65 and are removed from their position as a tenured professor.

5.2.3 Potential outcomes

In the case of health care reform, there are different potential outcomes that can be discussed: total costs of healthcare, access to health care, distribution of costs and benefits, quality of healthcare, etc. Which outcome will weigh most in the design of policies is a political decision.

When we want to understand how rules, attributes of the environment, or attributes of the community change an action situation, careful attention must be given to how participants value certain outcomes. If there is a market where goods are exchanged at known prices, one could assign a monetary value to the goods. If there are taxes imposed on the exchange of goods (a sales tax), one could represent the outcomes in a monetary unit representing the market prices minus the tax. If one wanted to examine the profitability of growing rice as contrasted to tomatoes or other cash crops, one would represent the outcomes in terms of the monetary value of the realized sales value minus the monetary value of the inputs (land, labor, energy, fertilizer, and other variable inputs).

To examine the effect of rules, one needs to distinguish the effect of material rewards from financial values. For example, the physical amount of goods produced during a particular time period is different from the financial rewards to workers and owners for that time period. If no goods are sold, the financial rewards for the owner might be negative, but the worker may still receive a reward in exchange for the hours worked to produce the goods. Besides monetary values and physical quantities of goods, participants also have internal values, such as moral judgment, that they use to examine potential outcomes. Gun ownership can be evaluated based on the numbers of different types of guns owned, the monetary value of the gun collection, and the moral value placed on gun ownership.

Frequently the outcomes are assumed to be the consequence of self-conscious decisions, but there can also be “unintended outcomes." For example, oil spills in the Gulf of Mexico are not an intended outcome of operations of oil companies (Figure 5.6).

Figure 5.6: The BP Deepwater Horizon oil rig ablaze.

5.2.4 Actions

Participants assigned to a position in an action situation must choose from a set of actions at any particular stage in a decision process. An action can be thought of as a selection of a setting or a value on a control variable (e.g., a dial or switch) that a participant hopes will affect the outcomes. The specific action selected is called a choice. A complete specification of the actions, taking in all possible variations of the action situation is called a strategy. It is important to note that it might not always be clear to participants what all the valid actions are in an action situation. A switch may clearly indicate two different positions, but sometimes participants are much more innovative in the use of possible actions in an action situation. Calling somebody with a mobile phone can be very expensive in some places. As a result, people may use the technology in a different way than was intended by the manufacturers. For example, in some communities, signaling systems evolved where the receiver of a mobile phone call can understand the message from simply counting the number of rings of the call and the information about who the caller is from the display. Users in such situations may seldom use their mobile phone for an actual voice call.

What is the change of outcome for the mobile phone service provider when people use this strategy?

Figure 5.7: Possible actions.

5.2.5 Control

The extent to which participants have control over aspects of the action situation vary widely. Obviously, the position that a participant occupies affects the power of this participant (her ability to affect the actions of other participants and outcomes). The level of control a participant has can, therefore, change over time, for example, if she changes her position. Barack Obama acquired a new repertoire of actions and control when he assumed office on January 20, 2009. And this repertoire has changed over the years from being a lecturer and giving grades at the University of Chicago, a community organizer, a member of the Illinois Senate, and a member of the U.S. Senate. Each position held certain duties and rights. Leaving a position also means losing the duties and rights that hold to the specific position, as happened too with President Obama on January 20, 2017. The issue of control, or 'controllability' is a very general feature of coupled infrastructure systems as we will discuss in further detail in Chapter 8.

5.2.6 Information about the action situation

What is the information participants have in an action situation? In an extreme case they have complete information and know the number of participants, the positions, the outcomes, the actions available, how the actions are linked to outcomes, the information available to other players, and the payoffs available. If they know what other participants will do, participants are said to have perfect information. Of course, perfect information is an extreme case especially when people make their decisions privately. Often there is no perfect understanding of how actions will lead to outcomes, or what others plan to do. Even if people communicate and negotiate what everyone will do, the actual actions may turn out different since people make mistakes or cheat.

In many situations, there is asymmetric access to the available information. For example, in work situations, a boss cannot know exactly what employees are doing. That is why providing an incentive to increase productivity is a challenge. The same holds for insurance companies. Your insurance company does not have perfect information about your driving abilities and health conditions, but makes an informed guess based on statistics of historical events. Would you like your insurance company to have access to your genetic profile? What about your driving behavior? The car insurance company Progressive allows customers to join a voluntary program where a device is installed in your car to track your driving style. One can save a significant amount on their car insurance with proper driving style.

As with 'controllability' mentioned above, the capacity to gather information and access to that information is also a very general feature of coupled infrastructure systems. The general notion for the availability of good information is referred to as 'observability' in general systems theory as we will discuss in Chapter 8.

5.2.7 Costs and benefits

To evaluate the outcomes of the actions taken in the action situation we have to look at the costs and benefits. These costs and benefits accumulate over time. Not all participants will experience the same costs and benefits. Sometimes the positions that participants hold affect their cost and benefits since it affects the compensation, penalties, fees, rewards, and opportunities. A physician receives a monetary benefit from doing a treatment while the patient will pay to receive an improvement in their health condition. Even if participants hold the same position, like players on a sports team, their rewards vary as defined by their individual contracts.

If we study action arenas we need to make a distinction between the physical outcome and the valuation that a participant assigns to that outcome. In economics, the value assigned by participants is often referred to as utility. Individual utility is a summary measure of all the net values to an individual of all the benefits and costs of the outcome of a particular action situation. Utility might increase with an increase in profit, but depending on the study at hand, it may also include elements like joy, shame, regret and guilt.

For example, driving above the speed limit can save you time. However, if you are caught you will have to pay a traffic fine (Figure 5.8). You may challenge a ticket by appearing in court, yet this will take time and may have other costs associated with it. Paying the fine (accepting guilt for the traffic violation) could also result in higher car insurance and accumulate points on your driving record. If you accumulate too many points on your driving record, your license may be suspended. Not paying a ticket in time will lead to additional penalties.

Figure 5.8: The monitoring of speeding.

5.2.8 Linking action situations

In reality, people make decisions in different action situations that are often linked together. Rarely do action situations exist entirely independently of other situations. For example, new laws in the

U.S. need to be approved by the Congress and the Senate before the President may sign it. Signing a bill is meaningless unless the bill has successfully passed through the Congress and Senate action arenas.

Given the importance of repeated interactions to the development of a reputation for reciprocity and the importance of reciprocity for achieving higher levels of cooperation and better outcomes over time, individuals have a strong motivation to link situations.

Action situations can be linked through organizational connections. Within larger organizations, what happens in the purchasing department affects what happens in the production and sales department and vice versa. Sometimes action situations are structured over time. For example, a tournament or sport competition is a description of how players (e.g., tennis) or teams (e.g., basketball) will proceed through a sequence of action situations. In other examples, action situations are not formally linked. Farmers who have successful innovative practices in deriving better profits are frequently copied by others.

Another way in which action situations can be linked is through different levels of activities. We can distinguish three levels of rules that cumulatively affect the actions taken and outcomes obtained:

- *Operational rules* (or rules-in-use in the IAD Framework) directly affect day-to-day decisions made by the participants in any setting. These can change relatively rapidly—from day to day.
- *Collective-choice rules* affect operational activities and results through their effects in determining who is eligible to be a participant and the specific rules to be used in changing operational rules. These change at a much slower pace.
- *Constitutional-choice rules* first affect collective-choice activities by determining who is eligible to be a participant

> and the rules to be used in crafting the set of collective-choice rules that, in turn, affect the set of operational rules. Constitutional-choice rules change at the slowest pace.

An example of an operational-level situation is a group of fishers who decide where and when to fish. At the collective-choice level the group of fishers may decide on which seasons or locations to implement bans on fishing. At the constitutional-choice level decisions are made regarding the conditions required in order to be eligible for membership in the group of fishers.

Figure 5.9 illustrates the different levels of rules related to a class at a university. Within a classroom, decisions are made based on the rules set in the syllabus. Day-to-day decisions include what the assignments are for next week, who will give a talk, and when students can come to office hours. In order for a regular course to be approved, a committee (upper right) will review the proposed syllabus and make a recommendation to approve or not approve the course. The committee also solicits comments of departments that provide similar courses to avoid potential conflicts. The university senate will come into play when new degrees are proposed (middle left). Finally, the upper administration of the university will be involved in decisions that have university-wide impact, such as changing tuition rates. Such a tuition raise will have to be approved, at least for public universities, by a state level committee.

Figure 5.9: Different levels of rules related to a university class (clockwise): a) day-to-day assignments; b) syllabi are approved by committee; c) university senate approve new degrees; and d) upper administration makes decisions that have a university-wide impact, such as tuition rates.

5.2.9 Outcomes

It is difficult to predict the outcomes of rule changes made in action situations. Changing the rules in one action situation may have consequences in other action situations. The difficulty of predicting the consequences of changes shows that we have to closely observe what is happening before rules are changed and after rules are changed. This suggests that we should view policies experiments, and closely observe these experiments in order to learn and have a better understanding of what will happen in a similar case in the future.

Besides the difficulty of predicting outcomes, how to evaluate outcomes is also often not immediately evident. There are different criteria that one can use to evaluate the outcomes:

- Economic efficiency—what are the costs relative to the benefits?

- Equity—how are costs and benefits distributed among the participants?
- Accountability—are participants in leadership positions accountable for the consequences of their decisions?
- Conformance to general morality, i.e. procedural justice—are the procedures fair, is cheating detected, and are promises kept?
- Sustainability—how do the outcomes evolve over time? And what are the consequences of decisions on the underlying system?

In order to evaluate the outcomes one needs to evaluate trade-offs associated with the different criteria. If some groups are affected differently than others, it will be important to define procedures in the collective-choice or constitutional-choice rules to address such differences. For example, the outcome of changing the criteria for student-loans will not have the same consequence for each individual student. It would be important to consider the different types of outcomes for different types of participants and develop agreements regarding how to evaluate such outcomes.

5.3 Critical reflections

The concept of **action arenas** was the main topic of this chapter. An action arena consists of an **action situation** that defines the structure of interactions, actions and outcomes, and the individuals, organizations or nations who may participate in the action situation. When two or more participants interact, there is an action arena where participants hold **positions**, and can make decisions. Not everybody in an action situation can take the same **actions**, or has the same level of **information**. The consequences of the actions are the **outcomes** of the action situation, which can be **evaluated** differently by each participant in the action situation.

5.4 Make yourself think

1. What positions do you hold in different action situations? Provide some examples.
2. What is an action situation you experience regularly? What are the possible outcomes in this situation? What actions can you take? Be sure to distinguish between actions you may take and choices you do make.
3. Do you have an example from your own personal experience where you have experienced the same action situation but with different participants that led to a different action arena and a different outcome?

5.5 References

Ostrom, E. (2005) *Understanding institutional diversity.* Princeton, NJ: Princeton University Press.

CHAPTER 6

Design Principles to Sustain the Commons

Key Concepts

In this chapter we will:

- Be introduced to design principles for effective institutions
- Learn about boundary rules, monitoring, graduated sanctions and input of local participants to collective-choice arrangements
- See how violations of design principles may lead to corruption and rent-seeking

6.1 Introduction

What is the best set of rules to govern a particular type of shared infrastructure? This is the ultimate question asked by policymakers involved with managing common-pool resources and providing public infrastructure. Unfortunately, there is no such optimal set of rules. What we have learned over the years is that there are some "design principles" that help explain why some communities are successful while others are not.

These design principles are based on a systematic study of many case studies of fisheries, irrigation, groundwater, and forestry systems. Information was collected from each case study regarding the size and composition of the community, the formal rules and norms in use, how the system was monitored and by whom, the conflicts that arose, and how the resource system evolved over time.

Elinor Ostrom and her team studied hundreds of these case studies in the 1980s and proposed the design principles in her classic 1990 book Governing the Commons. She initially focused on determining which rules were best but was unsuccessful in identifying a particular set of rules that were "best" in all circumstances. Instead, she turned her efforts toward identifying eight underlying design principles that characterized case studies of long-lasting common-pool resource systems. The design principles she identified were mostly met in these long-lasting systems, but were absent in those that collapsed.

The concept of eight design principles was an initial proposal in 1990. Twenty years later, analyses of about 100 case-studies provide evidence that the design principles hold up when challenged with data (Cox et al., 2010; Baggio et al., 2016).

6.2 Institutional design principles

The design principles derived from case studies of long-lasting systems of common-pool resource governance:

1. **Clearly defined boundaries**. The boundaries of the resource system (e.g., irrigation system or fishery) and the individuals or households with rights to harvest resource units are clearly defined.
2. **Proportional equivalence between benefits and costs**.

Rules specifying the amount of resource products that a user is allocated are related to local conditions and to rules requiring labor, materials, and/or money inputs.

3. **Collective-choice arrangements**. Many of the individuals affected by harvesting and protection rules are included in the group that can modify these rules.
4. **Monitoring**. Monitors, who actively audit biophysical conditions and user behavior, are at least partially accountable to the users and/or are the users themselves.
5. **Graduated sanctions**. Users who violate rules-in-use are likely to receive graduated sanctions (depending on the seriousness and context of the offense) from other users, from officials accountable to these users, or from both.
6. **Conflict-resolution mechanisms**. Users and their officials have rapid access to low-cost, local action situations to resolve conflict among users or between users and officials.
7. **Minimal recognition of rights to organize**. The rights of users to devise their own institutions are not challenged by external governmental authorities, and users have long-term tenure rights to the resource.

For resources that are parts of larger systems:

8.**Nested enterprises**. Appropriation, provision, monitoring, enforcement, conflict resolution, and governance activities are organized in multiple layers of nested enterprises.

We will now look at each design principle in more detail. Although the design principles are intended for natural resources, we will illustrate their use with other kinds of systems too.

6.3 Examples of design principles

6.3.1 Well-defined boundaries

The first principle relates to two types of boundaries: the biophysical boundaries of the resource system and the boundaries that determine which individuals or households have rights to harvest the resource. Having clear boundaries is critical to addressing the problem of free-riding. If there are clear boundaries indicating who can appropriate which resources, it will be clear who is following the rules or not when harvesting activities are observed.

These two types of boundaries are demonstrated in our roadway systems. Not having a valid drivers' license is a clear violation of a boundary rule of driving a car on the road. Roads themselves benefit from clear demarcations on the road to define the lanes.

Such clear demarcations also help in sports to define the boundaries of the common space for competition. What would John McEnroe do if there were no clear lines on the tennis court? In many sports, players wear a uniform and a number that indicates that they are allowed to participate. You cannot just run in the New York marathon. You have to sign up and wear your start number and start in your assigned position. Also, many internet services require that you register in order to be able to use those services. Watching a streaming Netflix movie in the U.S. is possible if you have paid your monthly fee, but it is not possible to do this on the same account on your laptop in China. The reason for this is that the rights for consuming Netflix content are clearly defined for certain countries for which Netflix owns copyrights. When your laptop uses an IP address outside the U.S., Netflix will not provide access to the content.

When resource users create boundaries, they can make use of the ecological context. For example, defining specific areas with easily observable landmarks where one can harvest from the resource makes it easier to monitor and enforce.

Clearly defined boundaries are not enough for a successful institutional arrangement. One also needs to be able to defend boundaries against potential intruders. But the research shows that clearly defined boundaries are a prerequisite for success.

Figure 6.1: Nowadays technology helps to determine whether a tennis ball is in or out.

6.3.2 Proportional equivalence between benefits and costs

Do the rules allocate the benefits from the resource in proportion to the costs of effort people put into harvesting and maintaining the resource? If some users get all the benefits and pay a small proportion of the costs, other users might not be willing to follow the rules over the long term. This design principle relates to fairness. Yet some inequality in the benefits people derive from the resource are acceptable provided the cost they bear in relation to the benefit is proportional. For instance, users may accept that the individuals who put a lot of effort into building an irrigation system get a better plot of land to grow his/her crops.

Salary levels are a typical topic of debate, especially in periods of economically challenging times. For the sustainability of an organization, it is not wise to lay off employers while the boss keeps her seven-figure salary. Similar levels of sacrifices are expected over the whole organization.

There are important differences between countries in how costs and benefits are defined. For example, because the labor market in the U.S. is more flexible it is easier to lose a job and there is only a limited safety net. Within Europe, social security provides a minimum level of income for people who have no job. How a society allocates its costs and benefits—as exemplified in their rules—is largely defined by prevailing social norms. Fairness within the U.S. relates to receiving benefits based on effort, while in Europe fairness relates to equality among people independent of effort. This is a gross simplification, but as the authors can attest—since both of us have lived on both continents—there are important differences in social norms that affect the rules on allocation of costs and benefits.

6.3.3 Collective-choice arrangements

People who are affected by institutional arrangements should have a way to participate in making and modifying the rules. Action situations where local resource users are able to devise rules are able to better tune those rules to local circumstances. Further, when participants make the rules together, they often receive more careful consideration by participants. When rules are viewed as imposed by an elite, participants are less likely to comply with them.

Universities typically include students in their governance systems, and sport federations have athletes on their boards. These are all ways to include relevant knowledge in the decision making process—the perspective of students and athletes—and increase the likelihood that the rules are accepted by the broader community.

In larger organizations, or even states, countries and the global community, the ability of people to participate in rule crafting seems difficult. Sometimes we may think the only thing we can do is to vote for who represents us in the decision-making process. One of the big challenges we face in modern times is the large scale of the communities we participate in.

6.3.4 Monitoring

How is monitoring of the rules organized? Effective monitoring is not only a matter of counting the number of guards, referees, and policemen and making sure this number is high enough. The subtleties of what these monitors actually do to monitor and enforce the rules and the incentives they face are critical. If guards in a national park are not paid well, it is not surprising that they might accept bribes to look the other way when poaching is happening. It is often more effective to have local people as monitors. This may assure conditionally-cooperative resource users that somebody is genuinely checking on the conformance of others to local rules. In urban areas neighbors sometimes organize themselves in a neighborhood watch instead of hiring security personnel.

To understand whether monitoring will be effective, we need to understand what the incentives are for a monitor. Will a monitor be paid independent of whether rules are broken or not? Will a monitor be affected by rule breaking (are they a local resource user themselves). Will somebody notice whether the monitor is doing his or her job or not? If there is a lot of variation in harvest levels, people may not notice in the short run whether rules are broken or not.

Sometimes monitoring is so important for the functioning of the system that high investments are made in monitoring. In some highly productive fisheries in the Artic, each boat has an official of the federal agency NOAA (National Oceanic and Atmospheric Administration) on board. Does this sound outrageous? This is like having a police officer riding in every car to make sure drivers don't break the rules. What about all the official and unofficial referees (general public) during a football match?

6.3.5 Graduated sanctions

People make errors. When you make a mistake you get a warning. If you keep bullying a player on the other team you may receive a yellow card. And if you continue ignoring the rules you may get sent off the field with a red card or even be expelled from the league.

For many action situations there is a graduated sanctioning system. One reason is that rules are not always commonly understood or known, and getting a warning when a rule is broken may remind people of the actual rules in use. Another reason is the potentially high costs of strict enforcement. What would happen to a sport if there was strict enforcement of the rules? There are norms of fair play, and a tolerance of players exploring the boundaries of the formal rules. This is especially important if the assessment of whether a rule is broken necessarily involves some subjectivity.

6.3.6 Conflict-resolution mechanisms

The goal of conflict resolution mechanisms is to provide access to rapid, low-cost, local opportunities to resolve conflict among users or between users and officials. Rules, unlike laws of nature, have to be understood in order to be effective (the laws of nature function whether or not we understand them—right?). There are always situations in which participants can interpret a rule that they have jointly made in different ways. By devising simple, local mechanisms to get conflicts aired immediately and resolutions that are generally known and accepted in the community, the number of conflicts that reduce trust and cost time and money can be minimized. If individuals are going to follow rules over a long period of time, some mechanism for discussing and resolving what is or is not a rule infraction is quite necessary to the continuance of rule conformance itself. For example, within Wikipedia there can be eruptions of editing wars and designated editors can mediate between the different parties to resolve the conflict.

Not all disputes within the U.S. appear in front of the Supreme Court. Most conflicts can be resolved informally by having a good discussion in a neutral environment. By having a drink with your colleague or neighbor, you may discover that the conflict is mainly caused by a misunderstanding. When conflicts are not resolved in informal ways, more formal procedures are possible, such as through your homeowners' association, your company, the court system, etc. The importance of cost effective conflict resolution cannot be understated. Consider the resources expended on litigation in the U.S. Litigating minor conflicts in hope of financial gain is enormously costly to society.

6.3.7 Minimal recognition of rights

When local users can organize themselves to craft their own rules, do national and local government entities recognize and respect these arrangements? The lobster fishers in Maine organized themselves in the 1920s and 1930s after the lobster population almost collapsed. The rules devised by these organizations were informal arrangements among fishers. In the 1990s the federal government wanted to reorganize fishery regulations along the east coast of the U.S. and have all fishers in all states comply with the same regulations. The well-functioning lobster fishery system would have been negatively affected by this, and significant efforts by fishers and scientists resulted in their informal arrangements receiving legal recognition, permitting them continue as they had within the framework of modern regulations.

6.3.8 Nested enterprises

When systems are larger, it may be necessary to have systems of governance at different levels. What might be needed is a “polycentric” (many centers of governance and authority) system. Every neighborhood may need to have policemen to patrol the streets, but not each neighborhood needs to have a crime lab. One crime lab for the whole city might be a better solution in terms of having the specific expertise available at a reasonable cost.

Polycentric systems emphasize approaching problems at the right level and ensuring that all parties with some control over outcomes (centers of power) are involved. This may mean that some collective action problems are addressed locally, while others are addressed at a regional or national level. For example, in resolving disputes on the use of water from the Colorado River, it is not productive to have Los Angeles and Phoenix debate this topic alone. Several states and the Bureau of Reclamation need to be involved in order to develop a meaningful plan for allocating the available water.

There are various challenges with polycentric systems. What is the right level of governance for each problem? Some lower-level communities might not be able to self-organize, or they may be dominated by local elites. How do higher-level authorities facilitate the local level governance systems to succeed?

In a polycentric system, some units are general-purpose governments while others may be highly specialized. Self-organized resource governance systems in such a system may be special districts, private associations, or parts of a local government. These highly specialized governance units are nested in several levels of general-purpose governments that provide civil, equity, as well as criminal courts.

A university is often organized as a polycentric system. Each department has a certain level of autonomy in offering courses. New courses can be offered, but to get them permanently on the books and count for college-wide credits, they have to be approved at the college level. Some basic requirements hold for all the majors, while the rest of the course work can be tailored to the specifics of the major.

Besides departments, there are specialized organizations and services such as fraternities and sororities, research institutes, financial aid offices, and libraries. Although there is a university president who oversees the whole university, most units have a lot of flexibility within the general constraints set by the higher levels of authority. If there is a lack of autonomy such that each decision at the local level requires approval from "above," a large organization will grind nearly to a halt because of the transaction costs associated with decision making and many decisions will be made by higher-level officials without proper knowledge of the detailed practical problems associated with the decision.

6.4 Using design principles in practice

The design principles were originally proposed as hypotheses based on analysis of several case studies. In the more than 30 years since they were developed, they have held up to scrutiny. Although some people may interpret the design principles as blueprints for designing robust institutional arrangements, they are not. They are observed regularities derived by looking at cases after the fact. So how can we use the design principles in practice?

One way to use design principles is to translate them into questions concerning how to improve institutional arrangements for governing the commons. For example, for local resource users we can ask:

- How can we better define the boundaries of this resource and of the individuals who are using it so as to make clear who is authorized to harvest and where harvesting is authorized?

- How can we clarify the relationship between the benefits received and the contributions to the costs of sustaining this system?
- How can we enhance the participation of those involved in making key decisions about this system?
- Who is monitoring this system and do they face appropriate incentives given the challenge of monitoring?
- What are the sanctions we are authorizing and can they be adjusted so that someone who makes an error or a small rule infraction is sufficiently warned so as to ensure longer-term compliance without our trying to impose unrealistic sanctions?
- What local and regional mechanisms exist to resolve conflicts arising over the use of this resource?

For design principles seven and eight, questions need to be addressed at a higher level of governance.

- Are there functional and creative efforts by local appropriators to craft effective stewardship mechanisms for local resources that should be recognized?
- How do we create a multiple-layer, polycentric system that can be dynamic, adaptive, and effective over time?

These are not, of course, the only questions local resource users and officials should ask in an effective process of crafting institutional arrangements, but they can be thought of as a good beginning.

6.5 Threats to sustainable use of the commons

Even as we look carefully to the design principles for guidance, there are various threats to the sustainable use of common-pool resources. We list a number of them below based on what we have learned from many case studies.

6.5.1 Rapid exogenous changes

Rapid changes in technology or population numbers can become a challenge to effective governance of the commons. A new technology that enables fishers to catch more fish with the same amount of effort, may render rules on where and when to fish useless unless restrictions on gear are implemented. The provision of music, movies and books has experienced many challenges during the last few decades with the emergence of new physical devices to digital files. These challenges have caused the need for change in the original business structure of creative activities in order to remain viable. A musician cannot depend on the sales of music records anymore, but must find other ways to earn revenues from his or her creations.

6.5.2 Translation failures

Informal arrangements can be translated into official rules. For example, when writing the bylaws of a homeowners association the rules are often based on informal practices. Over time the reasons why some formal rules are written in the books may be lost, which may lead to problems.

For example, when a simple majority rule is used to make decisions, one can push forward with important decisions when the minimum number required is reached, but this might not be best for the viability of the community. A slight majority means that almost as many members of the community oppose the decision.

Leaders who rely on minimal majorities for too many decisions may find themselves having to use coercion and/or corruption, rather than general agreement, to keep themselves in power.

6.5.3 Blueprint thinking

The fact that some rules work out well in one action situation does not mean that those same rules will work well for other situations. The rules and regulations for urban planning in Boston might not work well in Phoenix given the many differences in the biophysical and social context. Obviously, rules for tennis would not work for basketball. However, rules of the game of ice hockey may not work well for field hockey even though these sports share many similarities. Nonetheless, we see blueprint thinking frequently, especially when large organizations implement many projects, as is the case in development agencies. The World Bank may implement projects on community development in many places in the world, but requires each project to use the same blueprint policies to receive funding.

6.5.4 Corruption and rent-seeking

When individuals in power have the opportunity to allocate resources, there is always the possibility of corruption. In these situations, various actors may collaborate to harvest subsidies or large infrastructure investments. When a bridge needs to be built, will the money be spent to build it according to the specifications, or can the inspector be bought off to save on expensive, high-quality construction materials?

6.5.5 Lack of large-scale supportive institutions

Small-scale communities can be very effective in self-organizing and sustaining their shared resources but will eventually experience challenges in the long term if they do not have the support of larger-scale institutions. For example, when efforts are coordinated at a large scale, scientific information can be collected and analyzed with expertise and resources that are not possible at the small-scale (it is difficult to build a particle accelerator by yourself). Farmers, for example, may receive help from highly trained professionals on new technologies and methods. Agricultural extension agencies provide a supportive role in disseminating knowledge, information and experience that farmers would not have the time to gather as individuals working alone.

6.6 Critical reflections

Based on the analysis of many case studies of fisheries, forestry, irrigation systems, and other long-lasting social-ecological systems, design principles can be identified. Those design principles are not blueprints for design but are guiding principles to analyze institutional arrangements and help researchers and practitioners to ask appropriate questions to improve the governance of the commons as needed.

6.7 Make yourself think

1. What are examples of boundaries that are not clearly defined?

2. Neighborhood watch programs are an example of self-monitoring. What are the pros and cons of such programs? How about neighborhood Home Owners Associations (HOA's)? Have you ever heard someone say about the new house they bought "Well, the good thing is that there is NO HOA".

3. When economic times are tough, politicians often recommend centralization of government function to save money. Why is centralization of governance structures in cities not necessarily an effective way to save monetary resources?

6.8 References

Baggio, J. A., Barnett, A. J., Perez-Ibarra, I., Brady, U., Ratajczyk, E., Rollins, N.,Rubiños, C., Shin, H.C., Yu, D.J., Aggarwal, R., Anderies, J.M., and Janssen, M. A. (2016). Explaining success and failure in the commons: the configural nature of Ostrom's institutional design principles. *International Journal of the Commons*, 10(2), 417–439. DOI: http://doi.org/10.18352/ijc.634

Cox, M., Arnold, G., & Villamayor Tomás, S. (2010). A review of design principles for community-based natural resource management. *Ecology and Society, 15*(4), 38.

Ostrom, E. (1990). *Governing the commons: the evolution of institutions for collective action.* Cambridge University Press.

Ostrom, E. (2005). *Understanding institutional diversity.* Princeton, NJ: Princeton University Press.

PART III

PART 3: RESILIENCE AND ROBUSTNESS

CHAPTER 7

Feedbacks and Stability

Key Concepts

In this chapter we will:

- Be introduced to system dynamics, feedbacks and resilience
- Learn that systems can have different possible outcomes, and can tip from one stability domain to another
- Become aware that concepts like resilience and robustness are frequently used in sustainability studies

7.1 Flipping Lakes

In this chapter we introduce core concepts of systems thinking. Those are important since our framework to study coupled infrastructure systems is very much influenced by those concepts. When we want to study the governance of coupled infrastructure systems, those systems operate in an environment exposed to disturbances and surprises. The world we live in experiences a lot of variability, whether it is weather variation, infectious diseases spreading within your community, accidents and unexpected outcomes of sport events. Life happens, and if we think about governance, we have to take into account that there is a lot of variation in the world around us that could impact the way we best govern a system. Assuming the world is nice, orderly and predictable, is a recipe for disaster.

To explain the concepts introduced in this chapter, we start with an illustrative example of lakes. Lakes are islands on land. Lakes are a favorite study object for ecologists since they are relatively self-contained ecosystems in which a range of plant species, fish species, and biochemistry interact. There are many different types of lakes and one can study which attributes of lakes generate different patterns of species abundance. One of the key areas of study is the process of eutrophication. If a large quantity of nutrients enters a lake due to runoff from the heavy fertilizer use on nearby farms, the lake water can suddenly flip from being crystal clear to looking like pea soup. This is an example of an ecosystem that can exhibit different long-run patterns of species composition (or “states”): a clear lake with little algae and a green lake dominated by algae (Figure 7.1).

Figure 7.1: On the left is an example of a crystal clear blue lake, while on the right you see a eutrophic lake.

One ecological mystery that has interested limnologists for a long time is why lakes suddenly flip from a clear blue to pea soup. The study of resilience of ecosystems has provided some insights on why this happens. Lakes, especially shallow lakes, have tipping points related to the amount of nutrients they can process beyond which the lake flips. When a lake turns to pea soup, it is not only less attractive to swimmers, it also creates an environment with reduced biodiversity that favors weeds and limits the number of fish species. We know that we need to control nutrient use to avoid creating undesirable states in nearby lakes, but the use of fertilization for agriculture also has benefits. Understanding how to avoid flipping the clear lake system into pea soup is critical and can provide lessons for other types of problems. For example, we may want to avoid pushing the climate system toward dangerously rapid change, or causing coral reefs to flip from a healthy state with many fish species to one dominated by slimy weeds.

7.2 Introduction to Feedback

In this chapter we discuss systems thinking and the concept of resilience. **Systems** are composed of component parts that interact with each other. For example, a herd of cows consumes grass from a pasture. The component parts in this case are the herd of cows and the pasture. As cows consume grass the biomass from the pasture is reduced. On the other hand, the cows produce manure which fertilizes the pasture leading to an increase in biomass. This simple example illustrates how the component parts interact and cause each other to change. Another example is a person controlling the temperature of the water while in the shower (Figure 7.2). If the person wants to increase the temperature, she will increase the volume of hot water and/or reduce the volume of cold water. There might be some delay between when a faucet is adjusted and hot water actually begins coming out of the faucet. An impatient person might open the hot water faucet too much and burn herself. Getting the right temperature requires that the person adjusting the faucet reacts appropriately to the information gathered from the shower (water temperature) in order to adjust the controls (opening the hot water faucet) appropriately. The interaction between the person and the shower is a system based on feedback. The notion of feedback is illustrated in Figure 7.2. The two component parts of the system, which are often represented as “state variables,” are here represented by boxes and the interactions between the variables by arrows. You can imagine signals flowing in a circle. A temperature signal flows from the shower water to the person who translates it into a position signal for the faucet knob, and the cycle begins again. This cycling of signals is the reason for the term “feedback loop.”

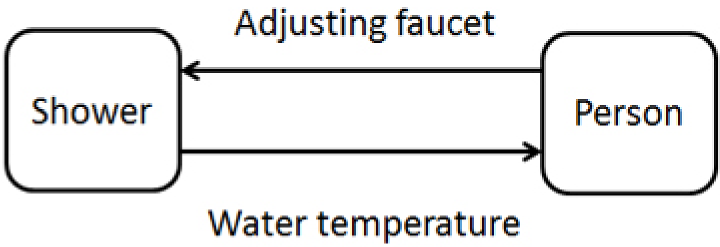

Figure 7.2: Hot water faucet adjustment from a systems perspective.

In the language of "state variables," a change in variable A causes a change in variable B which subsequently impacts variable A again. Feedbacks can be positive or negative (Figure 12.3). A positive feedback occurs when a change in one variable, after going through the feedback pathways, returns to induce an additional change in that variable in the same direction as the original change; if the additional change is in the opposite direction, we call that a negative feedback.

An example of a positive feedback is money on a savings account. As long as you don't take any money out of a savings account, the money will generate interest which is put back on your account and the next year you earn interest over a larger sum. Hence the interest you earn on interest from the previous year is an example of a positive feedback. An example of a negative feedback is the shower example. If the water is too hot, you reduce the amount of hot water, and if it is too cold you increase the amount of hot water in order to push the temperature toward your preferred value or **set point**. Here, "negative" refers to deviations away from the set point. The person generates negative feedback when they respond to the **positive deviation** by adjusting the faucet to **reduce the deviation**.

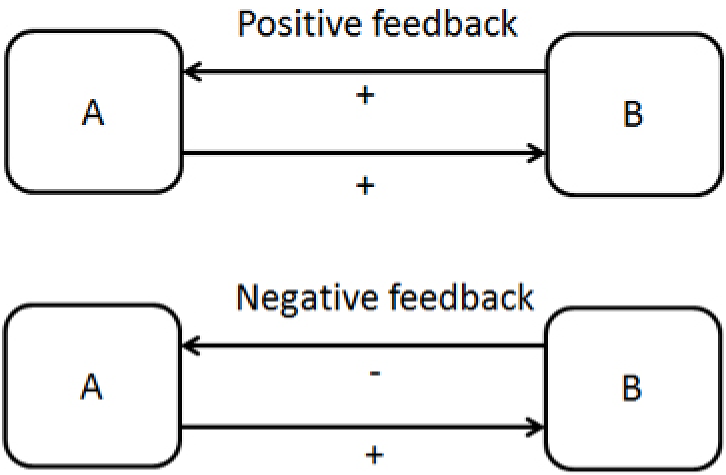

Figure 7.3: Examples of systems with positive and negative feedbacks.

If we look back at Figure 4.5 on the Institutional Analysis and Development Framework we can also consider the IAD a systems representation of governance problems. The outcomes of the action situation feed back into the contextual variables or directly into the action situation again. In this chapter we discuss the systems perspective more explicitly since there are a number of related concepts that are increasingly used in sustainability studies.

7.3 Resilience

Humans are part of a system of human-environmental interactions. Humans influence the rest of the system by appropriation of resources (i.e., removing system elements), pollution (i.e., adding system elements), landscape alterations (reconfiguring system elements), etc. There are characteristics of systems that help us to understand how they may evolve over the long term and how they are affected by these human activities.

Consider a young forest where the trees are small and there is sufficient light and nutrients for the trees to grow. The trees in the forest initially grow fast and the forest starts to mature over time. The growth rates of the trees slow down when the trees get bigger, block the sunlight for each other and compete for nutrients. Add to this idealistic description of tree growth the fact that forests cope with many types of disturbances such as pests, forest fires, and tornados and this situation becomes more complicated. If a forest experiences a fire why does it cause more damage in some forests than others? How does management of the forest influence the size of forest fires? And if forest fires are a natural phenomenon, shouldn't we allow them to burn freely? These are major challenges for agencies that manage forests. Since 1944, the U.S. has used an icon named Smokey the Bear to promote the suppression of fires. But it is one thing to try to prevent humans from starting fires, it is quite another to suppress all fires.

In fact, due to the suppression of forest fires, forests in many places in the U.S. have built up large fuel loads. This fuel consists of dead wood that is not removed by regular forest fires. This has consequences. When forest fires happen in forests where fuel has accumulated for decades, fires are intense and burn all trees, young and old, and even the soil. As a result, the forest will not recover. This is in stark contrast to forests that have frequent, smaller, lower intensity fires that regularly reduce the fuel load (hence, only small fires can burn). These types of files do not burn all the trees or harm the soil, therefore forests can easily recover. Hence too much suppression of forest fires can reduce the resilience of forests such that they cannot recover from the disturbance of an inevitable fire.

The concept of resilience in ecological systems was first introduced in 1973 by the Canadian ecologist C.S. "Buzz" Holling in order to describe the persistence of natural systems in the face of disturbances such as fires and pest outbreaks (Figure 7.4). A single system can have multiple types of states, for example a lake can be either clear or green and murky, a rangeland can have a mix of healthy tall grass and with a few trees and shrubs, or it can be covered with noxious weeds. If a system state is resilient, the system remains in that state even if it is exposed to disturbances. If a system state loses it's resilience, for example due to fire suppression, decades of nutrient loading in lakes or overgrazing a rangeland, it may not be able to recover from even a small disturbance, which would cause it to flip into a very undesirable state.

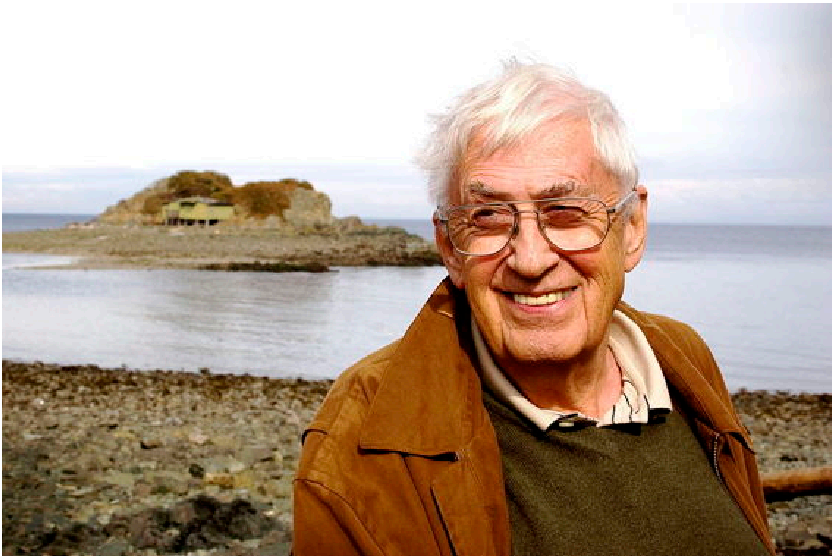

Figure 7.4: C.S. "Buzz" Holling.

The concept of resilience can be applied to many ecological systems. As discussed above, ecosystems often have multiple stable states. With the term stable state we refer to a certain configuration of the system—such as a healthy productive ecosystem with a lot of biodiversity—which can cope with variability such as rainfall, storms, droughts etc. An alternative stable state could be an eroded unproductive ecosystem. But there are limits to the size of the disturbances a system can cope with while in a particular stable state. If the system is in a desired stable state—such as the healthy productive state—we often want to keep it like that. Resilience can be defined as the magnitude of the largest disturbance (e.g., fire, storm, flood, nutrient shock) the system can absorb without transforming into a new state. The problem is that human activities can reduce the resilience of the system and make it more vulnerable to smaller and smaller disturbances such that it flips to another stable state (Figure 7.5). If a system is in an undesirable stable state we may want to restore the ecosystem but such states might be very difficult to get out from. Those undesirable states might be very resilient.

Figure 7.5: A representation of a system, the fish, that flips from one state to another state, the bowl.

Rangelands, found in arid areas all over the world, are another example of systems with multiple stable states. To illustrate the possible multiple stable states we focus on the example of

Australian rangelands. When Europeans came to Australia they started to use rangelands to raise sheep and cattle. Before European settlement, natural grazing pressure was the result of native herbivores such as kangaroos, wallabies, and wombats and was relatively low. When settlers added sheep and cattle to the system, the grazing pressure increased significantly. Moreover, European settlers installed watering points (simple troughs fed by water pipes) in the landscape to provide water for their sheep and cattle. Not only did these watering points benefit the sheep and cattle, they also benefited the kangaroo population which further increased grazing pressure.

In many areas of Australia, the properties (ranches) are very large and the density of grass is very low. As a result, the impacts of overgrazing may not be directly visible in the short run. Figure 7.6a shows an example of a healthy grazing area, which looks quite different compared to the green meadows in Europe. Nevertheless, the farmers made a good living out of this production strategy. But they not only increased the grazing pressure, they also suppressed fire. As a consequence, woody weeds started to blossom (Figure 7.6b) which outcompeted the grass and made the landscape useless for sheep farming. It will take decades before the woody weeds will disappear through a natural cycle, and it is too costly (given the size of the properties covering hundreds or even thousands of acres) to remove the weeds mechanically. As a consequence, a significant area of Australia's grasslands have now flipped from a grass-producing sheep-supporting landscape into a woody, weed-dominated wasteland.

Figure 7.6: On the left we have a typical Australian grassland (a). The landscape on the right shows a pasture dominated by the invasive species Scotch thistle in south eastern Australia (b).

7.4 Tipping points

In this section, we focus on a specific element of systems with multiple stable states, namely tipping points. How can we explain that systems suddenly change their behavior and change their configuration? Remember that systems consist of components that are connected via feedback. Those feedback relations can change. For example, grass can compete with woody shrubs in rangelands when grass has a particular density. When the ratio of grass to woody shrubs crosses a certain threshold, grass cannot compete anymore, and woody weeds take over. This threshold in the ratio of grass biomass to woody shrubs is the tipping point. Ecological systems are not the only systems that have tipping points. We can find tipping points in social systems as well (e.g., a peaceful protest flips to a violent riot). In his bestselling book, Malcolm Gladwell (2000) uses the example of the New York subway system. When there is a lot of graffiti on the subway vehicles and trash on the ground, people are more likely to be polluters. In the 1980s, the city government started cleaning the subways every day whenever they noticed new trash and graffiti. In a clean subway car, people are less likely to put new graffiti or throw trash on the ground or on the floor of the subway car. The feedback dynamics have changed, and cleaner subways cars stimulate cleaner behavior—an example of positive feedback.

How do we know we are near a tipping point? This is a critical question, but unfortunately we cannot answer it very well. A tipping point is something we cannot directly observe—it is hidden until we pass it. But if the only way to discover it is to pass it, and the point of discovering it is to **not** pass it, we are in a bit of a quandary. In fact, we only observe certain features of the system that are the indirect result of getting **near** a tipping point. One feature we can exploit is the fact that near a tipping point, somehow the system is “balanced” (see Figure 7.7). When the system is balanced, very small disturbances persist for a long time. Imagine a marble on a perfectly flat surface. The tiniest push will make the ball roll for a long time. Imagine, on the other hand, a bowl with steep sides, with the marble resting at the bottom of the bowl. If you give the marble a push away from the bottom, it will return rather quickly and stop moving. Thus, as systems approach tipping points, the rate at which they rebalance themselves after a small disturbance slows down. This is called “critical slowing down” and is one method (albeit an imperfect one) to discover whether a system is nearing a tipping point. Discovering when complex social and ecological systems approach tipping points is extremely difficult, and remains an active area of research.

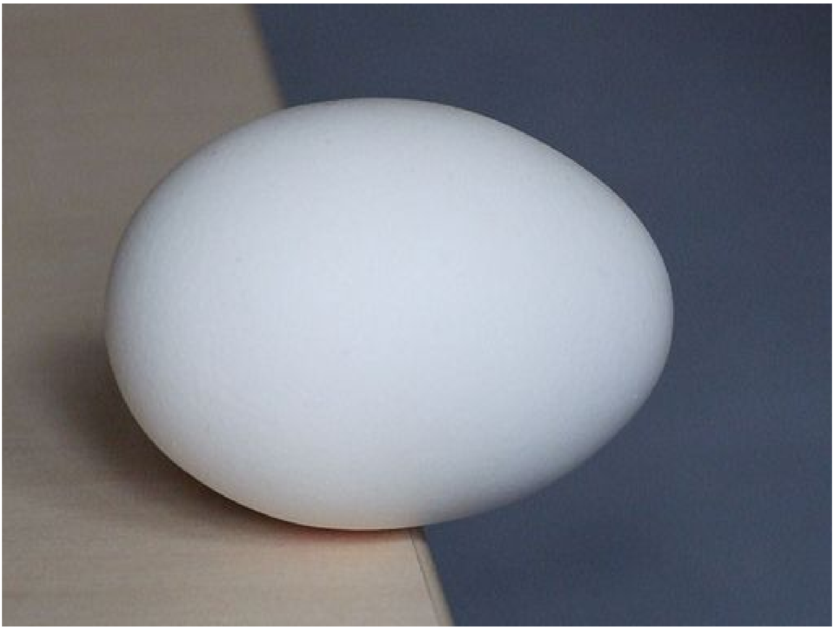

Figure 7.7: An egg balancing on an edge. How much disturbance will cause the egg to fall and experience a new stability domain?

Along with the increasing use of the term resilience in sustainability debates we also see the use of the term robustness. What is robustness and how does it differ from resilience? The concept of robustness comes from engineering and is used to design systems. For example, engineers develop control systems for airplanes such that the ability of the airplane to fly is "robust" to mechanical failures, turbulence, wind shear, etc. In their design of a robust airplane, engineers include various backup systems to avoid the situation in which the "system" (i.e., the plane and the people in it) flips from cruising at a constant altitude of 30,000 feet (10 km) to a free fall. But engineers also know that there are costs related to robustness and tradeoffs need to be made.

Do we build a wall 10 meters high around New Orleans to avoid damages from the next major hurricane? As we design systems to reduce our sensitivity to damages (i.e., to be robust) from disturbances, we have to make choices about what disturbances to consider. To be robust to one type of disturbance may create vulnerabilities to other types of disturbances. For example, utilizing concrete canals instead earthen canals for irrigation may reduce the number of wash outs, but may also reduce the ability of farmers to be adaptable and spatially reconfigure water flows using temporary mud walls in earthen canal systems to deliver water to the farmers' fields under variable circumstances.

Engineers argue that systems can be robust, yet fragile. They can become more robust to big fires but more vulnerable to small fires. Hence difficult choices have to be made.

7.5 Managing performance of systems

In our description of the IAD framework we discussed evaluation criteria associated with various outcomes. Interactions between the participants in the action arena lead to particular outcomes. Those outcomes are evaluated somehow. For example, did the participants achieve their goals? Are policy targets reached? Did the interactions lead to a fair allocation of resources? Based on the evaluation of the outcomes, the interactions in the action arena continue, and/or participants learn and change the rules-in-use. This chapter is about systems, and systems are about feedback and control. As such this chapter provides a more general perspective of the IAD framework.

Let's discuss the example of the IAD framework applied to taking a course (Figure 4.6). The interactions of the participants lead to a grade over the course of the semester. The specifications on how the grade is calculated are specified in the syllabus. When the professor and the students generate new grade information after an exam, they evaluate this information. This can lead to a continuation of the interactions in the action arena, but may also lead to a change in the attributes of the course participants (some may drop the course, or start studying more), or a change in the rules-in-use (the professor makes adjustments to the course material for the remainder of the course).

An example for natural resources is the use of groundwater. A city may use groundwater to provide its residents and industries with the water they need. The groundwater is replenished when it rains. If in the long term less water is extracted than is replenished, the groundwater level remains the same. However, a problem in many urban areas is that water demand is increasing rapidly, while the supply of water remains the same. As a result, the groundwater level will decline. If one measures the groundwater level on a regular basis one will observe this decline. How will the city government respond to this decline? At which level of groundwater decline will new policies be implemented (i.e., changes in the rules-in-use)? Will those policies focus on increasing supply or reducing demand? If the city will not respond in an adequate way, residents may revolt against water shortages or higher water prices, and may even leave the city (this may seem an unlikely scenario to many readers, but this actually occurred in the year 2000 in Cochabamba, Bolivia). Every response may generate new or expose existing hidden fragilities. For example, reducing water demand may cause problems with the pipes, such as solid waste building up when the flow rate of water through the pipes is reduced. Another example is that importing of water, such as bringing water from the Colorado river via canals over hundreds of miles through the desert to the city, makes a city vulnerable to changes in climate in other parts of the country.

Managing a dynamically changing system is difficult. We can control the temperature of our shower at home pretty well, but may burn ourselves if confronted with a different shower in a hotel during travel. What if lots of people are trying to adjust the faucet at the same time? As a thought experiment, consider a bunch of participants in the shower. First, a goal, the desired water temperature, has to be defined through a collective choice process.

Will all participants have a say, or are only certain participants in the action arena allowed to define the goals? Suppose there is a common goal, how will information about water temperature be used to adjust the faucet. Not all participants will receive the same feedback since not all participants can be under the showerhead. Do the people who adjust the faucet get reliable information from those who experience the hot water from the shower? This example shows the complexity of controlling a dynamic system when there are different participants who have different goals and positions. The institutional arrangements can enable or hinder the ability of groups to reach long term goals.

Earlier in this chapter, we mentioned that it is very difficult to know when a system is reaching a tipping point. We only know it for sure once we have passed it. How can we manage complex systems if we have incomplete information about the system? Scholars who have studied resilience and robustness of systems come to a number of insights that might be helpful for managing systems:

- Maintain diversity within the social and ecological components of the system. This includes biodiversity, but also institutional diversity. This diversity contains alternative solutions expressed in DNA or institutional arrangements. Avoid monocultures. In agriculture, a crop may be affected at a global scale if all seeds come from the same source and this particular variant becomes vulnerable to a pest. Likewise, we don’t want to have the same institutional arrangements in all jurisdictions. With institutional monocultures we cannot learn how others have addressed a similar problem in a different way.
- Maintain modularity of systems. Nobel Laureate Herbert Simon used the example of the watchmaker to illustrate the importance of modularity. Suppose a watch consists of 1000 parts. One approach to watch design is to assemble all 1000 components in one sitting. If the

watchmaker is disturbed or makes an error during the assembly process, she has to start again from scratch. Another design has modules and the watchmakers can assemble the modules, and then put the modules together. If a disturbance happens the watchmaker only needs to recreate one module. Modules also relate to the governance of social-ecological systems, and therefore we have states, counties and watersheds as units of governance in which new technologies and policies can be experimented with, without impacting the rest of the system.

- Finally, it is important to keep options open. Maintain redundancy, by which we mean that it is important to maintain some breathing space for the system. If everything is organized in a very efficient way, a disturbance could eradicate a keystone species, a charismatic leader, or the one source of revenue. It is important to have some fat in the system so that a disturbance can be absorbed. Have two operators of the energy distribution system so that the system can still continue if one of the operators is sick. Have multiple suppliers of energy so that a cloudy day reducing solar energy will not lead to a blackout of the energy system. .

Managing the performance of a system is very hard. It requires practice, continuous learning, and maintaining diversity, modularity and redundancy. In fact, there is a large, very well-developed technical field called control theory that focuses on how to use feedbacks to manage systems which we will take up in the next chapter. To close this chapter, we relate the concepts of resilience and robustness back to sustainability.

7.6 Resilience, robustness, and sustainability

In recent years the concepts of resilience and robustness have been increasingly used in the debate about sustainability. How do they relate to each other? Sustainability refers to a goal one aims to achieve. Sustainability guides the discourse on the interaction between human societies and the environment. There are many dimensions of sustainability, varying from avoidance of depletion of natural resources, avoidance of inequality and stimulation of quality of life for everyone and striving for a just society. Resilience and robustness ideas can be used to define system properties that may help decision-makers to achieve sustainability. Robustness focuses on feedback systems with clearly defined boundaries. Robustness comes from engineering and robust-control systems. It can be used to address questions about how to control a system to reach a target, such as sustainability? Robustness enables us to think about decision making, which information to use, how fast to respond to changes, and to think about trade-offs in decisions to be robust to certain shocks but not to others.

Resilience provides a framework to think about how multiple systems, each operating at their characteristic temporal and spatial scales, interact across scales. Human decision making can affect the resilience of a system by changing the shape of a particular stability domain. This can be intentional with a goal, for example, catalyzing a transformation of a fossil fuel economy towards a solar powered economy. Hence, resilience of a system in a particular stability domain is not always desirable and human activities can shape the long term dynamics of the system.

7.7 Critical reflections

With a systems perspective we consider the components of the system and their dynamic interactions. The IAD framework we discuss in this book is a systems perspective of human behavior, institutions, and the environment in which they are embedded. Systems also have characteristics such as resilience and tipping points, which we can observe in social as well as ecological systems.

7.8 Make yourself think

1. The next time you take a shower, reflect on your ability to control the temperature.

7.9 References

Anderies, J. M., Folke, C., Walker, B., & Ostrom, E. (2013). Aligning key concepts for global change policy: robustness, resilience, and sustainability. *Ecology and Society, 18*(2), 8.

Gladwell, M. (2006). *The tipping point: how little things can make a big difference.* Little, Brown.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 1–23.

Simon, H. A. (1996). *The sciences of the artificial.* MIT press.

Walker, B. & Salt, D. (2012). *Resilience thinking: sustaining ecosystems and people in a changing world.* Island Press.

CHAPTER 8

Robust control

Key Concepts

In this chapter we will:

- Introduce the concept of signals and systems
- Explore a short history of control theory in sustainability
- Map a standard control system and discuss the key concepts of stability and robustness in designed feedback systems.

8.1 Introduction

The word 'control' has the connotation that we act on a system to determine an outcome. That is, you are in 'control' of your life. You can 'control' focus on designed systems.This is anthropocentric. No need to think of 'control' with its implied notions of 'agency'. Why are you trying to 'control' something? Because someone or something else is controlling you (hunger, fear, shame, etc.). So let's forget about control, and think about what it really is: the process of taking some information from a system, using it to make a decision, then acting based on that decision.

This basic idea is captured in Figure 8.1. Information flows around the loop. Somehow the system that is being controlled (the 'dynamic system') such as your car, your body, a machine, an ecosystem, and economy, whatever. This dynamic system is referred to as such as it isn't very interesting to act on a static system. The dynamic system is evaluated by a 'sensor system', i.e. is being measured in some way. Your stomach has a sensor system that creates a signal that you interpret as 'hungry'. This hungry signal enters your 'evaluation/action system' and you take action, i.e. go get some food that then feeds back (alters) the dynamic system in some way, i.e. some biomass moves from an ecosystem into your stomach. And this loop functions continuously to keep you in an 'alive' state. This 'alive' state is an extremely complex state captured by the organization of molecules in your body.

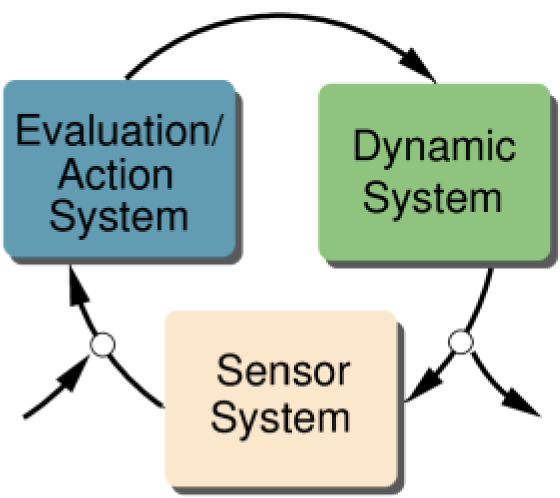

Figure 8.1: Sensor-Decision-Action loop.

These simple examples illustrate a much deeper point: all persistent structures, any observable pattern, are created and maintained by networks of such simple feedback structures shown in Figure 8.1, depicted in Figure 8.2.

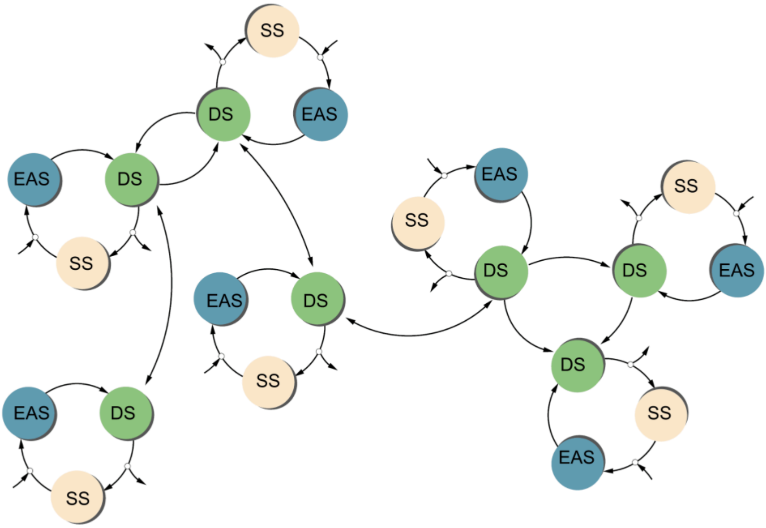

Figure 8.2: Network of regulatory feedbacks. All complex , persistent structures are made up of such networks.

In your body, there is a feedback loop to keep your temperature, your blood sugar level, your balance, etc. within some reasonable range. Together, they produce you as a persistent structure. When one of these loops goes wrong, it can cascade through the whole system and cause catastrophic failure. The idea of "governance" or "policy" or "management" is really about making the rules (institutions) and creating the hardware (organizations, monitoring systems, etc.) to implement those rules.

8.2 Control theory basics

In control theory, there are always four basic components:

1. A goal, e.g. produce a certain flow of resources such as in

a sustainable fishery, maintain a system state such as the atmospheric carbon dioxide concentration,

2. A controller, e.g. a mechanism that takes information about the system, compares it with the goal and takes action accordingly,
3. The system being controlled, e.g. an ecosystem, a factory, a car, an airplane, the earth system.
4. A sensor, e.g. a way to measure the system state, e.g. what is the fish biomass in a fishery, what is the water level in an aquifer, what is the carbon dioxide concentration in the atmosphere, what is the phosphorus level in a water body, or system flows, e.g. how many fish are being caught, how much water is being extracted, and how much carbon is being emitted into the atmosphere per unit time.

Figure 8.3 illustrates the "block diagram" depiction of control systems that are common in the controls literature. For a generic control system, the blocks are typically called the 'controller block', the 'plant block' (think of a manufacturing or a chemical plant), and the sensor block. It doesn't matter what the blocks are as long as the blocks do what is listed as above. That is, control theory and the block diagram can be applied to **any system**, from a cell, to your body, to a spaceship.

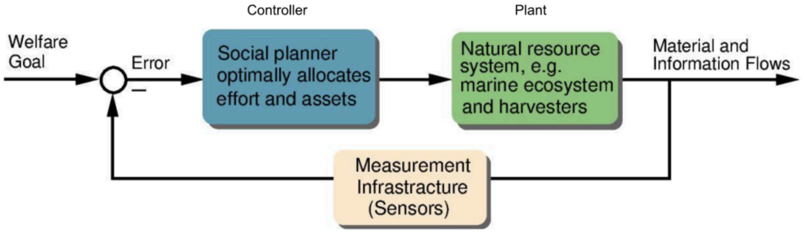

Figure 8.3: Block diagram for a control system in which we have adapted the blocks for a sustainability problem.

The absolutely essential feature of a control system is that information flows in a loop! Let's walk through the diagram. First is some information about a goal that *flows into* the control system. That goal information is compared to information about the recent state of the system. This is the arrow from the sensor that flows into the circle. Usually the 'comparison' is the difference between the actual state of the system and the goal state and this difference is the 'error'. Imagine you set your cruise control in your car at 60. Some sensor in your engine, motor, or wheels senses the speed. Is it 60? Let's say it is 58. The error = goal – actual = 60 – 58 = 2. So an error of 2 is fed into the controller which has a rule (a policy) that says 'if the error is positive, accelerate'. Now the signal "accelerate" is sent from the controller block to the plant block. What is the 'plant' in this case? The plant is the car and the landscape (the weight of the car and the landscape it is on will determine how it will move naturally without any force from the engine or brakes) and the car's engine and brakes which provide the control to change the velocity of the car (note that the controls have to match the goal – our goal is velocity and we must be able to change the variable related to the goal).

Let's now suppose the next reading of the speed is 60 (the plant has done its job to accelerate the car). Then the error is 60-60=0. Zero is fed into the controller. There is a rule in the controller that says "if the error is zero (we have achieved our goal), do nothing, i.e.

neither accelerate nor decelerate. Next, suppose as you are driving, you come to a downhill section in the road. Your car naturally wants to accelerate because of gravity – an intrinsic feature of the 'plant'. Now the sensor detects a speed of 65. Now the error is 60-65=-5. The error of -5 is sent to the controller which has a rule like 'if the error is negative, decelerate. The signal 'decelerate' is then sent to the plant which reduces the throttle to the engine and/or applies the brakes and the car slows down. In this way, an error detection, error correction loop (EDECL) can maintain the goal speed. There are many subtle problems with such EDECLs that must be solved and the methods for doing so are the core off control theory.

In what follows, we will use the 'car speed control' example. While it may sound simplistic to attempt to relate the control of complex systems to cruise control in a car, the problems are, in principle, the same. Control is fundamentally about keeping some system in a desired state by speeding it up or slowing it down, i.e. managing a rate or a collection of rates. Think of a fishery. We manage it by speeding up or slowing down the harvest rate. We manage the climate by slowing down the carbon emission rate or slowing it down to negative rates, i.e. removing carbon dioxide from the atmosphere. We manage aquifers by speeding up or slowing down the extraction rate. Groundwater recharge is, of course, slowing the extraction rates to negative levels, i.e. negative extraction is recharge. Thus, in the end, the relatively simple problem of controlling the speed of a vehicle in one direction has the essential features of much more complex control problems. We now explore key challenges and insights from control theory using this example.

8.2.1 System stability

One key problem feedback control must address is stability. That is, can the system hold its state close to the goal? In the car example, if we adjust the control based on the instantaneous error, we cannot reach the goal. Why? Because when the error is zero, the controller doesn’t do anything. Thus, only when the speed moves away from the goal does the controller react. This means that such a controller can only keep the speed within a range of the goal. That is, it will keep accelerating and decelerating as the speed moves up and down. The same is true for your thermostat in your house. You may have noticed that your heater or air conditioner cycles on and off according to what is called the ‘residual’, i.e. the error. Most heating systems have a residual of 2 degrees. So if you set the thermostat at 70 degrees in the summer, the air conditioner will turn on at 72, cool your house to 68 and then turn off. The temperature will then naturally rise to 72 when the thermostat will turn back on. So the temperature in your house will fluctuate between 68 and 72, but will never actually settle at 70. In the case of a home climate management system, this fluctuation is a fundamental property of the heating system: you can only turn a heat pump on and off. You can’t run it at any speed, it only has a limited number of speeds. With the car example, it is due to other factors as we can run the engine at any speed we want and we can brake to any speed we want. Rather, fluctuations are due to something engineers call *bandwidth* – i.e. how fast the controller and plant can respond.

The bandwidth issue is manifest in physical systems due to **inertia**. Inertia, in simple terms, is the fact that the energy required to change the velocity of an object depends on its mass and velocity. Quickly changing the velocity of a moving object can require an enormous amount of energy. Thus, the ability of the plant to produce energy fundamentally limits the bandwidth, or speed of response of any system. In the car, this is the horsepower of the engine – i.e. the common advertisement that a car can go from 0-60 mph in so many seconds, and the braking distance (or time). So, no matter how good the controller and sensors are, our capacity to keep a car at a particular speed is limited. On a flat road with no wind, with good controller design, we can keep a car extremely close to the set speed. But we know we can’t just rely on real-time information or we will chase our tail – think of trying to get your shower temperature just right based on how the water temperature feels now. In control theory, the most common type of controller, by far, is called a PID controller. P stands for proportional – you adjust your responses in proportion to the **present** error. I stands for integral – you adjust your response based on the sum of **past** errors. Finally, D stands for differential – you adjust your response based on your estimate of how the error is changing – usually the present error minus the previous time step error divided by the time step. This is a measure of **future** error. Most control strategies are then a weighted sum of P, I, and D strategies. In a perfect world, feedback control loops can be constructed to achieve any goal within the biophysical limits of the plant and limits to how fast the controller can process information. For an example of just how powerful feedback control is in ideal conditions, see this youtube video. Unfortunately, we do not live in a perfect world.

8.2.2 System robustness

Figure 8.3 depicts a control system for ideal conditions. In the imperfect world we live in, we often don't know how the plant works, the system is exposed to exogenous shocks, we can't get perfect measurements, and we can't process information arbitrarily quickly. These factors introduce mistakes and delays into the feedback system which can wreak havoc. Figure 8.4 shows the block diagram for real-world feedback control systems that we face in the real world. First, we can't even agree on a goal. Do human carbon emissions really contribute to climate change and do we limit carbon emissions or not? What should we limit them to? Are the scientists being overly reactive? Are we really overexploiting global fisheries? Most scientists say yes, others are not so sure. What level of harvest is sustainable? Is present economic inequality too high? Is high inequality good (some say it motivates entrepreneurship) or is it bad (others say it stifles economic growth). What level of inequality is good? What gini coefficient should society shoot for? Should 20 percent of the people be allowed to own 80 percent of the world's economic assets? And the list goes on and on……

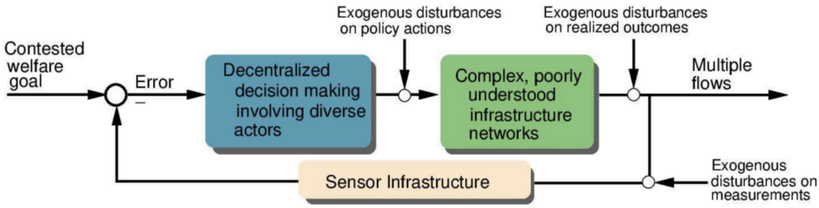

Figure 8.4: Real-world feedback system encountered in all sustainability problems.

Now, let's use the car driving example to work our way around the control loop. First, suppose that there is not one, but 5 people in the car who somehow elect a driver. The elected driver may have their own goal about where to go and how fast, but this may not represent the other passengers' goals. Next, suppose that somehow the group agreed on where to go, and how fast (to the movie at 45 mph). Now suppose after the journey begins, some in the group think that the car is going too fast. They have received some sensor information they believe indicates the car is not going 45. One yells 'slow down', the other 'speed up'. The driver says "but the speedometer is indicating 45", a third says "I don't believe the speedometer". Resolving this disagreement takes time, and introduces a delay between when information is received and when it is acted on, e.g. we are emitting too much carbon dioxide and should have acted 20 years ago. Now it may be too late and, if not too late, much more expensive. Now consider a situation in which the car randomly slows down when the throttle is pushed. Sometimes it randomly turns left when the steering wheel is turned to the right. Sometimes the brakes randomly don't work. These are examples of uncertainty in the plant. Suppose that the driver attempts to turn left but the passenger grabs the wheel and pulls it to the right. This is an example of an exogenous disturbance on policy actions. Suppose that an extremely strong headwind buffets the car and effectively slows it down as it tries to accelerate. This is an example of an exogenous disturbance on realized outcomes – yes the car is accelerating as per the drivers signal from the throttle, but it can’t do anything about the wind. Finally, suppose that mud is splattered on the windshield by a passing truck and the speedometer stops working correctly. These are exogenous disturbances on measurements.

While the scenarios above may seem far-fetched, many industrial systems must reliably operate in very challenging conditions. The field of **robust control** focuses on designing feedback control systems for such conditions. In the discussion of PID controllers above, the details are far beyond the scope of this book and we just focused on the main principles of carefully weighing how we use present, past, and future information to achieve system stability. Robust control system design is even more challenging and the details are, again, far beyond the scope of our discussion here. We just want to leave you with to main principles from robust control:

1. With unlimited bandwidth (essentially the capacity to respond instantaneously), it is possible to keep a system operating within an arbitrarily tight interval given limitations on the **observability** and **controllability** of the system. That is, we can build extremely robust systems, but this can be costly.
2. Our ability to increase robustness is limited by the intrinsic dynamics of the system. Just like there is a law in physics that energy is conserved, there is a law in robust control that **robustness is conserved**. That is we can convert energy into different forms, i.e. kinetic or potential energy, but the total amount of energy is fixed. Similarly, we can spread robustness across different kinds

> of shocks, but the total ‘amount’ of robustness is fixed. In practical terms, we can direct robustness capacity toward reducing the sensitivity of the system to high frequency variations, but then the system will become vulnerable (very sensitive) to low frequency variations.

So, in summary, we have illustrated the main ideas from control theory: balancing the use of present, past, and future information to achieve stability and using rapidly responding feedbacks to counter certain kinds of shocks to the system knowing that this will cost in terms of weakening the capacity of the system to deal with different kinds of shocks. For an excellent example of the application of these ideas to managing the COVID-19 pandemic, see this youtube video. What do these ideas mean for sustainability more broadly?

8.3 Control theory, governance, and sustainability science

Buckminster Fuller was a futurist and inventor who popularized the term ‘spaceship earth’. Kenneth Boulding then wrote an essay titled “The Economics of the Coming Spaceship Earth” in 1966. Well, at the time of writing of this book some 60 years later, the coming spaceship earth has arrived. If we think of Earth as a spaceship (and it is easy to make this analogy as we do below), then we can see how the example of keeping a car at a certain speed is quite general. The only difference is that when we pilot spaceship earth, the goal is not keep a certain speed but, rather, keep other variables within a ‘safe operating space’ while maintaining a certain level of human welfare. If one considers one aspect of the choice of speed of a car one of a safe speed, then the cruise control system of your car keeps the speed in a safe operating space while maintaining a certain level of the driver’s and passenger’s (the society in the context of transportation in a car) welfare measured in terms of the subjective feelings of the passengers about how long their journey will take. So, again, we see the similarity between the relatively simple problem of controlling speed to that of sustainability, or in other words, controlling spaceship earth.

With humanity's departure from the Holocene wherein humans predominantly adapted to global dynamics and entrance into the Anthropocene wherein we increasingly control them, we have transitioned from being passengers on to piloting `Spaceship Earth'. As pilots, we become responsible for the life support systems on our `Earth-class' spaceship. Until very recently Spaceship Earth has been running on autopilot regulated by global feedback processes that emerged over time through the interplay between climatic, geo-physical, and biological processes. These processes must necessarily have the capacity to function in spite of variability and natural disturbances and thus have developed some level of resilience in the classic sense of Holling–the ability to absorb and recover from perturbations while maintaining systemic features.

The capacity of these regulatory feedback networks to provide system resilience is limited. The Planetary Boundaries framework makes these limitations explicit and defines, in principle, a safe operating envelope for Spaceship Earth. The pilots face two challenges: knowing the location of the `default' operating envelope boundaries and understanding how these boundaries change with changing operating conditions. Pilots typically have an operating manual that provides this information which enables them to better utilize the resilience of their vessel (Earth System resilience). Because we don't have that luxury, the concept of resilience becomes critical: the art of maintaining life support systems under high levels of uncertainty–flying our Earth-class spaceship without an operating manual. Earth System science is, at its core, the enterprise of uncovering the operating manual. Unfortunately, our capacity to experiment is quite limited: the time required is enormous and the number of independent copies of Earth is limited. We can't test resilience by transgressing global thresholds and observing how the system behaves in new states and potentially recovers, e.g. `snow-ball Earth' and the `tropical states' of Earth's past.

With a manual, a crew, and a captain, the `resilience' question would boil down to the competence and risk aversion of the captain and the competence of the crew. This resilience would encompass

- piloting the spaceship so as to avoid shocks (e.g. avoid asteroids, maintain safe speeds, etc.),
- developing knowledge/skills to quickly repair existing systems if shocks can't be avoided,
- the capacity to improvise and create new systems when existing systems can't be repaired,
- and conducting routine maintenance so as not to destroy the ship through usage.

The first two elements constitute a key element of robustness, or specified resilience, described above. The third element (general resilience) is much more difficult to invest in. It requires the development of generalized knowledge and process to cope with rare and difficult-to-predict events. The fourth element has been the focus of most environmental policy thus far with a tendency to do just enough to get by.

Now consider our Earth-class spaceship. There is no captain or crew (there is no driver of the car – we all know the problem of the backseat driver). Subgroups of passengers are restricted to certain areas (e.g. the upper or lower decks). Some groups have access to more ship amenities and those with less access often support the production of amenities for those with more. As with the Titanic, the impact of shocks is very different for first- and second-class passengers. There is minimal maintenance of life support systems and, in particular, the waste management subsystem. Complaints about life support systems are met with agreements that it should be fixed but disagreements about who should pay. Worse yet, there is no operating manual so no one knows how to effect repairs, or their costs. No one knows how to set the cruise control, and no one knows how to fix the engine or brakes if they fail.

To an outside observer, our situation might seem absurd – like a crazy group of people in a car arguing about where to go and how fast to go there while the car is careening toward a cliff. If given the choice, many rational passengers on board would disembark. While some extremely wealthy passengers seem to be making an attempt, disembarkation is not realistic. So what do the passengers do? One key difference between the spaceship metaphor and our journey is that there is no destination. Further, the ship's journey is much longer than our lives so the ship becomes our home. `Good piloting' is tantamount to effectively managing life support systems–the `Earth System' (ES)–while ensuring the wellbeing of and preventing critical conflict among the passengers–managing the `World System' (WS). And because we must do this without an operating manual, we must build World–Earth System (WES) resilience (WER). We must be able to define WER to characterize how close these critical systems are to breaking down and model it to explore mechanisms that enhance or degrade WER.

So, based on what we know about control theory, what considerations should be built into our policies?

8.4 Critical reflections

Robust control is a concept from engineering, but can be applied to other types of systems too. Suppose we wish to navigate the system towards a certain goal (e.g. 60 miles/hour, BMI < 25, net zero carbon emissions), we need to take actions to reach that goal. Due to imperfect understanding of the system and external events, we receive information that indicates we have to adjust our actions to reach our goal.

A robust control perspective can be applied to governance, explicitly acknowledging that we do not have perfect knowledge, and have to create error-detection-error-correction type policies so we keep on track of our collective goals. Goals can only be in reach if we have the bandwidth to make adjustments to the “speed” at which the system changes. We also have to make tradeoffs regarding which kind of disturbances to prepare for, since there is only limited bandwidth available. In Chapter 14 we will discuss the bandwidth problem for reaching net zero carbon emissions by 2050.

8.5 Make yourself think

1. Do you have any ‘robust’ policies in your life? I.e. setting your alarm clock 30 minutes early to make sure you get up in time (your body has ‘sleep inertia’, right?)
2. Try to think of one critical feedback system in your body that regulates some important quantity. There are many to choose from!

PART IV

PART 4: COUPLED INFRASTRUCTURE SYSTEMS FRAMEWORK

CHAPTER 9

Coupled Infrastructure Systems

Key Concepts

In this chapter we will:

- See how to combine systems concepts with the IAD framework
- Explore how different type of infrastructure (natural, human-made, soft and hard) are part of a general framework to study human-environment interactions

9.1 Introduction

In this chapter, we will discuss an extension of the IAD framework that includes some of the insights from systems science. Furthermore, it incorporates our long-term experience with studying current and historical irrigation systems around the world. We view irrigation systems as a model for many problems societies experience just as the fruit fly is used as a basic model in genetics. By this, we mean that studying collective action problems in irrigation systems will teach us a lot about solving collective action problems in many other societal contexts. We will expand on this in the next chapter. Moreover, we argue that many societal problems can be studied from a coupled infrastructure systems perspective.

In Chapter 2 we discussed different types of infrastructure such as hard infrastructure (human-made brick and mortar infrastructure), soft infrastructure (human-made "software" to use other types of infrastructure), natural infrastructure (hard infrastructure that is not man-made), human infrastructure (knowledge and skills) and social infrastructure (social relationships). In the next sections we discuss how those different types of infrastructure relate to and depend on each other.

9.2 Collective action and infrastructure

There are a number of collective action problems related to the creation, maintenance, and use of infrastructure. One key problem is the question of who is going to pay for the creation of the infrastructure or, put another way, how will the cost be shared? Farmers in rural Nepal may pay by providing labor to the construction and maintenance of the irrigation system. In many western societies we have governments that collect taxes and use the resulting revenue to pay professionals for the creation of infrastructure.

The regulations regarding who pays can lead to perverse effects. For example, in the U.S. the federal government pitches in to pay the majority of the costs for highways, while the local governments pay less than half the costs. As a consequence, local governments are often eager to increase the number of highways to promote economic development. The subsidizing of road expansion was instituted after World War II to improve transportation and accessibility and, indirectly, economic development. A consequence of this policy is that roads are cheap for users and this enhances the demand for using them. One reason for the low price for using roads is the relative lack of maintenance, which is paid for by local governments. So, local governments want to improve economic performance to send signals to voters to get re-elected, but forget that roads must be maintained. This activity may generate short term benefits, but generates a long-term burden for taxpayers. One way to generate more funds for road maintenance is through the "user pays principle" and to raise taxes on gasoline (an indirect way to charge drivers to use the roads). Of course, this is a politically challenging proposition for elected officials who cannot afford to be seen raising taxes. It also creates problems when the vehicle fleet is becoming more fuel efficient. Higher fuel efficiency means that road use can increase while the tax revenues to maintain the roads goes down. This is especially true of electric vehicles who do not pay for gasoline and thus do not pay a user fee. As the proportion of electric vehicles grows, governments will have to find other ways to finance maintenance than through gasoline taxes.

A second collective action problem for infrastructure is to define who gets access to its use. In many irrigation systems there is a natural asymmetry between the upstream and downstream users of the canal system. Farmers have to solve the collective action problem of how to deal with this asymmetry. They may create a rotation system to reduce the impact of asymmetry. When taxpayers contribute equally to the creation of the infrastructure, this does not mean that they all have access. Taxpayers contribute to higher education, but not everyone who pays taxes has access to the knowledge infrastructure that is created. There are criteria for students to be admitted. The reason that societies invest in higher education is that everyone benefits indirectly by having a highly educated population (such as physicians, engineers and lawyers).

We have taught in different countries and have experienced the impact of how higher education is organized. In many countries, such as European countries, higher education is subsidized. Admission criteria are more challenging than some other countries, such as the USA, where public universities cover only a very small percentage of their expenditures by tax money. Students in the

USA pay high tuition to enjoy higher education. This impacts the way education is implemented. If education is subsidized, there is an incentive for universities to be selective about who enters and progresses through the educational programs. When students pay high tuition, students are clients, and there is an incentive to accommodate them to be successful in the program. In subsidized higher education, governmental criteria determine how money is spent and what educational programs may get more students. When students pay huge tuition fees, the expected income from getting degrees will drive which programs are popular. Both systems may, in theory, lead to similar outcomes in which people are educated for jobs that need higher education. However, we may expect that there are significant differences in how the benefits of higher education are distributed across different groups in society.

9.3 Coupled infrastructure systems

In this section, we present an extension of the IAD framework that includes some of the specific problems related to interacting infrastructures. The framework was developed by Anderies, Janssen, and Ostrom in the early 2000s to facilitate their study of irrigation systems. While the notion of infrastructures creating action arenas are implicit in the IAD framework, the details are not explicit. The intent of the framework was to make explicit how the different types of infrastructure discussed above come together to structure the "action arena" and, more broadly, networks of action arenas that constitute real-world systems in which humans interact with shared built and natural infrastructures. To introduce the framework, we will focus first on a shared resource that is used by a number of resource users. This has been the canonical view for many of our small communities who interact with their resources. However, this raises the question of who creates the rules of how the shared resource will be used.

In this framework, we explicitly include two levels of action arenas, namely the operational level and the collective choice level. The third component of the framework consists of the public infrastructure providers, who are the ones who create the rules for the resource users. In small communities all resource users might come together on a regular basis in the evening to discuss the challenges in governing their shared resource. There might be a chair, a treasurer, and some other roles within the group of public infrastructure providers, but they are all resource users and thus have a stake in creating rules to improve the performance of using the shared resource.

In larger systems, individuals represent other resource users, typically in committees that deal with provisioning of the public infrastructure. These representatives are often selected through a collective choice arrangement, such as elections. Decisions by the committee about what types and how much public infrastructure is provided are also made using agreed-upon collective choice arrangements (e.g., Robert's Rules of Order). This could be the state forestry committee that makes decisions on how to cope (i.e., how to allocate resources) with invasive species and who set rules on property tax benefits for landowners who plant new trees. The general assembly of the United Nations is a more extreme example where each member, a nation, is represented by an ambassador in making policies at the international level.

In studies where the Coupled Infrastructure System framework has been used, a common finding is that the link between resource users (mainly human and social infrastructures) and public infrastructure providers (mainly human, social, and soft human-made infrastructures) is a critical one. The bigger the distance, the less the practical knowledge (a particular type of human infrastructure) from resource users is included in creating institutional arrangements. A lack of practical knowledge may lead to policies that do not fit the reality that resource users experience and therefore policies may not be effective. On the other hand, local communities may not have the specialized knowledge (a type of human infrastructure) needed to solve certain problems on their own, and therefore creating institutional arrangements where representatives of many localities are involved can be beneficial.

The fourth component of the framework is the public infrastructure which includes mainly hard and soft human-made infrastructures. The public infrastructure providers may have decided on new institutional arrangements, but they may need a bureaucratic apparatus to implement and enforce those rules. Tax collectors, property inspectors, and guards all mobilize essential human infrastructure to implement the soft infrastructure of various types of coupled infrastructure systems. Canals, pipes, bridges, and satellites are part of the hard infrastructure of various types of coupled infrastructure systems.

The infrastructure could influence the resource directly, for example by improving the capacity of a landscape to capture water, or monitoring the state of the forest by remote sensing. The infrastructure can also interact directly with resource users, namely by assigning allowable actions (licensing), by monitoring the actions of resource users relative to allowable actions, or by providing information to users such as weather forecasts.

The framework (Figure 9.1) distinguishes four components, namely the shared resource system (natural infrastructure), resource users, public infrastructure providers and the public infrastructure. We can integrate this with the IAD framework such that the interaction between the four components constitutes a set of action arenas at the operational (resource users) and collective choice (public infrastructure providers) level. The external context defines the biophysical conditions of the shared resource system and the public infrastructure, the attributes of and the rules in use among the resource users, and the public infrastructure providers.

On short time scales, the interactions of the four components lead to outcomes. We are especially interested in how the interaction between resource users and public infrastructure providers leads to infrastructure that facilitates productive outcomes. On longer time scales, the interactions of the four components generate feedbacks that generate persistent patterns over time, for example, the inter- and intra-generationally fair use of shared natural infrastructures like the oceans and climate systems, i.e. features that underlie sustainable societies. Thus, we are also especially interested in the robustness of coupled infrastructure systems, building on the concepts introduced in the previous part of the book.

Inequality is an important component of the functioning of coupled infrastructure systems. We know from historical research and experimental studies that inequality may have negative consequences for the ability of groups and societies to solve collective action problems. For example, where do we place the nuclear power station, or which economic sectors will have to reduce their water use to avoid the major consequences of a drought? If we have to reduce our carbon footprint will the average Joe have to forgo their holidays (carbon emissions from travel) while rich households can buy additional carbon emission rights? What happens if rich neighborhoods get off the grid by powering their houses with solar energy and driving around in Teslas? Those who do not go off the grid will now have less capacity to maintain an already aging infrastructure.

What if an elite group in society is better represented in the category of public infrastructure providers compared to a lower income group? How will this affect the kind of policies that are developed, what types of public infrastructure are produced (i.e., defense versus health care and education or environmental protection) and how the fairness of those infrastructures are perceived?

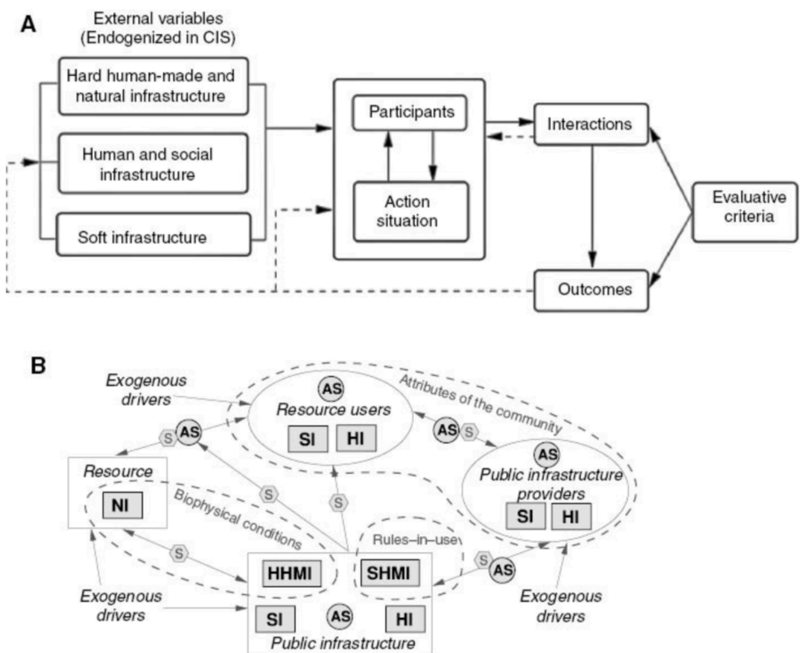

Figure 9.1: A. The IAD framework with different types of infrastructure as external variables. B.The coupled infrastructure system with IAD components overlaying the framework. The resource system and the public infrastructure are directly related to the biophysical conditions. The resource users and the public infrastructure providers are directly related to the attributes of the community. And the public infrastructure is directly related to the rules-in-use. AS = Action Situation, NI = Natural Infrastructure, SI = Soft Infrastructure, HI = Human Infrastructure, HHMI = hard human made infrastructure, SHMI = soft human made infrastructure, S = spillovers.

9.4 Use action arenas to study coupled infrastructure systems

The creation of the coupled infrastructure systems framework does not mean we abandon the IAD framework. On the contrary we argue that the IAD framework, and the action arenas in particular, are especially suitable useful to study coupled infrastructure systems. As mentioned in the previous section, there are questions on how participants make their decisions to invest in infrastructure and solve collective action problems. How to create incentive structures for elected officials (as public infrastructure providers) to facilitate the creation and maintenance of public infrastructure that leads to beneficial outcomes for the resource users? In other words, how to avoid rent seeking and corruption which is tempting since public infrastructure providers are often in a position of specific information and power in decision making and frequently out-of-office the time potential design flaws of public infrastructure can be observed. The answer depends on the specific case, but it could help to have public infrastructure providers who have a stake in the success of the outcome (members of the community vs outsiders), have transparency broad participation in the design and implementation, and design indicators measuring progress and process.

Another type of action situation is the appropriation of resource users of the resource. The action situation may help to identify how actions of the participants are monitored, whether monitors have an incentive to police strictly but compassionately, and whether resource users and monitors can monitor accurately the state of the resource. Will capacity building be needed (investments in human and social infrastructure)?

Basically, the coupled infrastructure systems framework is a more dynamic perspective of the IAD framework with multiple actions arenas. We have seen students after the introduction of infrastructure perspectives separate it from the IAD framework, but that is not the intension. Our perspective of infrastructure is very much about collective action and human decision making, not bricks and mortar. The examples we discuss in the coming chapters may hopefully contribute to immerse the different frameworks.

9.5 Robustness of coupled infrastructure systems

Coupled infrastructure systems experience many types of disturbances. For example, weather, insect outbreaks, wildfires, and earthquakes can impact the shared resource system as well as the hard public infrastructure. External changes imposed by higher levels of governance can impact the soft public infrastructure, resources users, and public infrastructure providers as well as cause changes in prices of inputs and outputs, infectious diseases and technological innovations. "Robustness" refers to the capacity of a particular coupled infrastructure system to cope with such shocks and continue to maintain persistent structures and patterns of organization that deliver benefit streams over time.

If a coupled infrastructure system is to be robust, a disturbance should not fundamentally disrupt the functionality of the system and the system should regain its basic performance relatively quickly. Earthquakes can cause major damage. In the early months of 2010 there were two major earthquakes. A magnitude 7.0 earthquake destroyed the hard and soft infrastructure of Haiti on January 12, 2010. Years after the earthquake many people still live in camps, some with only basic sanitation. The total death toll is not known but is believed to be around 200,000. In contrast, on February 27, 2010, Chile experienced an earthquake with magnitude 8.8, which is much stronger than what Haiti experienced. Yet the total number of fatalities was 497, mainly due to a tsunami caused by the quake. A year after the earthquake, most of the damage, including damage to roads and bridges, was repaired.

What might explain the difference in responses to the earthquakes in the different countries? First, note that Chile has very strict building guidelines to improve the ability of buildings to cope with earthquakes (i.e., to be robust to earthquakes). Since there are so many earthquakes in Chile, one has to build with the right materials and construction design. There is a general awareness across the population about the danger of earthquakes, and individuals, families, and organizations regularly practice what to do when there is a major earthquake, that is, they have invested in knowledge and emergency response protocols to increase robustness (what kind of infrastructure is this?). In the less economically developed Haiti, earthquakes are less frequent than in Chile. Thus, there is much less experience with major earthquake disasters. Because of this lack of experience (an element of human capital) there was no attention or resources allocated to mitigate the effects of potential earthquakes.

The weak soft infrastructure hindered the ability of Haiti's government to implement effective disaster risk-reduction measures which reduced the robustness of the hard infrastructure to earthquakes.

Chile has a more robust coupled infrastructure system to cope with earthquakes compared to Haiti. Due to the frequency of earthquakes and the occurrence of the largest magnitude earthquake ever measured (9.5) in 1960, the Chilean government created strict building guidelines to reduce the impact of future earthquakes. Since building robustness has a cost, one has to define priorities to guide how resources are allocated. It is not uncommon that after a major disaster, new regulations are put into place to reduce the impact of rare, major shocks, whether they are earthquakes, floods, or forest fires. Regardless how much and in what capacities governments invest, coupled infrastructure systems cannot be robust to every possible shock. Scholars who study these systems thus speak about systems being "robust yet fragile." A system can be designed to be robust to one type of shock but can, as a consequence, become vulnerable to other types of shocks. The simplest example is the sea wall that protects a community from annual storm surges but makes it more vulnerable to rare surges that happen once a century. At a more basic level, the resources used to build the sea wall cannot be used to invest resources in becoming robust to another type of shock.

These examples illustrate that the "robust yet fragile" (recall this feature of feedback control systems discussed in Chapter 8) nature of coupled infrastructure systems play out in multiple ways.

In recent years the U.S. has experienced major damage due to hurricanes such as hurricane Katrina (New Orleans), hurricane Sandy (New York City), hurricane Ian (Florida). Those hurricanes demonstrated the vulnerabilities of coupled infrastructures, especially due to flooding. Those vulnerabilities were well known in the scientific and engineering communities, but were not considered important enough for governments to act on. As mentioned before, being robust to specific threats requires priority setting.

With the anticipated climatic change over the next century, we expect more frequent and/or more intense hurricanes. As a result, vulnerable urban areas are now rethinking what it means to be robust. Does this require a different way to produce and distribute clean water, energy, and information? Do we continue to invest in cities which are in vulnerable areas, especially those impacted by the rise in sea level, such as New York City? Would it be best to abandon the types of coastal natural infrastructures that support iconic cities and stage a slow, directed resettlement? How many resources should be spent by all taxpayers to protect a small proportion of the population that lives near vulnerable coastal areas? These questions highlight the challenging choices and trade-offs public infrastructure providers must make as they allocate scarce resources to develop and maintain different types of infrastructure that constitute the coupled infrastructure systems upon which we all critically depend for almost all aspects of our welfare.

Figure 9.2: New Orleans after Katrina.

Another example of a consequence of lack of infrastructure investments is the power crisis in February 2021 in Texas. About 5 million people lost their power for several days while there was a major winter storm. This led to hundreds of people dying due to lack of health and other services that need power. The reason for the power outage was the lack of winterizing of power sources. The vulnerability of the Texas energy system to winter storms was known, and warned for by Federal agencies, but no investment in preparedness was made. The power crises cost about 200 $ billion. It is easy to blame the Texas government and energy companies after the event, but they have to make decisions about what to invest in for many types of potential threats. Whatever investments will be made to improvements of the robustness of the Texas power system, there will always be vulnerabilities.

9.6 Common features of infrastructure systems

When using infrastructure concepts for ecological, social and human attributes of systems, can we apply features we may associate with physical infrastructure to other types of infrastructure? We explore this question in this section. For physical infrastructure, one needs to construct the actual brick and mortar structures and, after completion, repair wear and tear to maintain the productivity of the infrastructure. Typically, infrastructure has a particular flow capacity, how many cars can cross the bridge per hour, how much water can flow through drainage pipes, how many bytes can flow down a cable, etc. We may distinguish between a base load, the common load during the normal operations of a system, and a maximum load, the maximum demand that the infrastructure can support. Using more than the base and peak load could lead to extra stress to the system, potentially leading to burst pipe, collapsing bridges and cracking surfaces of roads.

In the design of infrastructure, one needs to consider what the capacity constraints are, which impact what kind of robustness will be available. We can apply this to other types of infrastructure (Table 9.1). Soft infrastructure can be overrun with too many rules and regulations that do not fit the system. Creating too many rules, too much bureaucratic processes, may lead an organization, whether it is a community organization or a federal government to a stand still. When actors in action situations have to comply with the soft infrastructure, this requires time, and especially if those rules are not well understood or accepted, this may lead to mistakes and a lack of compliance.

With natural infrastructure, humans may make adjustments to improve the flow of resources. This may require continuous maintenance by mowing, trimming, cleaning up of the natural infrastructure. An agricultural field may increase productivity by using tilling, pesticides and artificial fertilizers, but too many crops will reduce the natural productivity of the system. Natural regeneration is needed and taking into account the stress put on the system.

Same with human infrastructure. As you are creating skills in human infrastructure to study coupled infrastructure systems, one may need to use it and apply the skills, in order to maintain those skills. Learning too many concepts at once may reduce retention of knowledge. People differ in their ability and practices to study, but in general cranking all your knowledge and skills into your brain and muscles just before an exam, performance or race might lead to bad outcomes (overworked, overtrained).

With social networks the focus is on connections. Each connection added to your social network leads to additional time to maintain connections. Although people may have hundreds of friends on Facebook, in practice people will spend their quality time to a small fraction of those connections. Not having the time to keep up with maintaining those connections could lead to distrust and a lack of reciprocity, when you need it.

Table 9.1: Applying infrastructure terminology across the types of infrastructure.

	hard	**soft**	**natural**	**human**	**social**
creation	construction	Drafting rules and regulations	Niche construction	Training, education	Networking, making connections
maintenance	Repair of wear and tear	Keeping rules on paper inline with rules in use	cleaning, trimming, watering	Using skills, Keep practicing	Stay in touch
Base load	Typical flow	Typical organization	Normal productivity	attention	attention
maximum load	Maximum capacity pipes	Maximum organization capacity	Maximum net primary productivity	Maximum hours of work	Maximum number of connections to maintain
Stress outcomes	Black out, flooding, pipe bursts	Stand still	Reduction of regeneration	Mental health	Lack of trust and reciprocity

9.7 Critical reflections

Infrastructure may be taken for granted, but it is critical to generate the services and resources we need for our daily lives. There are different ways we can organize the creation and maintenance of infrastructure, and the institutional arrangements (soft infrastructure) that affect the robustness of coupled infrastructure systems.

9.8 Make yourself think

1. Who paid for the creation and maintenance of roads you use to get to campus?
2. How is the electricity generated that you use at home?
3. Maintenance of roads is paid largely from gasoline tax. What are the consequences of more energy efficient and even electric vehicles for the continued maintenance of the road infrastructure?

9.9 References

Anderies, J., Janssen, M.A., & Ostrom, E. (2004). A framework to analyze the robustness of social-ecological systems from an institutional perspective. *Ecology and Society, 9*(1), 18.

Anderies, J. M., Janssen, M. A., & Schlager, E. (2016). Institutions and the performance of coupled infrastructure systems. *International Journal of the Commons*, 10(2), 495–516. DOI: http://doi.org/10.18352/ijc.651

LePatner, B. B. (2010). *Too Big to Fall: America's Failing Infrastructure and the Way Forward.* UPNE.

Stiglitz, J. (2012). *The Price of Inequality.* Penguin UK.

CHAPTER 10

Coupled Infrastructure Systems for Water

Key Concepts

In this chapter we will:

- Demonstrate the CIS framework applied to diverse water systems
- Learn that water systems only function properly if the various types of infrastructure are maintained.

10.1 Introduction

In this chapter, we discuss a number of case studies of irrigation, drainage, and sewage systems from a coupled infrastructure perspective. These cases provide a rich set of examples of how to apply the CIS framework. In all these cases humans created physical infrastructure to aggregate water in the landscape, store it, and transport it to canals and pipes, and dispose of the water at a designated location. This allows societies to regulate water availability and reduce the chances of floods and droughts.

The creation, maintenance, and use of this hard infrastructure requires collective action. The earliest records of irrigation that have been found date to about 8000 years ago in the Middle East (contemporary Iraq and Iran) and cases of independent development of irrigation systems are found all over the rest of the world. There is a wide variety of ways in which infrastructure is built ranging from mud-based canals and simple wooden diversion structures, to concrete canals and computer operated diversion gates.

Karl Wittfogel (1957) argued that the need to solve the many problems associated with complex irrigation systems is one of the sources of complex societies. In his view, large bureaucratic systems are needed to coordinate labor and supplies required to build, operate, and maintain large irrigation systems. These bureaucratic systems, he argues, were then later extended to the rest of society. Although this is an interesting and plausible argument, it rests on a basic assumption: top-down intervention is needed for the operation of large, complex irrigation systems. From many studies of irrigation systems in the last few decades, we see a different story emerging that suggests that this is not necessarily the case. We will discuss various cases in this chapter where decentralized bottom up governance leads to high performance of hydrological systems.

The governance of water requires the maintenance of various types of infrastructure. Key infrastructure types include natural infrastructure such as watersheds and the human-made infrastructure of dams, canals, pumps, and treatment plants. Water governance is arguably one of the most important issues we face today. We use water every day to drink, to cook, to shower, to wash, to flush, to clean, to irrigate, to swim in, etc. The importance of water makes it a central object of many social dilemmas: too little clean water, too much polluted water, unequal distributions of water, and variability in water availability.

One drop of water can be used multiple times for different uses, so technically water itself is not subtractable in the sense that the same drop of water used by a city resident to wash their dishes can be used later to water a golf course. This is exactly the same as the situation with a book in a physical library. One person can read a book now, then someone else can read it later. The key limit is that the resource cannot be used for two different purposes at the same time. So, if clean freshwater is being used for washing, flushing, drinking, etc., that same quality of water is not available for other activities at that time. Even though we can use creative coordination mechanisms and technology to recycle shower water to flush toilets, nevertheless, usable water is frequently discarded as wastewater to the sewer system or the natural environment. At that point, such water needs to be recollected and treated for reuse. Hence inefficient use of water by individuals leads to a shortage of clean fresh water for the group. Thus, water use is a social dilemma, or as social dilemmas associated with common-pool resources are sometimes called a "commons dilemma."

10.2 Phoenix, Arizona, USA

The city of Phoenix was founded in 1861. Since then, the population in the Phoenix metropolitan area has grown to about 4.8 million people (2020). The city has a subtropical desert climate with temperatures up to 118° F (48° C) in the summer, and average annual precipitation of 8.3 inches (210 mm). As places to live go, Phoenix is a harsh climate to be sure (Figure 10.1). This raises the question of why do so many people live in the desert? Where do they get their water from?

Note that both authors have lived in the Phoenix metropolitan area since the early 2000s.

Figure 10.1: Phoenix valley.

When the Mexican-American war ended in 1848, Americans started to explore the west in search of riches. One such explorer, Jack Swilling, while on an outing near the White Tank Mountains in 1857, noted an abandoned river valley. This valley, where modern-day Phoenix is located, in fact, has excellent terrain, fertile soils, and an excellent climate for farming. All that was required was water.

Swilling was not the first to recognize the farming potential of the valley. The Hohokam people lived in the valley for more than 1000 years and created 135 miles (217 kilometers) of irrigation canals. The Hohokam were very successful farmers and engaged in an extensive trade network that covered a significant portion of what is now the state of Arizona. There are many archaeological features that attest to the scale of Hohokam irrigation society. After about 1070 C.E. Hohokam society began to change and by 1450, the Hohokam abandoned the valley. The reasons for the abandonment are not known, but might relate to a period of major droughts and severe floods, destroying important physical infrastructure. It is the remains of Hohokam irrigation systems that sparked Swilling's imagination for what the potential of the valley might be. Swilling built a series of canals following the Hohokam system, thus founding Phoenix. The name Phoenix was chosen to reflect the fact that it is a city born from the ruins of a former civilization. Even today, canals providing Phoenix with water follow the ancient canal systems of the Hohokam.

Initially, Phoenix was a productive agricultural area for cotton and citrus with year-round sun and plentiful irrigation water from the Salt and Gila Rivers. In 1911, the Roosevelt dam was created east of the valley, which provided a more predictable source of water and, with it, the opportunity to grow to a population of 150,000 people. By damming the Salt River, water could be accumulated in the mountains near Phoenix and distributed via irrigation canals. As a result, the Salt River bed that runs through Phoenix is dry most of the time.

In 1922, the seven U.S. states that are part of the basin of the Colorado River created an agreement, called the Colorado River Compact. This agreement defines the allocation of water rights among the states of Colorado, New Mexico, Utah, Wyoming, Nevada, Arizona, and California. Based on historical rainfall patterns, the flow of the Colorado river was equally divided between the upper-division states (Colorado, New Mexico, Utah, and Wyoming), and the lower division states (Nevada, Arizona, and California).

Defining the rights to Colorado River water created opportunities for irrigation, and led to projects like the creation of the Hoover dam to harvest Colorado River water. Arizona was dissatisfied with the agreement and did not ratify it until 1944. Negotiations dragged on for almost 20 more years until specific disagreements with California were settled in the Supreme Court in 1963. The verdict specifies that California can use 50% of the river flow up to a maximum of 5.4 cubic kilometers annually, Nevada 0.4 cubic kilometers and Arizona the remainder of the lower Colorado river flow.

After solving the allocation problem, the way was cleared for the construction of the largest aqueduct system within the U.S., the Central Arizona Project. Three hundred and thirty-six miles of canals bring the water from the Colorado River to the urban areas in the central and southern regions of Arizona (Figure 10.2). The project started in 1973 and took 20 years to complete. Since the 1950s the population in the Phoenix metropolitan area has rapidly increased. This was made possible by the availability of affordable air conditioning. With the rapid growth of the city, agricultural land was transferred into urban use. Although urban use includes golf courses, swimming pools, and the domestic use of water, water use per ha in an urban setting is still considerably less than for agricultural land, which uses a lot of water for irrigation.

Figure 10.2: Central Arizona Project canal.

Nevertheless, since the 1990s there has been a drought that has led to lower water flows in the Colorado River. As water demand continues to grow, the city will either be forced to increase the use of groundwater or to direct ever more water away from agriculture to meet the demand. In order to evaluate the water use challenges for the future of Arizona, Arizona State University has developed a simulation model that enables observers to explore the consequences of droughts, population increase, and water policies. You can explore different scenarios of water use projections here.

In the beginning of 2022, a Tier 1 Water Shortage declaration was s put in effect since the water levels in lake Mead dropped below 1075 feet, and in 2023 a Tier 2a Water Shortage declaration was put in effect since the water level in lake Mead is expected to drop below 1050 feet. Crossing those predefined levels triggers water allocation reductions that amount to 21% of the allotted water to Arizona from the Colorado river, which has to be absorbed by the agricultural community, as specified by regulations from the 1960s. As one can imagine, the current water crisis leads to major disputes how water is used by some land owners, growing water intensive alfalfa to execute their individual water rights, and major investments, more than 1 billion dollars, by the state government to bring additional water sources to the state. Interesting is the focus on increasing water supply rather than on reducing water demand. It is an open question what a water secure future for Arizona looks like with the continuing change in water supplies driven by climate change.

The story of the development of Phoenix is one of how water scarcity was overcome with shared infrastructure. There is no way that a small group of individuals could build the Roosevelt Dam. Previous efforts by smaller groups to build canal systems often failed. Getting water to the arid west required a monumental effort. In this case, a central government was essential to provide public infrastructure due to the scale of the problem. This is not always the case. Sometimes smaller groups can solve such large-scale problems, as we shall soon see.

10.3 The Netherlands

The Netherlands is a small country in Western Europe. The average income per person is one of the highest in the world. However, 25% of the Netherlands is actually below sea level (Figure 10.3). Furthermore, some of the biggest rivers in Europe cross the Netherlands on the way to their final destination in the North Sea.

Figure 10.3: Map of the Netherlands with portion below sea level.

In contrast to Phoenix, the Netherlands is a place with an abundance of water. Over time the Dutch have had to solve various social dilemmas to keep their feet dry.

Before 800 C.E., the inhabitants of the precursor of the Netherlands used non-structural measures to keep their feet dry. Such measures like man-made hills or abandoning areas in times of danger were the result of decisions made by individual households. Due to increased population pressure, developments in technological know-how, and finance, there was a rapid increase in the development of structural water control measures after 800 C.E. Such measures included dikes and sluices. Construction and maintenance of these structures required cooperation within communities.

Farmers whose lands directly bordered the dikes agreed to commit themselves to the necessary construction work and maintenance activities. Coincident with the construction of dikes, drainage activities began to be developed as well. To make the lowland area inhabitable, it was necessary to get rid of the extra water. Small dams and sluices were built and maintained, based on similar agreements as for the flood protection systems between direct beneficiaries. A noticeable difference with regard to input for dike maintenance and small dams and sluices was the fact that in the case of the latter, all beneficiaries had to pay for the benefits received. These dikes, dams, and sluices are all quintessential examples of public infrastructure.

Originally, the local communities in the countryside were in charge of all general collective interests and took responsibility for water management as well. Around 1100 C.E., however, a new adaptation occurred as water management tasks gradually began to become separated from general public tasks. The reason is likely due to the increase in the number and severity of flooding events as well as a growing interdependence and complexity of the hydraulic works that began to stretch beyond the local scale.

Starting at the end of the 11th century and the beginning of the 12th century, the first public bodies charged with governing local and regional water management appeared on the scene and the phenomenon of the water boards was born. The purpose of the water boards was to construct and maintain the necessary hydraulic structures, providing safety through dikes and dry feet through drainage (Figure 10.4). Their establishment was recognized by the higher, regional authorities who still held themselves responsible for good water management but who resigned from their administrative duties.

Each of the water boards differed in their design and implementation of physical structures as well as rules. They were also confronted with different problems. They were not always successful in preventing floods or draining areas effectively. During the period known as the “Republic of the United Provinces of the Netherlands” from 1581 to 1795, there were severe floods and extensive peat-digging (for fuel), which caused unintended artificial lakes and diverse management problems. Still, the water boards survived this period. One of the main reasons for the long-term adaptation and survival of the water boards is the institutional arrangements upon which they are based. The design of rules was based on the shared norms and values of the population. Although the water boards were not always successful in maintaining safety and dry feet, they were maintained anyway because changing them would be costly. The benefits associated with switching to new and unfamiliar institutional arrangements in an effort to improve performance may have been outweighed by the costs of operating these new institutions. Thus, the water boards maintained the familiar institutional arrangements, which they knew how to operate and that they could adapt. Perhaps the roots of the shared norms in contemporary Dutch society goes back to those people who found ways to make the land liveable by developing institutions based on reciprocity.

Figure 10.4: Windmills at the Kinderdijk, near the hometown of one of the authors.

Since the Napoleanic occupation of the Netherlands in the early 1800s, there has been an increased centralization of water governance over time. Although water boards are still independent organizations, a ministry of water management was created in order to coordinate water management over the entire country. In 1953, a major flood in the south of the Netherlands killed 1800 people. This event led to an increased effort to protect the increasingly urbanized Netherlands from potential floods. As a result, there has been a huge amount of investment in infrastructure made to reduce the risks of flooding over the past 60 years.

Interestingly, major challenges for the future of the Netherlands do not come from the sea. The canalization of the river Rhine has made the river more suitable for the transport of goods to

Germany via ships, but it has also reduced the natural buffering capacity of the river. Removal of swamps that naturally would be areas to buffer excess water, now leads to rapid transport of water down the river during rain events. As a consequence, floodings now happen more frequently in the river delta. Upstream countries need to increase their buffer areas to reduce the flooding risks in downstream countries. This situation raises an important point: there are often inherent trade-offs when choosing among performance, robustness (the capacity to cope with change) and robustness to different types of shocks. So the Netherlands has become fairly robust to weather shocks from the sea but, in so doing (occupying more and more low-lying land) it has become more vulnerable to weather shocks from continental Europe (flooding of the Rhine due to rain). Can you run the same mental experiment with the situation in Arizona?

The history of the water boards shows a continuous tinkering with rules at different levels of organization and spatial scales. Disturbances like floods and the unintended consequences of peat digging have triggered the development of new rules and structures. The Dutch water boards illustrate how local-level governance structures may evolve into a resilient collaboration of multi-level governance structures when national institutions recognize the importance of smaller governance units and work with them rather than destroying them.

10.4 Bali, Indonesia

As we have seen, irrigation requires coordination and cooperation. One has to build infrastructure in order to move water around. This infrastructure needs to be maintained in order to function properly. By maintenance we mean cleaning of the canals and repairing damage to levies and diversion structures. Once the infrastructure is in place, water needs to be shared. In most cases, farmers who have their plots of land near the source of the water have preferential access to the water. In order for downstream farmers to get the water they need, the upstream farmers need to restrict their use of water. But given our earlier discussion of social dilemmas, why would they do this? An iconic example can be found on the island Bali, Indonesia.

Before we discuss this case, we would like to indicate that an increasing share of food production is dependent on irrigation where water is distributed from sources (rivers, groundwater, lakes, etc), to individual plots of land. About 70% of the global freshwater supply is used to irrigate small plots (over 90% of plots worldwide are less than 2 hectares or about 4 football fields). In order to get available water to the right location at the right time, a coupled infrastructure system is required which may vary in scale and complexity depending on the biophysical and cultural context.

Bali is one of the islands of Indonesia that has had a complex and very productive irrigation society for about 1000 years. Hundreds of irrigation communities called subaks are connected via waterways that begin at a lake near the top of an old volcano (Figure 10.5). Canals connect this water to subaks downstream. This complex irrigation system has been studied in detail by anthropologist Stephen Lansing and has been made famous in a book titled *Priests and Programmers*.

Figure 10.5: Subak irrigation.

The irrigators have to solve a complex coordination problem involving water distribution and pest control (they didn't have pesticides 1000 years ago). On the one hand, control of pests is most effective when all rice fields in a particular subregion are on the same schedule for planting rice. This is due to the fact that the pests (insects called planthoppers) are limited in their ability to move (or disperse) on the landscape. If large enough areas are kept fallow (areas without plants) between planted areas, the planthoppers can't cross them because there isn't any food to keep them alive while they cross (i.e., the fallow areas are "food deserts"). This keeps pest outbreaks localized if they occur at all. On the other hand, the terraces (see Figure 10.5) are hydrologically interdependent, with long and fragile systems of weirs, tunnels, canals, and aqueducts used to control where water goes, making it challenging to get the water to all the fields while maintaining large enough fallow areas between planted areas, all in a limited growing season.

To balance the need for coordinated fallow periods and water use, a complex calendar system has been developed that determines what actions should be carried out on each specific date. These actions are related to the spiritual practice of making offerings to “water temples” at several levels: small temples at the rice terrace level, the temple at the village level, and the temple at the regional level, associated with the Pura Ulun Swi, “the Head of the Rice Terraces” (this is the temple of the high priest Jero Gde, the human representative of the Goddess of the Temple of the Crater Lake, the main source of water for irrigation). These offerings of water and other items were collected as a counter performance for the use of water that belonged to the gods. These ritual practices trigger the calendar actions (i.e., people make offerings at particular times after which they can plant, etc.).

Balinese society consisted of many kingdoms before the conquest of territory of Bali around 1900 by the Dutch. The Dutch saw these offerings made to the various temples in a different light, namely as a royal irrigation tax. The fact that during the nineteenth century there were quite a number of kingdoms in Bali was a sign that the institution of kingship had weakened over time from one powerful kingdom to a number of smaller kingdoms. Therefore, the Dutch wanted to restore centralized government; in particular they wanted to use a revived royal irrigation tax to improve the irrigation system. The Dutch administrative reorganization failed, partly due to lack of funding, but also because historical analysis conducted during the 1930s demonstrated that there was no evidence that Bali had ever had a centralized government. Although Indonesia became independent from the Netherlands after World War II, many aspects of the colonial bureaucratic system were adopted by the new independent government.

During the late 1960s, the Indonesian government made self-sufficiency in rice production a major goal for national development. In the same period the Green Revolution began in Asia. The Green Revolution involved the spread of new rice-growing technologies that promised a dramatic increase in rice production. Bali was one of the first targets of the Green Revolution. In contrast to the earlier Dutch attempts to modernize rice production in Bali, this time the engineers were well funded.

The function and power of the water temples were invisible to the planners involved in promoting the Green Revolution. They regarded agriculture as a purely technical process. Farmers were forced to switch to the miracle rice varieties which would produce three harvests a year instead of the two that could be achieved with traditional varieties. Farmers were motivated by governmental programs that subsidized the use of fertilizers and pesticides. The farmers continued performing their rituals, but now they no longer coincided with the timing of rice farming activities. Soon after the introduction of the miracle rice, a plague of planthoppers caused a huge amount of damage to the rice crops. A new rice variety was introduced, but it was followed by another pest plague. Furthermore, water shortages began to occur because there was nothing to replace the rituals (which were now out of step with plantings) which had been the basis for the efficient allocation of water.

During the 1980s, an increasing number of farmers wanted to switch back to their old ritual-based system, but the engineers interpreted this as religious conservatism and resistance to change. Steve Lansing quotes a frustrated American irrigation engineer "These people don't need a high priest, they need a hydrologist!" (Lansing 1991 p. 115). It was Lansing who unraveled the function of the water temples, and was able to convince the financiers of the Green Revolution project on Bali that irrigation and rice cultivation was best coordinated at the level of the water temples. Lansing built a computer model of the artificial ecosystem, and showed that for different levels of coordination, from farmer level up to central control, the temple level was where decisions could be made to maximize the production of rice.

As this story suggests, the complex irrigation system on Bali and the role of the temples in operating it has evolved over a long history of local adaptations at different levels of organization and different spatial scales. The water temples played a significant role in the coordination of the use of water, but also in providing technical advice and mediating water use conflicts between different subaks. By making offerings to different temples, the farmers were made aware of the interconnections between the water flows at different scales. Due to Lansing's insight and analysis, some of these systems have evolved still further and avoided the fate of many self-organized systems of this kind when experts declared them defunct and constructed new infrastructure without paying much attention to local property rights, ecology, culture, and traditions.

10.5 Sanitation

A huge portion of potable water household use is designated for sanitation. It is remarkable that we use a valuable resource, water, to dispose of another valuable resource, human waste. Human waste is a valuable resource, largely as a natural fertilizer. Hunter-gatherers had a natural disgust for human waste, did not like the smell, did their necessities outside the campsite, covered it, and moved the camp around regularly to avoid any problems human waste could cause. Although they were unaware of disease transmission related to human waste, this behavior avoided disease spread.

When humans started aggregating in higher densities, a more elaborate system of recycling night soil was established in various locations. In Edo, the precursor of Tokyo, an elaborate poop economy was created to harvest the productive night soil to fertilize the nearby farmland. When a person was visiting another household, the person likely ran back home if they felt an urge for a number 2. Leaving a dropping at another household was considered a gift. The benefits of recycling night soil worked well in

Edo since Japanese had (and have) a strong culture of hygiene, they cooked their vegetables, had a sophisticated composting system, and heated their water to drink tea. This avoided – without their knowledge – the problem of spreading germs.

In Europe, urbanization has increased the distance between farmers and producers of human waste. This has reduced the economic viability of the reuse of human waste as fertilizers. Europeans started using water to dispose of human waste, leading for example to the Great Stink in London in 1858 due to a large volume of untreated human waste and warm weather. The discovery of the germ theory disease and the availability of cheap germ-free fertilizer alternatives (first guano, then artificial fertilizers) made the reuse of human waste less desirable.

If we focus on the current situation, about 25% of the human population has no proper sanitation facilities. This is a major causal factor in the spread of diseases. Providing these populations with a Western approach to human waste disposal is neither desirable nor sustainable. Human waste has a lot of potential to be reused for fertilizers and energy creation. The water intensive Western approach to human waste disposal is no longer option with the concurrent decrease in the availability of freshwater around the world and increase in human populations (and their waste). In 2011, the Gates foundation started an initiative to reinvent the toilet. There are many new approaches being explored to reduce water use and collect human waste at scale in a way that provides good jobs (instead of suppressing lower casts to do dangerous and “shitty” work). See, for example, the various examples from container-based sanitation that lead to reuse for fertilization and energy production.

Although sanitation might not be a typical dinner conversation, managing where the dinner ends up is a key issue for humanity that needs to become less wasteful. Changing this will require not only technical innovation but also changes in norms and perceptions of what is acceptable regarding where the dinner ends up. The changes in the coupled infrastructure systems relating to the management of human waste, as reported in the popular book by Zeldovich (2021), demonstrate the dynamics of economic, social and engineering processes that lead to continuous change of the infrastructure.

10.6 Critical reflections

Hydrological systems such as drainage and irrigation systems, can be found around the world for millennia, and are good examples of coupled infrastructure systems. Those systems demonstrate how the biophysical and cultural context impact the kind of engineering and institutional solutions that are possible.

10.7 Make yourself think

1. What are examples of hydrological systems near you?
2. What water distribution problem does the hard infrastructure aim to solve?
3. How is this hydrological system be maintained?
4. How much water does your household use, and where does the water come from?

10.8 References

Lansing, J.S. (1991) Priests and Programmers: Technologies of Power in the Engineered Landscape of Bali, Princeton University Press.

Wittfogel, K. (1957). Oriental despotism; a comparative study of total power. New York: Random House

Zeldovich, L. (2021) The Other Dark Matter: The Science and Business of Turning Waste into Wealth and Health, Chicago University Press.

PART V

PART 5: OUR CRUMBLING CRITICAL INFRASTRUCTURE

CHAPTER 11

Critical Infrastructure in Peril

Key Concepts

In this chapter we will:

- Describe Critical Infrastructure
- Explore how critical infrastructure is impacted by disturbances

11.1 Critical infrastructure

Critical infrastructure typically refers to physical infrastructure like the network of highways, connecting bridges and tunnels, railways, utilities and buildings necessary to maintain normalcy in daily life. Physical infrastructure allows the transportation of people, goods, water, sewage, electricity, oil and gas, information, etc. This infrastructure is referred to as critical since its malfunctioning could have major consequences for the economy and safety of people.

The USA has an agency devoted to managing the function of critical infrastructure, namely the Cybersecurity and Infrastructure Security Agency, whose mission it is to understand, manage, and reduce risk to our cyber and physical infrastructure. CISA distinguishes sixteen sectors and they mainly relate to hard infrastructure like dams, manufacturing, energy and transportation, but also include mixed infrastructure systems like health care and the financial system.

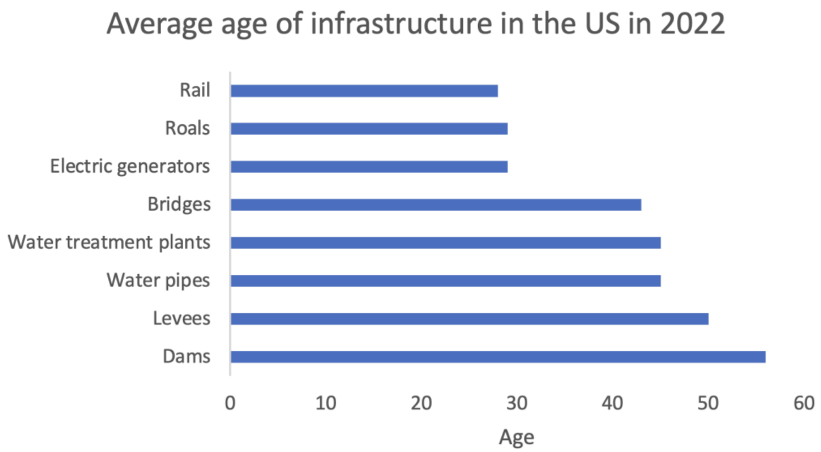

Figure 11.1: Average age of different types of infrastructure in the USA. Source: Statista.

In the rest of this chapter we will discuss some examples of critical infrastructure, especially within the USA, from a CIS perspective and the challenges they may experience in the near future.

11.2 Vulnerability of Critical infrastructure

The main challenge for critical infrastructure is that physical infrastructure deteriorates over time, and the deterioration of infrastructure is not always clearly visible. Managing infrastructure thus requires inspections to identify potential vulnerabilities. For example, around 8% of the bridges in the USA are identified as structurally deficient, but are still in use.

Every four years, the American Society for Civil Engineers (ASCE) evaluates the state of the American Infrastructure. The report card for American Infrastructure after the most recent evaluation is a C-. Types of infrastructure that get a C or higher are bridges (but see above the share of structurally deficient bridges), rail, ports and solid waste. Obviously the ASCE has a stake in having a critical evaluation of the infrastructure and recommending more work. Nevertheless, the state of its infrastructure has been a concern in the USA for recent administrations. Improving the state of infrastructure is one of the few bi-partician priorities as witnessed by the significant Infrastructure investments during the Biden administration.

In the following sections we discuss a number of examples from a coupled infrastructure perspective. We will explore potential vulnerabilities of the hard infrastructure to anticipated changes in the social and biophysical environment such as climate change and artificial intelligence.

11.3 Transportation Infrastructure

ASCE gave American road infrastructure a grade of “D” since the roads “are often crowded, frequently in poor condition, chronically underfunded, and are becoming more dangerous.” The condition of roads is impacted by weather conditions (frost, rain, heat), and the stress imposed by the combination of the weight and number of vehicles on the road. A country like the USA is car-oriented, which impacts the way cities are developed and the level of widely accessible public transportation that is provided within and between cities and towns. Despite the physical layout of cities centered around cars, there are traffic jams. The average time delay for commuters is steadily increasing to more than 40 hours a year.

The quality of roads is measured each year via a random sample of roads in each state using the International Roughness Index, and the results show that quality has been improving over the years. This seems inconsistent with the experience of the users who complain about deteriorating roads. This is even a popular complaint by US Presidents in State of the Union addresses.

One reason for this discrepancy about the measured and experienced quality of roads is that there has not been a systematic evaluation of the quality of the roads users actually experience, only a random sample of road surfaces in various states. This may change when new technologies enable car drivers to use an app to monitor the road quality they experience. Getting information from the road user's perspective instead of the road provider's perspective may lead to better feedback on the quality of the roads.

There is, however, a systematic underinvestment in road infrastructure, largely due to the lack of tax collected to keep up with the increasing costs of maintenance. For example, the Federal Highway Administration's Trust Fund functions as the mechanism to generate revenue, mainly from fuel taxes, and distribute the resources to approved highway projects across the USA. The fuel tax per gallon has not been increased since the 1990s, and with increasing fuel efficiency of cars (and now also electric vehicles), it is no surprise that since the 2000s, expenditures have been higher than the revenues, requiring various interventions to keep the Highway Trust Fund operational.

In sum, US citizens depend on cars and roads, but the US population as a whole is not investing in proper maintenance at the local and federal level.. The prospects for road infrastructure are not great with some of the changes that are expected to impact road infrastructure in the near future. Novel or extreme weather conditions due to climate change could impact the road quality. But perhaps the biggest vulnerability of the functionality of the road infrastructure are the rapid changes in vehicle transportation.

In recent years we have seen an increase in miles traveled due to deliveries of packages from online shopping and buying rides via Uber and Lyft. We also start seeing autonomous vehicles, and improved road assistance technology that could impact the capacity of the roads. In theory, those ride technologies may increase the capacity of roads since artificial intelligence supported coordination between cars allow them to drive closer to each other without causing traffic jams.

Those changes in technology and use of cars may impact how people commute and where they live. In fact, the COVID-19 pandemic has caused a change in human behavior that is expected to last, namely an increase in working from home. This may lead to changes where people want to live and thus demand for road infrastructure.

The path dependency of road infrastructure is an important factor in how new technological solutions can be implemented. Phoenix, the city where we are located, is designed as a grid system, where each mile there is a main road, and typically on the corner of each grid there is a gas station. Such a layout stimulated an urban sprawl of single floor single family houses. Such an urban design does not lend itself to transitioning to a high density urban area dependent on public transportation. On the other hand, old European cities were built very compact and do not accommodate a high dependency on cars. We see more diverse use of transportation systems, from public transport to scooters and bicycles.

Those different urban layouts may impact how technology could be implemented. Phoenix is a test location for autonomous vehicles due to its grid system and good weather. It is unknown whether autonomous vehicles will become reliable enough to drive within old European cities. Both spread out American cities and compact old European cities will have high costs to adjust design to new technologies. Since most of the new urban development is happening in the global south, how will transportation infrastructure, especially roads, be implemented?

Most cities in the global south provide a mixture of public transportation and motorized road transportation. The rapid increase of motorized vehicles, motors and cars, has led to problems with air pollution leading to the death of millions of people each year. Cities in China and India are known for their smog during certain times of the year. In China, this has stimulated the transition to electric vehicles to reduce air pollution.

The experience with transportation systems in the global south is mixed as experienced by the authors. The metro and train systems in China and Japan are crowded but reliable, while the roads in those countries have continuous traffic jams. It is difficult to see how an increase in transportation capacity can be accommodated. Traveling in India is a life changing experience since there seems to be no acknowledgement of the rules of the road. The high density of motorcycles, cars and cows makes the roads dangerous and polluted. . Also here it is difficult to see how an increase in transportation demand can be accommodated, especially since technological solutions like autonomous and electric vehicles are less suitable.

In conclusion, road infrastructure experiences long term challenges in maintaining sufficient capacity for transportation demand and controlling air pollution. Although public transportation is a desirable solution to address those challenges, the trends are an increased individualization of motorized transportation.

11.4 Water and Waste Water Infrastructure

We have noted that all infrastructure types share fundamental features and one of the most important functions of infrastructure is to move stuff, whether people, goods, electrons, information bits, or water from one place to another. Thus water infrastructure is a special kind of transportation infrastructure that moves, you guessed it, water from one place to another. It is a curious combination of natural (rivers, and lakes) and built (canals, dams, weirs, tanks) infrastructures that moves water from the tops of mountains to our crops and taps. It has a curious feature in that water can be used over and over again for different purposes and a critical attribute of water is its quality. So water infrastructures typically involve moving clean water to a given location and move dirty water away from that location. Although efforts to create infrastructure to provide safe drinking water and dispose of waste waters goes back thousands of years, it was the 1800s during which a rapid increase of this type of infrastructure began to be observed in urban areas in the USA and Europe. Nowadays most of the world has access to safe drinking water, although safe drinking water from the tap is restricted to mainly North America, Western Europe, Japan, Australia, New Zealand and Saudi Arabia.

The huge investments in water infrastructure in the 1800s and early 1900s in North America and Europe has led to an aging infrastructure with an average age of a water pipe in the USA of 45 years. About 10-20% of the water is leaking and thus wasted in the USA. This is significant but considerably lower than more than 50% in various cities in the global south. In fact, the state of the infrastructure is difficult to measure since pipes are underground, and we mainly find out when pipes burst.

A tragic example is the water crisis in Flint, Michigan, where starting in 2014 the drinking water was contaminated by lead caused by short cuts made to save money with an aging water distribution system to provide drinking water to the community. There are many communities like Flint, Michigan, where the drinking water infrastructure is outdated and could lead to public health challenges. Regulation changes in 1986 and 1996 prohibit the use of materials that could lead to lead in drinking water, but still there are an estimated 6 to 10 million service lines in the USA. To be able to provide safe drinking water to customers in the USA, the price of water is expected to continue to rise substantially to cope with the aging infrastructure.

Climate change is expected to impact the water infrastructure at different levels is significant ways. Changes in precipitation and temperature lead to long term droughts and rapid flooding events. The current infrastructure is not able to cope with those changes. This may lead to changes in capacity (to increase capacity to buffer major rainfall events, also be increasing natural infrastructure like wetlands), and reduction of demand. Demand reduction is partly possible by technology (drip irrigation, low water use toilets, efficient shower heads) but also changes in behavior (change in landscaping to reduce need for irrigation). Increased amount of recycling of waste water is happening in various places in the southwest of the USA. This technological solution also experiences social challenges since people have a resistance to drink recycled pee, while the quality of the recycled drinking water is superior.

In the global south water infrastructure is more diverse and more unequal. If you have piped water, you are not recommended to drink it without boiling the water first. Poor neighborhoods often do not have piped water and rely on water trucks or buying bottled water. Waste water is often not treated leading to pollution and eutrophication of waterways.

In sum, the aging water infrastructure in North America and Europe will be expensive to update and adapt to changing climate reality. There is increasing availability of water infrastructure in the global south, but there is significant inequality in availability and performance.

11.5 Energy infrastructure

Like water infrastructure moves water, energy infrastructure moves electrons. Similar to water, modern energy infrastructure, especially electricity, stems from the late 1800s and early 1900s. The centralized electricity system provides a reliable supply of energy made possible by the creation of standards (like the alternating current (AC) from Nicolas Tesla to distribute electricity). The creation of a centralized infrastructure allowed it to connect to large electricity generation facilities such as hydroelectricity generating dams, nuclear power plants, coal power plants, etc.

Electricity is much more difficult to store than water, requires batteries (the analogue of a reservoir in water systems, so one needs to produce electricity when there is demand (imagine trying to 'produce' water on demand). The combination of sources varying from low cost inflexible nuclear energy, to flexible but costly fossil fuel power generation, allows the centralized electricity system to function without frequent power outages.

About 40% of the energy inputs (fossil fuels, gravity in hydroelectric systems, wind, solar radiation) are used for electricity generation, the rest of fossil fuels and renewable energy is used directly for mainly transportation and industrial purposes. Think about the gasoline used for cars, and fossil fuels used to power large factories.

Like other critical infrastructure, energy infrastructure is aging, and there is a major shift in sources of electricity generation that will have a major impact on the vulnerability of the infrastructure. The increasing use of solar, wind and other decentralized energy sources makes the centralized electricity grid more difficult to control. Especially solar and wind power are not reliable sources, and other electricity generation needs to kick in when the sun sets or the wind calms. Those are expensive solutions. In fact, during some parts of the day there might be a surplus of electricity that has to be wasted (grounded), while other times expensive solutions have to be used.

Another trend is a decentralized energy system where neighborhoods or households go off the grid, relying on solar power with batteries to store electricity surplus. A vulnerability created by this approach is that those who go off the grid are often wealthier than the average American and do not continue to provide financial support for the public electricity grid. This will make the financial sustainability of the public electricity infrastructure more vulnerable.

Another dependency is the use of water for cooling and hydroelectric power. With water shortages in the southwestern USA (and with water temperature increasing), this could also impact electricity generation. This will be especially challenging during times when energy demand will be high (for running air conditioners). As such the interaction between water and electricity infrastructure increases the risk of blackouts during the summer months in the southwestern USA.

The war in the Ukraine revealed the vulnerability of international energy infrastructure where especially Europe depends on the natural gas and oil provided by Russia via an international system of pipelines. European countries started to decommission nuclear and coal power plants to meet environmental targets, and import gas and oil. However, those decisions had to be reversed, together with a major reduction in energy demand from industry and households to keep the economic system functioning.

Those complex centralized and cross-border energy infrastructures can be found in most countries in the global south, especially with hydroelectric power. However, many rural communities bypass the centralized energy system by adoption of affordable solar energy that can provide sufficient electricity for key energy uses. An open question is whether those solar energy systems are sufficient with the rapid increase of demand for electricity when rural areas start urbanizing.

11.6 Food Supply Chain infrastructure

Where does our food actually come from? As water infrastructure moves water, energy infrastructure moves electrons, food supply chain infrastructure moves food, and all three infrastructures are intertwined. Many of us buy food in the supermarket and are not involved in food production. However, this is only possible due to the complex agricultural systems connected with an international food supply chain system. We focus in this section on the distribution of food after its primary production.

Another eye opening vulnerability we experienced during the COVID-19 pandemic is the vulnerability of the food supply chain. Due to COVID outbreaks in some major meat processing facilities within the USA, there was a scarcity of certain types of meat. Those meat processing facilities were extra vulnerable due to the close proximity of the workers making it convenient for the virus to spread. Other disruptions experienced in the USA during the pandemic was the lack of the migrant workers that help harvesting, especially in California. And a lack of trained truck drivers caused delays in the transportation of food stuffs.

The war in the Ukraine, the main source of wheat for many countries in Africa and Asia, is causing food insecurity at an international level. As you might guess, the problem is caused, in part, by disrupted transportation infrastructure that relies on waterways. Specifically, wheat could not leave the Ukraine due to blocked ports in the Black Sea. The story of the Ever Given is a stark reminder of the vulnerability of our food supply chains to vagaries of coupled infrastructure systems.

On the morning of 23 March 2021, this giant container ship, one of the largest in the world, suddenly ran aground diagonally while passing through the Suez Canal on its way to Rotterdam and blocked the entire canal. Because of its enormous size, shipping traffic was jammed in both directions for six days. Billions of US dollars' worth of trade on hundreds of vessels came to a standstill. This curious combination of private infrastructure in the form of an enormous ship designed for transport efficiency and shared infrastructure of the Suez canal illustrate 'bottleneck fragilities' in coupled infrastructure systems that can impact billions of people.

The 'bottleneck' in the Suez Canal example is the reliance on a single constellation of infrastructures: a single route and a single mode of transport. In food systems, such bottlenecks take various forms, especially in the natural infrastructure component of the system, most specifically, genes in seeds. Genes can be seen as information storage infrastructure (software code) and seeds as a device for storing the code. You store the 'genes' of your photos in a sequence of zeros and ones on a 'seed' in the form of a USB flash drive. The code for plants has been written over millions of years of trial and error to cope with many conditions. We edit the code to maximize output of a particular plant in very controlled conditions. Controlling those conditions requires a lot of other infrastructure, e.g. chemical fertilizers, water control, etc. One thing we can't control is the susceptibility of plants to pests that mutate constantly. If we rely on one food plant (one canal) that is decimated by a pest (the ship runs around), we are very vulnerable indeed. It doesn't take too much imagination to see how our livestock systems are vulnerable to disease outbreaks where livestock are moved around to owners who specialized in different parts of the life cycle in a sort of constant 'super spreader event'.

11.7 Critical reflections

Infrastructure, once constructed, lasts for a long time. Infrastructure needs to be maintained by a specialized workforce. The longevity of infrastructure makes it costly and time consuming to adapt to changing conditions. With a rapidly changing society and climate change, how can infrastructure be adjusted to new conditions?

11.8 Make yourself think

1. What disruptions of infrastructure systems did you experience during the COVID-19 pandemic?
2. Do you know what happened, when you experienced a disruption in electricity or internet connectivity in your household?

11.9 References

Janssen, M.A., J.M. Anderies, A. Baeza, H.L. Breetz, T. Jasinski, H.C. Shin, and S. Vallury (2022), Highways as coupled infrastructure systems: An integrated approach to address sustainability challenges, Resilience and Sustainable Infrastructure 7(2): 100-111.

Walker, B., Crépin, A. S., Nyström, M., Anderies, J. M., Andersson, E., Elmqvist, T., ... & Vincent, J. R. (2023). Response diversity as a sustainability strategy. Nature Sustainability, 1-9.

CHAPTER 12

Natural Infrastructure Exploited

Key Concepts

In this chapter we will:

- Explore examples of sustainable and less sustainable use of natural resources
- Demonstrate the importance of aligning the incentives with the biophysical context

12.1 Introduction

Humanity relies on the natural environment as a resource for food, shelter, materials, recreations, like all other animal species. Humans are different from other animal species in that we have developed cumulative cultural practices to develop technology and organize activities to extract ever more resources from the natural environment. Whether this is by extracting timber from forests, growing crops, mining for minerals, using water to dispose of human waste, fly to the other side of the world to enjoy natural wonders, or use other animals for testing medical procedures. Natural infrastructure is comprised of the network of organisms and physical processes that process information and materials by capturing photons from the sun to assemble the resources we extract. We can extract more than other animals, since we have developed social, human, soft and physical infrastructures.

From a CIS perspective, it is important to recognize that natural resources can only be extracted and consumed because of our investments in other types of infrastructures. You may have heard about the term ecosystem services, which is often used in sustainability debates. It is suggested that nature provides some absolute level of services for human consumption such as plants cleaning air and filtering water, bacteria decomposing waste, bees pollinating flowers and plants, and trees protecting soil from erosion. We argue that the term ecosystem services is a misnomer and can be misleading since the natural infrastructure does not provide any services without the other types of infrastructure. One needs to have knowledge of the ecosystem, create norms and regulations to avoid overuse of resources, and have access to physical infrastructure like boats and chainsaws to extract the resources natural infrastructure creates.

In this chapter we discuss three types of natural resources that have been central to the development of sustainability science; fisheries, forestry and livestock, and discuss the interaction between different types of infrastructure for sustainable use of the natural resources. We also present a contemporary example of a more exploitative use of natural resources.

12.2 Fisheries

The Maine lobster fishery is a remarkable story of self-governance within the contemporary United States. The lobster fishery of Maine is organized into territories along most of the coast. Day-to-day fishing regulations (soft human-made infrastructure) are organized by harbor gangs. These harbor gangs are informal groups that enforce local customs. In order to fish for lobster you need to become a member of a harbor gang, which is the group of fishers who go lobstering from a single harbor (natural infrastructure). Members of this group can only set traps (hard human-made infrastructure) in the traditional territory of the harbor group. There are various fishing practices that each member is expected to obey (Figure 12.1). These practices vary from harbor to harbor. To become a member of a harbor community (a formal position) requires participation by family members in that community for several generations. People who are not born and raised in these harbor communities are considered outsiders and will have difficulty gaining the level of acceptance by the rest of the community that is required before the right to fish for lobster would be granted. In all harbor communities a person who gains a reputation for damaging others' gear or for violating conservation laws will be severely sanctioned. For example, if a fisher goes out to collect his traps, and discovers that somebody else has put traps in the same location, he may signal that this norm violation has been noticed by taking a lobster from the trap and leaving the trap open. If violations of accepted locations of traps continue, more severe measures can be taken, such as damaging gear and cutting the traps loose.

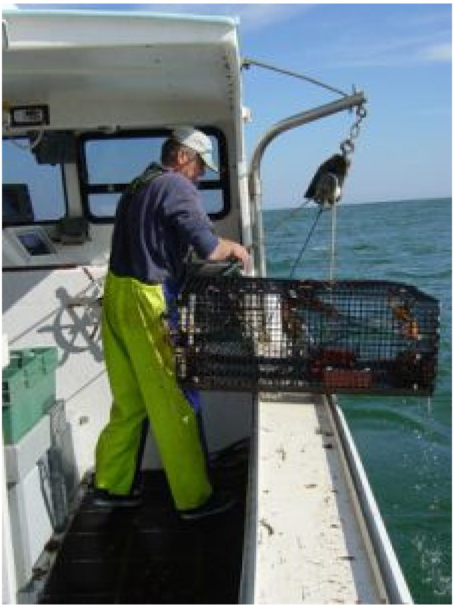

Figure 12.1: Lobster fisher.

If a fisher puts traps in another gang's territory, similar types of enforcement can be expected. The damages eventually make it unprofitable to continue breaking the informal rules, but are not so severe as to initiate a legal dispute. There are also formal laws in the state of Maine to protect the breeding stock and increase the likelihood that the regeneration rate remains high. The most important conservation laws are the minimum and maximum size measures, a prohibition against catching lobsters with eggs, and a law to prohibit the taking of lobsters which once had eggs and were marked (i.e., the V-notch law) (Figure 12.2). When a fisher collects the catch from his trap, he measures each lobster. If it is too small or too big, it will be thrown back. Such a rule avoids catching young lobsters, allowing them to mature to an age at which they will start generating offspring. If a lobster is caught that is carrying eggs, it is a productive female. The lobster will get a V-notch in the tail, and will be thrown back. This is also the case if a lobster is caught with a visible V-notch. This policy ensures that fertile female lobsters are kept alive.

Figure 12.2: V-notch in tail of lobster. Notice the eggs.

Why would fishers not cheat? How could other fishers find out if a V-notched female was caught? How would they find out if a lobster that is too small was caught? It turns out that the cooperation of the middlemen (the lobster buyers) in the lobster industry is required to prevent cheating. These middlemen do not accept ineligible lobsters because their long-term financial viability depends on a productive lobster population. This gives them an incentive to help enforce the rules (by simply not buying such lobsters). Interestingly, neither the state nor any of the lobster gangs has tried to limit the quantity of lobster captured. Further, the state does not try to limit the number of lobster fishers, since this is already done at a local level. However, the state has been willing to intercede when issues exceed the scope of control of local groups. In the late 1920s, when lobster stocks were at very low levels and many local areas appear to have had compliance problems, the state took a number of steps (including threatening to close the fishery) that supported informal local enforcement efforts. By the late 1930s, compliance problems were largely resolved and stocks rebounded. Note however, that there are still too many unknowns about lobster biology to make the claim that change in management was the cause of the recovery. Figure 12.3 shows the decline of lobster catches from the 1880s until the 1930s. The absolute numbers dropped, as well as the catch per trap and per fisher. Since the 1940s the catch numbers have increased, with a rapid increase after 1990. Since the catch numbers per trap and per fisher are also increasing, this suggests that the lobster population is in good condition. If the rapid increase in total catches were caused by overharvesting we would expect a reduction in the catch per trap over time, since a decrease in the total number of available lobsters would cause an increase in the time needed to catch each lobster. Thus, each fisher would be catching fewer lobsters in a given season. But if more fishers are fishing, the total catch could remain the same, or even increase, while the stock is being depleted. These are the classic symptoms of overharvesting. Figure 12.3 shows, in fact, that the catch per trap and catch per license are actually increasing. This suggests that the lobster fishery is not being overharvested, at least at the present moment.

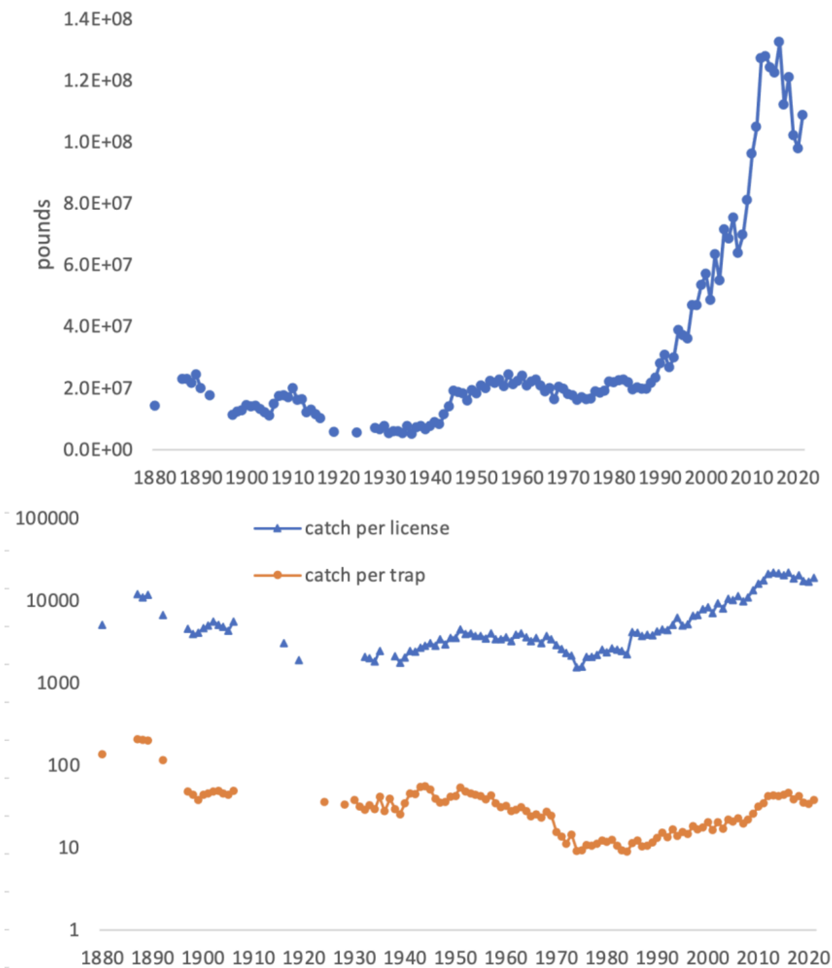

Figure 12.3: From top to bottom: (a) Lobster catch (in lbs) since 1880. (b) Lobster catch per trap and license based on data from the State of Maine historic data.

Recently, in response to changes that were breaking down the informal harbor gang system, the state has formalized the system by dividing the state into zones with democratically elected councils. Each council has been given authority over rules that have principally local impacts—trap limits, days and times fished, and so forth. Interestingly, the formalization of local zones was followed almost immediately by the creation of an informal council of councils to address problems at a greater than local scale.

The success story of Maine lobster fisheries is in stark contrast to the state of fisheries in general, and ocean fisheries in particular. The Food and Agricultural Organization of the United Nations (FAO) published a state of the world fisheries and aquaculture report in 2020 which shows an increase amount of capture of fish. Fish consumption has increased from 9 kg in 1961 to 20.5 kg in 2018. Most of the increase in fish production is caused by the increased importance of aquaculture (Figure 12.4). The natural infrastructure of marine resources shows an increased depletion over time (Figure 12.5) compensated by the creation of coupled infrastructure systems of fish production via aquaculture that are industrialized monocultures leading to the current situation in which about half of the current fish production comes from aquaculture.

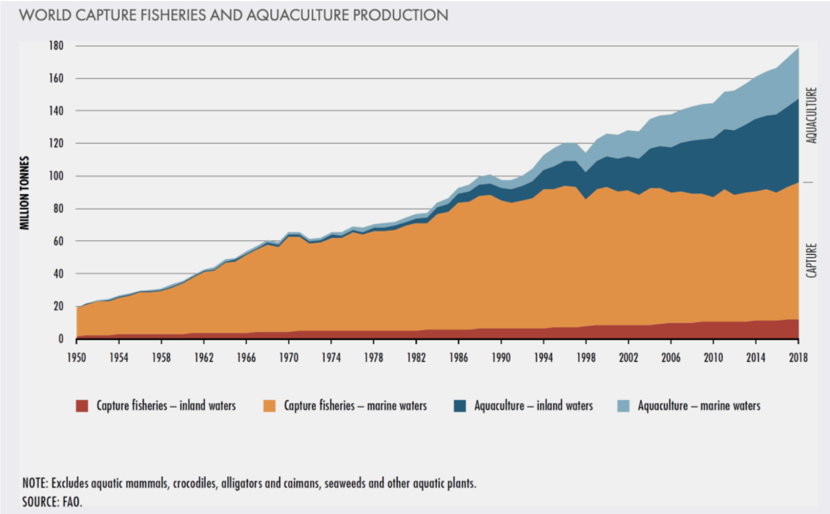

Figure 12.4: World capture fisheries and aquaculture production.

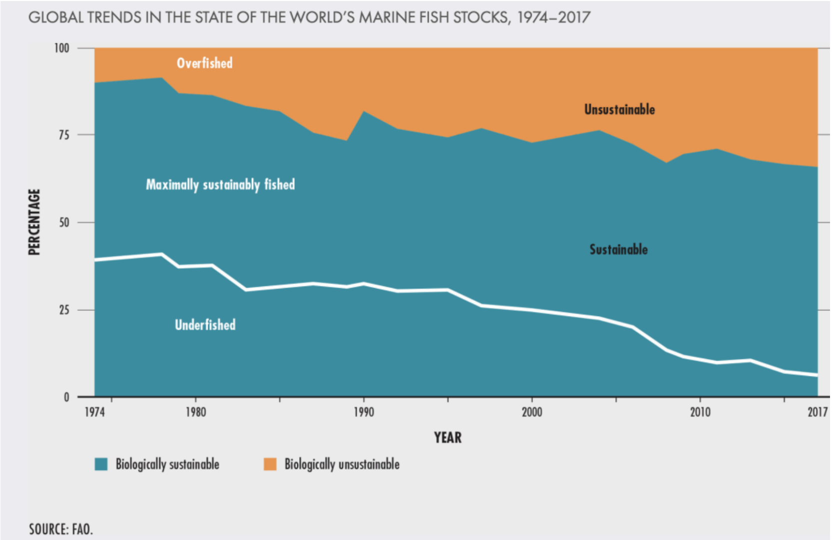

Figure 12.5: Global Trends in the state of the World's Marine Fish stocks between 1974-2017.

12.3 Forestry

What are effective ways to protect forests? Due to the importance of forests for biodiversity, there has been an increasing focus on creating parks and protected areas. One key concern is whether these areas are best protected by putting a fence around them or allowing human populations to continue to occupy them and help with conservation efforts. A related concern is whether designated areas become “paper parks,” i.e., areas set aside for protection on paper, but in practice, the lack of enforcement allows for a lot of poaching and illegal logging. Ostrom and Nagendra (2006) discuss long-term studies of land use change to test which type of management is most effective. They compared governmental, community, and private forests and found that the particular form of ownership is not important for the condition of the forests as measured by the quality and size of the trees. More important is whether boundaries have been well established in the field and are considered legitimate and whether regular monitoring and enforcement of rules related to entry and use exist (Figure 12.6).

Figure 12.6: Bicycles and trucks confiscated from people caught illegally removing large logs from the forests. Note the circular modification in the cycle frame (Inset) made to hold large logs of teak wood.

Whether the boundaries are considered legitimate depends on whether people have lived in the protected area before or if surrounding populations have used the resources over an extended period of time. If the boundaries are considered legitimate, how can they be monitored? For example, if indigenous populations are taken out of the protected areas (as has been the approach in several cases), one also loses potential monitoring capacity and it is not uncommon to see an increase in poaching in protected areas. Rather than relying in indigenous populations, perhaps paid guards are the answer? What are the incentives facing paid guards to monitor and enforce the rules? If guards are not paid well, they might be willing to accept bribes and not bother those who harvest illegally. This has also been observed frequently.

Community forests (as opposed to those run by the state with paid monitoring and sanctioning) can be effective since the population who benefits most from protecting the forests also monitors their use. In this case, because the cost of monitoring is aligned with its benefits, there is an incentive for high level monitoring effort, and those who are caught breaking the rules will experience social sanctions. On the other hand, community forests might be less effective in dealing with intrusion from outsiders due to lack of capacity and legal abilities to limit access (i.e., social sanctions won't be effective on outsiders). In this case, community forest managers need the assistance of the state. This is a clear example illustrating the importance of multilevel governance. Governmental and privately owned forests can be effective if sufficient effort is made in enforcement. But they might be more prone to corrupt guards. Further, when local people do not feel a sense of ownership and participation in the process, they are less willing to assist with monitoring (hiring enough guards is prohibitively expensive). This will lead to a lack of sufficient eyes to monitor the use of the forests.

The FAO also provided a state of the forest report in 2020. The changes in the forest stock varies across continents, in some areas there is a net increase of forest, especially Asia and Europe, due to active reforestation programs. In other regions, there is deforestation due to mainly increased conversion of forests to agriculture (especially to accommodate livestock production). There is also an increase amount of fragmentation of forest impacting habitats of animals and plants, and therefore biodiversity.

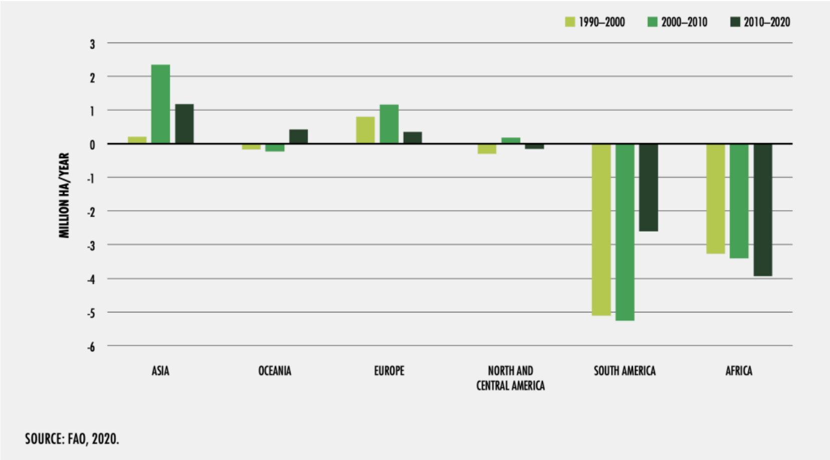

Figure 12.7: Net forest area change by region, 1990-2020 (million hectares per year).

Forests are net carbon sinks, but due to large scale deforestation and climate change this can change. Gatti et al. (2021) show that the Amazon is on its way to become a net carbon emitter. Building on that report, the Economist, reports that the Amazon, the lungs of planet Earth, is tipping to a net emitter. The causes are a combination of an economic crisis leading to illegal conversion of land from forests to agriculture and a Brazilian government who is not supportive of enforcing environmental policies.

12.4 Livestock

Since meat production is one of the main causes of deforestation, we focus the last example on livestock grazing. Although nowadays a large percentage of livestock production is done at an industrial scale, historically livestock production was done in harmony with natural and social infrastructure.

Törbel, Switzerland, is a village of about 600 people located in the Vispertal trench of the upper Valais canton. For centuries, Törbel peasants have planted their privately-owned plots with bread grains, garden vegetables, fruit trees, and hay for winter fodder. Cheese produced by a small group of herdsmen, who tend village cattle pastured on the communally owned alpine meadows during the summer months, has been an important part of the local economy. The earliest known written legal documents are from 1224, and provide information regarding the types of land tenure and transfers that have occurred in the village and the rules used by the villagers to regulate the five types of communally owned properties. On February 1, 1483, Törbel residents signed articles formally establishing an association to improve the regulation of the use of the alp, the forests, and the wastelands. The law specifically forbade a foreigner (Fremde) who bought or otherwise occupied land in Törbel from acquiring any right in the communal alp, common lands, or grazing places, or permission to fell timber. Ownership of a piece of land did not automatically confer any communal right (genossenschaftliches Recht). The inhabitants currently possessing land and water rights reserved the power to decide whether an outsider should be admitted to community membership (Netting, 1976, p. 139). The boundaries of the communally-owned lands were firmly established long ago, as indicated in a 1507 inventory document.

Figure 12.8: Cow in Törbel.

Access to this well-defined common property was limited to citizens, to whom communal rights were specifically extended. Here it is important to underscore why Hardin's use of the term "commons" is incorrect. The alpine meadows of Törbel are "commons" (Figure 12.8) in the sense that they consist of a common-pool resource over which there are no private property rights. It is property held in common with communal rights. Thus, as we discussed previously, "commons" is not equal to "open access," which refers to property with no rights attached. These pastures in Törbel are examples that not all "commons" end in tragedy as Hardin suggested. As far as the summer grazing pastures were concerned (the common-pool resource), regulations written in 1517 stated "no citizen could send more cows to the alp than he could feed during the winter" (Netting, 1976, p. 139). This regulation is still enforced today and provides for the imposition of substantial fines for any attempt by villagers to appropriate a larger share of grazing rights. Adherence to this "wintering" rule was administered by a local official who was authorized to levy fines on those who exceeded their quotas and to keep one-half of the fines for himself. Many other Swiss villages use the wintering rule as a means for allocating appropriation rights (frequently referred to as "cow rights") to the commons. This and other forms of cow rights are relatively easy to monitor and enforce. The cows are all sent to the mountain to be cared for by the herdsmen. They must be counted immediately, as the number of cows each family sends is the basis for determining the amount of cheese the family will receive at the annual distribution.

The village statutes are voted on by all citizens and provide the general legal authority for an alp association to manage the alp. This association, which includes all local citizens owning cattle, holds annual meetings to discuss general rules and policies and elect officials. The officials hire the alp staff, impose fines for misuse of the common property, arrange for distribution of manure on the summer pastures, and organize the annual maintenance work, such as building and maintaining roads and paths to and on the alp and rebuilding avalanche-damaged huts. Labor contributions or fees related to the use of the meadows are usually set in proportion to the number of cattle sent by each owner. Trees that will provide timber for construction and wood for heating are marked by village officials and assigned by lot to groups of households, whose members are then authorized to enter the forests and harvest the marked trees. Private rights to land are well developed in Törbel and other Swiss villages. Most of the meadows, gardens, grain fields, and vineyards are owned by various individuals, and complex condominium-type agreements are devised for the fractional ownership among siblings and other relatives of barns, granaries, and multi story housing units. The inheritance system in Törbel ensures that all legitimate offspring share equally in the division of the private holdings of their parents and consequently in access to the commons, but family property is not divided until surviving siblings are relatively mature.

Prior to a period of population growth in the nineteenth century, and hence severe population pressure on the limited land, the level of resource use was held in check by various population-control measures such as late marriages, high rates of celibacy, long birth spacing, and considerable emigration. The Swiss villagers have experienced the advantages and disadvantages of both private and communal tenure systems for at least five centuries, and they continue to use the communal tenure system. Although the yields are low, the land in Törbel has maintained its productivity for many centuries. Netting (1976) associates five attributes to land-use patterns with the differences between communal and individual land tenure. He argues that communal forms of land tenure are better suited to the problems that appropriators face when (1) the value of production per unit of land is low, (2) the frequency of dependability of use or yield is low, (3) the possibility of improvement or intensification is low, (4) a large territory is needed for effective use, and (5) relatively large groups are required for capital-intensive activities.

Figure 12.9: Gaddi shepherds with flock.

Not all owners of livestock own land, private or communal. Nomadic herders, or pastoralists, lead their livestock to graze around a large spatial landscape in order to be at the right place at the right time (Figure 12.9). In such cases, institutions have been developed to gain access to the land of various landowners. Our next example, the complex dynamics of Gaddi shepherds and their landscape in the Himachal Pradesh in India, demonstrates just such a situation. No particular place in this landscape is ideal for the maintenance of goats and sheep throughout the entire year. The only way that these animals can be cared for is to move them across a very large area with highly variable terrain. These pastoralists originally adapted their institutions to the harsh ecological conditions they faced in order to survive. They move their animals, goats, and sheep across a vast mountainous landscape within Himachal Pradesh. During the winter, they descend from the mountains and graze in the valleys and the lower elevation forests. The shepherds have made arrangements with agriculturalists (who own private plots of land) to graze on the stubble left after a harvest from private fields in return for the highly valued manure of the goats and sheep. In the summertime, it is too hot at lower elevations, so the pastoralists move into the mountains around the tree line. Lyall writes:

Snow and frost, in the high ranges, and heavy rain and heat in the low, make it impossible to carry sheep farming on a tolerably large scale with success in any part of the country. The only way is to change ground with the seasons, spending the winter in the forests in the low hills, retreating in the spring before the heat, up the sides of the snowy range, and crossing and going behind it to avoid the heavy rains in the summer (Lyall, 1872, p. 46; cited in Chakravarty-Kaul, 1998).

These seasonal movements are based on reciprocal relationships. The Gaddis shepherds invest a lot of time in social networking among themselves and with outsiders to provide access to grazing areas in return for manure and other goods and services. The informally-evolved rights of the Gaddis shepherds have never been formally recognized by the national government. In 1947, the Indian government adopted policies that reduced the shepherds' access to the usual grazing grounds by building dams to generate hydropower and by providing strictly private property rights to farming communities. This has resulted in more concentrated areas where livestock can graze, and may have contributed to erosion in the forested hilly regions. The government has accused the Gaddis of free-riding within this commons dilemma. However, the government had not recognized the efficient system that the participants in this action situation had already worked out; in fact, the shepherds and agriculturalists had developed an effective bargaining solution by trading manure for grazing rights. The shepherds adapted to temporal and spatial variability in their system by moving around the landscape in a particular, well-ordered pattern. Activities that hinder this movement pattern on the landscape hit the vulnerable point of this transhumance system (the seasonal movement of people with their livestock between fixed summer and winter pastures). When these movement patterns are affected, the shepherds are forced to use a smaller area which may, in turn, lead to overgrazing. Thus, the transhumance system is highly tolerant to seasonal variation through very specific institutional arrangements, but is extremely vulnerable to changes in access by social or physical barriers.

Global meat consumption has increased fivefold in the last 60 years. There is a strong correlation between meat consumption and income, and when countries like China get richer, we see a rapid increase in meat consumption. To produce this meat, it is not possible to rely on traditional pastoralism or extensive livestock production. As a result, livestock production has industrialized. Animals are held in industrial complexes and provided fodder produced elsewhere (leading to deforestation). The resulting manure is typically not recycled back as natural fertilizer, but emitted to the environment leading to nitrogen emissions. Intensive livestock production is also a source of zoonotic diseases as observed in Europe and Asia, where diseases jump successfully from livestock to humans.

There are various ways to produce beef, each of which may lead to different environmental impacts. On average, beef production is the most damaging for the environment (Figure 12.10). The required land use, water consumption, and emissions make it, by far, the most damaging type of meat consumption per 100 grams of protein. The development of plant based meat alternatives could make an important contribution to reducing the environmental impact for those who do not switch to a vegan or vegetarian diet.

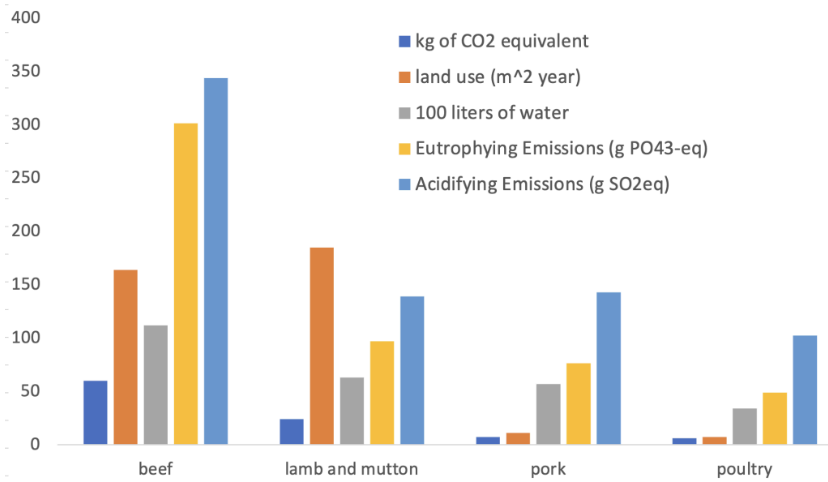

Figure 12.10: Environmental impacts per production of 100 gram protein. Based on data from Parlasca and Qaim (2022), Poore and Nemecek (2018) and Mekonnen and Hoekstra (2012).

12.5 Critical reflections

There are many examples of self-organized governance of natural infrastructure that have been successful over very long time periods. In the successful examples, we see that local communities play an important role, often crafting the rules and monitoring and enforcing those rules. Hence we see major investments in social, human and soft infrastructure to govern natural infrastructure. However, with rising human populations and changing consumption patterns, we see a shift to investing in more standardized hard infrastructure decoupled from investments in fair social, human and soft infrastructure. The lower connection of the biophysical and social context to produce the affordances of the natural infrastructure, lead to more environmental impacts.

12.6 Make yourself think

1. Do you make use of public parks? What is the state of the park? What are the rules, and how are they enforced?

2. Reflect on your diet and the environmental impact of your food choices.

12.7 References

Acheson, J. M. (2003). *Capturing the commons: devising institutions to manage the maine lobster industry.* University Press of New England.

Chakravarty-Kaul, M. (1998). Transhumance and customary pastoral right in Himachal Pradesh: claiming the high pastures for Gaddis. *Mountain Research and Development* 18: 5–17.

Gatti, L.V., Basso, L.S., Miller, J.B. et al. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature* 595, 388–393. https://doi.org/10.1038/s41586-021-03629-6

Lyall, F. (1872). *Final report of the regular settlement of Kangra*, 1868–72. Lahore, India. Government of the Punjab.

Mekonnen, M.F. and A.Y. Hoekstra (2012) A Global Assessment of the Water Footprint of Farm Animal Products, *Ecosystems* 15: 401-415.

Netting, R. M. (1976). What alpine peasants have in common: observations on communal tenure in a swiss village. *Human Ecology* 4: 135–146.

Ostrom, E. (1990). *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press.

Ostrom, E. & Nagendra, H. (2006). Insights on linking forests, trees, and people from the air, on the ground, and in the laboratory. *Proceedings of the National Academy of Sciences USA* 103: 19224–19231.

Parlasca, M.C. and M. Qaim (2022) Meat Consumption and Sustainability, *Annual Review of Resource Economics* 14: 17-41.

Poore, J. and T. Nemecek (2018) Reducing food's environmental impacts through producers and consumers, *Science* 360: 987-992.

CHAPTER 13

Building non-physical infrastructure is hard

Key Concepts

In this chapter we will:

- Provide examples of human, social and soft infrastructure
- Explore the vulnerabilities of underinvestment in human, social and soft infrastructure

13.1 Introduction

Humans invest time and energy to acquire knowledge and skills, build relationships, and develop rules and norms. In this chapter we will dig deeper into these three types of infrastructures and see an integrated perspective of sociality. We then discuss the implications for building and maintaining other types of infrastructure. As we will see in the next part of the book, there might be many technological solutions possible for a sustainable transition, but we may underestimate the effort needed to maintain these non-physical infrastructures.

13.2 Gaining Knowledge and Skills

Humans have to acquire a lot of knowledge to function in society. We learn by imitating behavior of others and individual experimentation. Cultural evolution happens in many species, but the infrastructure we have created to facilitate cultural knowledge transmission enables humans to be uniquely successful in creating complex societies on a sufficient scale to generate global impacts and user humanity into the Anthropocene (Henrich, 2020).

In early history, people learned in informal ways from family and community members to hunt, make artifacts, to prepare food, understand geography, identify poisonous plants, etc. During the last few hundreds of years, we have created educational systems that employed professionals to teach complex skills like reading, writing and arithmetic. Learning takes place during the whole lifetime, but tends to be concentrated in childhood. In contemporary societies children go to school from kindergarten to university and receive standardized education to prepare them for participation in the workforce.

There are various problems with creating those formal educational infrastructures. The schooling system we use today had its roots in early industrialization when there was a need to train a large population basic skills to participate in production systems. One may wonder whether those education factories are still adequate for contemporary times, where it might be more important that people learn how to learn during their whole life with constantly changing technologies. As demonstrated during the COVID-19 pandemic, schools function not just as a place to learn skills, but also as a daycare facility for working parents, and a place for children to socialize. In contemporary society where family members often do not live closeby, schools take over tasks which were done by the community historically. We also have to recognize the increased amount of mental health problems among children due to the pandemic, which may indicate that contemporary schools may not provide the support structure needed in times of crisis.

Human infrastructure needs to be maintained and continuously updated, especially in contemporary societies where we get confronted with new technologies and bureaucratic procedures. Nowadays one needs to be computer literate to manage your bank account, pay your taxes, etc. In a previous generation, you had to go to a physical bank and send a form via snail mail. The rapid advances of technologies causes older generations to be left out if they are not able to maintain their human infrastructure.

Skills that were taught in schools in previous generations might not be the best preparation for the next generation. With calculators, artificial intelligence, google, and other technologies, memorizing historical events or doing comprehensive arithmetic by hand becomes of questionable value. But how to use those new technologies in a proper way, and to recognize “bullshit” generated by those new technologies becomes essential.

Does the current coupled infrastructure system that is enabling the investment in human infrastructure well functioning? As we have seen with hard and natural infrastructure, there is a lot of standardization to scale up investments in human infrastructure. On the one hand, children around the world learn the same basics of arithmetic, language, science and art, enabling them to participate in the global economy. On the other hand, we may lose diversity of human knowledge and skills, to produce novel solutions for the challenges we face at the local and global level.

However, new technologies (computer based instruction) allow us to access individualized instruction to learn new languages, play guitar, and computational thinking. The provision of human infrastructure is in flux, and it is unclear what is the desired configuration for the future may be, and whether we are moving in the right direction.

13.3 Building Relations

Social infrastructure relates to the relationships people have. To build and maintain relationships takes time. Even though now you can have hundreds of "friends" on Facebook, actual meaningful social connections require investment of time, not just a click on a button. Increasingly physical places to socialize, whether they are bars, sport clubs, churches or other physical community building infrastructure, are being replaced by virtual alternatives. Where in the past people met their future partners at churches, dance halls, or bars, one is now swiping left and right to make connections. This might be more efficient, but it impacts the kind of social relationships we may build up.

Humans are social beings and we are wired the way we are by a long cultural and physical evolution going back millions of years. Technological advances and institutional changes allow us to interact globally in real-time, but that might not replace traditional ways of making and maintaining social connections. As demonstrated by the COVID-19 pandemic, the lack of face-to-face interaction has caused substantial mental health challenges. Zoom meetings allow for information exchange, but might be limited in other ways.

The benefits and opportunities that accrue to individuals from social infrastructure are becoming increasingly privatized. In past times, individuals relied more heavily on social connections. It took a village to raise a child, but the village also provided a safety net for care and food sharing. Nowadays, we purchase insurance to outsource the role of social connections. As such, there may be fewer incentives to maintain all those social connections. However this comes at a (social) cost. In a study we did in rice producing villages in China, Nepal, Thailand and Colombia we found that in communities more integrated with the market economy, there was less cooperation in public good dilemmas, especially when we included uncertain outcomes from the public good (Cardenas et al., 2017).

The privatization of benefits and opportunities associated with social infrastructure may also increase inequality. Those with financial resources can purchase a better safety net for child care, health care, elderly care, food sharing, etc., than those who have limited financial resources. This may lead to a reduced level of reciprocity from those with more resources to help those without many resources. Governmental policies may avoid some negative consequences by providing public schools, public health and social security programs. Nevertheless, the social dependencies have been reduced in modern society. This may limit our ability to respond to the sustainability challenges in our societies.

Within a globalizing world, investing in social infrastructure comes with additional challenges. With modern hard infrastructure, we get information from all around the world and can communicate with people in real-time who are physically faraway. Those diverse experiences enrich our lives, but not everyone speaks the same language, has the same religion, social norms, is in the same time zones, experiences the same weather, etc. Those differences increase the transaction costs associated with communication. Still there are only 24 hours in a day. Will we invest our time in social infrastructure with our physically local communities, or in our virtual global communities? Given the limited bandwidth we have in investing in social infrastructure, even with effective hard infrastructure, there are tradeoffs and consequences of the choices we make.

In the extreme, we can live physically among strangers, but engage via our hard infrastructure with people all over the world in personal and professional ways. In those circumstances we may have a lot of knowledge about the situations in other countries and cultures, but limited engagement with the physical location one lives in. On the other hand, one can focus investments in social infrastructure locally, and not know what is happening in the rest of the world. But in a globalized world, local economies are impacted by global markets.

Ideally, a community consists of a diverse population where some are more connected with the global community than others, and those diverse connections could help communities to have the appropriate knowledge and connections to flourish in a global economy. Unfortunately, we see frequent polarization between members of local and global communities, having underinvested in the different types of social infrastructure to be an active community member.

13.4 From shared norms to bureaucracies

In small scale communities, soft infrastructure relies on shared norms that are reinforced via shared narratives, religious rituals, and other types of rituals. There is no need in such communities to write up contracts. With the increasing size of societies, shared norms get formalized into written rules. Laws, contracts, and policy documents become increasingly complicated to anticipate or respond to the many situations and conflicts that could occur in large scale societies. This leads to specialized human infrastructure (lawyers, politicians), who may not necessarily may not share the experiences or have a deep understanding of the people they represent.

Archaeologist Joseph Tainter (1988) studied the collapse of societies and found that as societies become increasingly complex, they spend an increasing share of their resources to maintain the functions of this complexity. This may eventually become too taxing for a complex society and lead to its downfall. Tainter was especially influenced by the demise of the Roman Empire that became so big that it became too costly to maintain it’s borders. Increasing complexity in a society refers to more bureaucracy, and more defense and policing to maintain order.

We do not want to suggest a collapse of contemporary societies but the insights from Tainter could provide some lessons for our times. The success of our societies depends on the complex coupled infrastructure system we have built. We depend on physical infrastructure that brings us water, energy, information, and removes our waste. The schooling systems that prepare us to effectively function in the complex societies we have created, and the health systems that keep us healthy and allow us to have a longer expected lifetime are critical for our wellbeing. Clearly, the critical importance of these infrastructures suggests that planning and investing in all of them is worthy of our increased attention.

The increasingly complex coupled infrastructure systems that we are building require that humanity keep up the investments in human, and social, and soft infrastructure so that all humans have the proper knowledge to participate, the social connections to reciprocate efforts if needed, and the institutional arrangements that enable us to coordinate our activities and that fit the context of those activities. The increasing polarization and distrust in government we observe around the world might be signs that many people do not feel engaged with the society they live in. This is unfortunate since, for the challenges humanity is facing, we need cooperation and engagement.

13.5 Critical reflections

Humans have a long history of living in small scale societies that structure the way we interact with each other. We rely on building trusting relationships within local communities, and are good at identifying people who are different from us. This tribal behavior was important in ancient times but in a globalizing world they hinder us. We increasingly have to rely on hard infrastructure (communication technology) and formal regulations, to interact with an increasingly diverse community. Despite all the advances in technology, there is still limited time to invest in social relationships and, as such, we may struggle with building and maintaining non-physical infrastructure in contemporary societies. This might be the Achilles heel that may hinder or prevent potential transitions to a sustainable society.

13.6 Make yourself think

1. How much do you invest in the social infrastructure of local and global communities? Do you use social media for those investments?
2. Did you read all the details of the syllabus, which contain the many rules and regulations universities ask instructors put into the syllabus?

13.7 References

Cárdenas, J.C., M.A. Janssen, M. Ale, R. Bastakoti, A.M. Bernal, J. Chalermphol, Y. Gong, H.C. Shin, G. Shivakoti, Y. Wang, J.M. Anderies (2017) Fragility of the provision of local public goods to private and collective risks, *Proceedings of the National Academy of Sciences USA* 114(5): 921-925.

Henrich, J. (2020). *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. Farrar, Straus and Giroux.

Tainter, J. (1988). *The Collapse of Complex Societies*. Cambridge University Press

PART VI

PART 6: INFRASTRUCTURE IN TRANSITION

CHAPTER 14

Transition to a Zero-Carbon Society

Key Concepts

In this chapter we will:

- Explore a potential transition to net zero-carbon society
- Identify human, social, and soft infrastructures as the bottlenecks

14.1 Introduction

Since the late 1980s, climate change has been on the political agenda in many countries. The New York Times headlined on June 24 1988 that “Global Warming Has Begun, Experts Tell Senate”, Times magazine awarded Earth the planet of the year in 1989. The Intergovernmental Panel on Climate Change (IPCC) started in 1988 and published their first reports in 1990.

By that time the science was well enough established to know that action needed to be taken. If we look at the scenarios published in 1990, they explored different pathways to transition to a less carbon intensive economy (Figure 14.1). Only the Accelerated Policies Scenario would result in a global temperature increase of less than 2 degrees Celsius by 2100.

We have also included in Figure 14.1 the actual historical emissions, and it is remarkable that those are just above even the high emission scenario of the IPCC! It is remarkable since by the 1990s there was already sufficient knowledge about climate change to know that emissions had to be reduced significantly, and the issue had gained international political attention. Despite good intentions, emissions have grown more than anticipated for a even the worst case scenario.

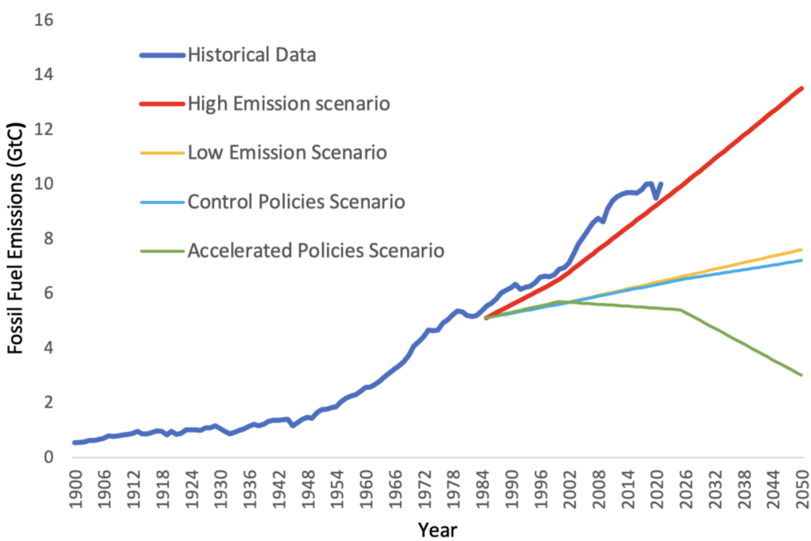

Figure 14.1: Historical emissions from 1990 until 2021 compared with the four IPCC scenarios published in 1990 for global fossil fuel emissions.

At the moment of writing this chapter several decades after these events, there is now a lot of attention being paid to a transition to a net zero carbon society. We will discuss the possibilities and challenges for such a scenario from a coupled infrastructure systems perspective. The reason for such a transition is to avoid global temperature rising to more than 1.5 degrees Celsius, which is considered the tipping point to irreversible effects from global warming such as ice sheet and permafrost melting and Amazon rainforest destabilization.

In Figure 14.2 we show the scenario for fossil fuel emissions that is needed to avoid dangerous levels of climate change. Since it is recognized that those emission scenarios cannot be reached by emission reductions only, current scenarios also include options to extract CO2 from the atmosphere by large scale tree planting or carbon capture and storage of actual emissions, for example from bio fuel based energy generation. This is why we discuss net emissions where one takes into account both the emissions and the negative emission technologies to counter the actual emissions.

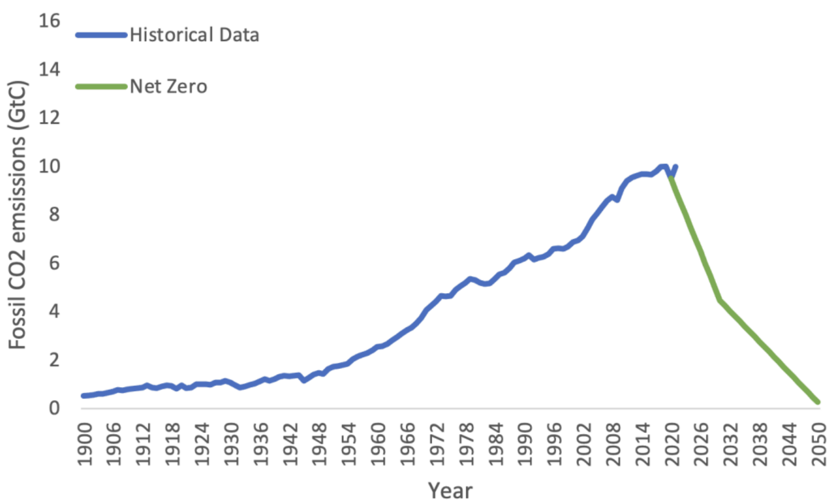

Figure 14.2: Emission scenario for the period 2020 to 2050 to avoid a global temperature rise of more than 1.5 degrees Celsius compared to historical emission trajectory. In comparison to Figure 1, the required decarbonization scenario is far steeper than if we had taken action earlier.

14.2 The problem of scale

One of the ultimate problems faced by governance bodies is the overuse of the carbon assimilation capacity of the atmosphere and the resulting global climate change. There are many uncertainties regarding the specifics of the consequences, and experienced weather changes people experience will likely be quite diverse and vary geographically. To reduce the extent of climate change we need to reduce emissions of greenhouse gasses substantially. Since the gasses remain in the atmosphere for years, it does not matter where emissions have taken place. Therefore, to have a measurable impact, emissions need to be reduced at a global level.

Since the early 1990s, several international negotiations have taken place to develop agreements to reduce emissions. In 1992, the United Nations Framework Convention on Climate Change was created, which had as its aim to "stabilize greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system." In 1992, CO2 concentration was 356 ppmv. At the time of the writing of this chapter, the concentration has increased to 419 ppmv. In fact, every few years new treaties are signed with ambitious plans to reduce emissions, but none have had a measurable impact.

Why is the problem of climate change so difficult to solve? Emission reduction at the global level requires that we not use cheap available fossil fuels such as oil, coal, and natural gas but, rather, use alternative energy sources and reduce our overall energy use. This will require major technological innovations and behavior change, which may affect both economic growth and our well-being. Should a country like the U.S. that has historically emitted the most greenhouse gasses make a bigger contribution to the solution compared to other countries? Will we allow countries like India and China to grow their emissions in the coming years since their emissions are low in per capita terms compared to other developed nations and because these countries are relatively less developed? On what ethical basis could we prohibit some Chinese families from enjoying a private car when such is allowed in other countries? Countries like the Maldives will disappear in the coming decades due to the rising sea level and other consequences of climatic change. Who will take care of the climate refugees? What will happen if countries do not do what they promised? Can we enforce the rules?

As you can see there is no simple solution to the greenhouse gas emission problem. Instead of giving up and continuing with business as usual, we can explore the problem from a polycentric perspective. There are many actors around the world who want to make a difference. There are states, cities, universities, and towns that have committed themselves to reducing greenhouse gas emissions. For example C40 is a network of megacities who are committed to implementing practical solutions to reduce emissions and create a sustainable future for their citizens. Many of these local actors are driven by the fact that their efforts to reduce emissions also contribute to solving other problems such as local air pollution and the cost of energy use.

Such "climate clubs" have the benefit of starting with actors who are motivated. Participants who want to join opt in and must agree with the rules of the club. Clubs can then set examples of solutions and exchange lessons learned from their local attempts to implement solutions.

14.3 Transition of infrastructure

Figure 14.2 shows a dramatic change in historical trends and illustrates what might be possible and what challenges we face. A transition to a low carbon economy will require fundamental changes in our society in all factors of life, from our diet to our transportation, housing and occupations. This will require major challenges in different types of infrastructure. In this section we will explore for each type of infrastructure some of those challenges and opportunities.

Hard infrastructure:

Hard infrastructure relates to greenhouse gas emissions via the generation of energy, and cement, the technology to grow and process food, the physical traffic infrastructure, etc. There are many new technological options available to reduce our carbon footprint varying from non-carbon energy production systems like solar, wind and biofuels, electric transportation, meat alternatives, retrofitting housing to improve insulation, low carbon options to produce cement, etc.

A major focus in recent decades has been to find technological solutions to reduce emissions, but this alone will not be sufficient to achieve a transition. Our current physical infrastructure of roads, energy and water depends on decisions made more than 100 years ago, such as the standardization of currents we use to transport electricity. Although low carbon alternatives are available, a fundamental change in the physical infrastructure requires new ways of approaching infrastructure such as using micro-grids, electric vehicles, alternative low-water use sanitation options, etc. Such a transition requires a behavioral and institutional change too.

Even if we adopt a lot of these new technologies, each solution has consequences too. Electric vehicles use energy, which could, in fact, be fossil fuels if fossil fuels are used to generate electricity. Electric vehicles use a lot of minerals and potentially polluting batteries. Wind energy impacts bird populations, hydroelectric power impacts water availability and displaces peoples, and going vegan will still generate CO2 emissions. Whatever technological solution we adopt, it will not be sufficient to address the bigger challenge, our high consumption lifestyles that are increasingly being adopted by a growing global population.

Soft infrastructure:

To stimulate adoption of new technologies and behavioral change among households, businesses and organizations, new regulations and incentives are needed. Often policies are focused on setting standards or provide price incentives (subsidies and taxes). Price incentives can be implemented quickly, but raising taxes will not get you votes and might be politically a challenge. Changing standards is a long-term challenge and the business sector would provide objections to the costs such changes will bring with it. For example, a change in fuel standards for cars requires car companies who do not meet the new standards to adjust their product and this leads to extra costs.

Perhaps the main challenge in changing the software infrastructure to address CO2 emission reductions is the need for an international coordinated effort. Although the need for a transition is acknowledged in international treaties and promises are made to reduce emissions, in practice nations generally do not comply with those proposed intentions. The main reason is that there is no accepted authority who could enforce international agreements. Most nations are sovereign, meaning that there is one centralized government that has the power to govern a specific geographic area. International enforcement of agreements will conflict with this sovereignty and thus international agreements are basically voluntary commitments. Given the commons dilemma nature of emission reductions, there is reluctance to change societies to become less materialistic.

Natural infrastructure:

Natural infrastructure will be impacted in different ways in a potential transition. Planting trees could contribute to carbon storage, and thus would be a way to generate negative emissions to meet the policy goals. Other types of farming that improves soil health & soil carbon storage via regenerative grazing, could make a contribution to negative carbon emissions.

The main challenge for the natural infrastructure is to deal with the unavoidable changes that are going to happen already and continue to happen, such extreme weather events, forest fires, sea level rise, droughts, and a reduction of biodiversity. To maintain desirable natural infrastructure functions, adaptation will be needed. This could mean changes in crop varieties, to cope with less water, improve reforestation to improve resilience of soils to rainfall events, and improve conservation corridors to improve resilience of species for changing habitats.

Human infrastructure:

Do we know what actions to take? Do people have sufficient knowledge about the climate crisis and do they know what actions to take? A large portion of the population does not identify climate change as real or something to worry about (Leiserowitz et al., 2021). Recommended behavioral changes decrease motivation of personal behavioral change (Palm et al., 2020). This is unfortunate since Dietz et al. (2006) demonstrated that a direct emission reduction of 7% could have been achieved without significant impact on human welfare. Even if people are intend to change their behavior, they may not do so. Think about the many New Year resolutions of doing more exercise and losing weight that are not met.

Knowledge about the problem is not the same as accepting the majority view and the consequences for behavioral change. Perhaps this is caused by the implications of behavioral change for people's lifestyle as well as the ease of ignoring the climate change problem, "Après nous, le déluge" (French for "After me, the deluge" meaning "Ruin, if you like, when we are dead and gone")

Not only do we need changes in our consumer behavior, we also will need a change in skills needed to facilitate the transition.

There is a demand for workforce in non-fossil fuel energy systems, alternative food systems, better insulated houses and changing transportation systems. To have this workforce available in the coming years, those individuals need to be trained now. We will also change the insurance system, since probabilities of weather related events are changing, preventing the insurance industry from having a sustainable business model by continuing business as usual.

Social infrastructure:

An important constraint in deriving a behavioral change are existing social norms about lifestyle, the role of government, science and the existence of climate change. If an individual would like to adopt a carbon neutral lifestyle, will this be supported by their friends, family and community? Social pressures to conform to social norms can also be an accelerator for change as happened with the reduction of smoking cigarettes (Nyborg et al. (2016).

To derive a fundamental socio-economic transition might be disruptive to communities and the way a behavioral change is stimulated is critical. In individualistic cultures like the USA, imposing behavioral change can backfire (Palm et al., 2020), while in collectivist cultures such as China, a top down persuasion could be effective in changing social norms.

Perhaps the biggest challenge is the huge demand for coordination and conflict resolution at different levels to implement a transition. A transition will lead to changes in locations where what will be done, and thus require changes in zoning and permits. This typically requires local inputs, and generates delays in the implementation of projects. So many changes will be needed in the way we live, work, organize, and produce, it is difficult to imagine how this would be possible without conflicts in society. The COVID-19 pandemic provided an example of a relatively simple problem, an infectious disease that could impact everyone's life, and the disruptive period to cope with the pandemic. Disagreements on how to prevent the spread, to implement a solution via vaccinations, are just initial signs that a just transition to a net zero carbon economy could expect many challenges once implementations will reveal different perspectives at the local and national levels.

14.4 Challenges to changing infrastructure

Change in hard infrastructure is slow. For example, bridges and sewer lines are designed to last 100 years or more, and coal plants and nuclear power plants more than 35 years. A natural replacement of infrastructure will therefore last decades. A more rapid transition will require the destruction and retrofitting of existing infrastructure.

All types of infrastructure are characterized by the build up and maintenance of the capacity of the infrastructure. Since there are limited resources, trade-offs have to be made which efforts to put in which kind of infrastructure. So far, governments at multiple levels have invested in changes in potential new hard infrastructure, but have not invested in creating momentum to implement a transformation of the actual infrastructure. To do this will require aligning a sufficient proportion of humanity behind the goals of a transformation in order for humanity to invest an enormous effort to coordinate such a transition.

A transformation in hard infrastructure requires the coordination on designs where and when infrastructure is to be placed, what new standards to adopt, and how to train the workforce to create, use and maintain new infrastructure. In democratic societies this would lead to negotiations between different special interest groups on zoning conflicts, standards and curricula. Even if people are motivated, it is a humongous task.

The current polarization in society and ineffective national governments, as highlighted by the response to the recent

COVID-19 pandemic, indicates we may have underinvested especially in social infrastructure. The conundrum for a just transition is the need for a rapid change but this cannot happen in a way that gives everyone a voice and people will accept the changes that are proposed to be implemented.

Although we should strive for decarbonization of our society, we have to prepare for a society with dangerous climate change. This may require hard and natural infrastructure to be able to cope with more weather extremes, humans and communities to learn to cope with living in a different climate with more heat waves and forest fires, and adjust insurance and migration policies to cope with the consequences of climate change.

14.5 Critical reflections

In this chapter we explore the desired transition to a net zero carbon economy. Although the need for this transition is widely recognized and technological solutions are available, it is difficult to see such a transition would be possible given the high demand on human, social and soft infrastructure.

14.6 Make yourself think

1. Do you know how your electricity company generates the electricity you are using? And does your electricity company plan to change their practices?

2. What need to happen to reduce the carbon footprint of your household, company and town?

14.7 References

Dietz,, T., G.T. Gardner, J. Gilligan, P.C. Stern and M.P. Vandenbergh (2006) Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions, *Proceedings of the National Academy of Sciences USA* 106 (44): 18452-18456.

Leiserowitz, A., C. Roser-Renouf, J. Marlon and E. Maibach (2021) Global Warming's Six Americas: a review and recommendations for climate change communication. *Current Opinion in Behavioral Sciences* 42, 97–103. https://doi.org/10.1016/j.cobeha.2021.04.007

Nyborg, K., J.M. Anderies, A. Dannenberg, T. Lindahl, C. Schill, M. Schlüter, W.N. Adger, K.J. Arrow, S. Barrett, S. Carpenter, et al. (2016) Social norms as solutions, *Science* 354: 42-43

Palm, R., T. Bolsen, and J.T. Kingsland (2020, "Don't Tell me what to Do": Resistance to Climate Change Messages Suggesting Behavior Changes, *Weather, Climate and Society* 12(4): 827-835, doi:10.1175/wcas-d-19-0141.1

Rogelj, J., O. Geden, A. Cowie & A. Reisinger (2021). Three ways to improve net-zero emissions targets. *Nature* 591(7850): 365-3.

CHAPTER 15

Potential futures for different types of infrastructure

Key Concepts

In this chapter we will:

- Explore various scenarios for the future of infrastructure using the CIS Framework

15.1 Introduction

In this chapter, we will explore potential futures for different types of infrastructure. The goal for doing this is to use the coupled infrastructure framework and reason for alternative futures using the perspectives from different actors into account. Since we live in uncertain times, the use of scenarios help us to explore the consequences of decisions we are facing.

15.2 The future of education

The first topic is human infrastructure. The main current approach to investing in human infrastructure is based on very old practices. Formal schools have been around for thousands of years and were found in China, Greece and the Roman Empire. The current emphasis of compulsory education goes back to Martin Luther who advocated that everyone needed to be able to read so that one could be able to read the Bible for themselves. Only around 1900 did compulsory education become common in various countries. The justification was to train children in the skills needed to become productive citizens of the nation. This coincided with the increased industrialization of society for which there was a need for trained workers who could read and write, do arithmetic, and execute tasks in an industrial production process. Until recently, only those from wealthy families could afford higher education.

The action situations of many schools are still based on the traditional notion that children are sorted according to age and under the leadership of a teacher who uses a standardized curriculum that all students in a state or nation need to follow. Typically outcomes are measured by written tests throughout the school term. This action situation does not take into account the diversity of students and the variety of skill sets needed by the nation in the future. The action situation selects for those students that perform well in doing the tests, and some of those students will become teachers for future generations. We, the authors of this book, are well aware of the problems associated with this action situation. Since we teach at Arizona State University, which likes to measure itself by what students achieve during their time in college instead of by whom we exclude, we have come across a very diverse population of students. Students do not always excel in the traditional system, but can blossom with additional accommodations. In fact, a more personalized pathway of learning seems to be appropriate instead of a standardized curriculum.

The current developments in education, especially via options of online learning, allows students to find relevant courses of their own interest which they can take at their own pace. The question is whether traditional schools and universities will continue to survive in the long term. At a minimum , if the focus of schools is knowledge transmission, the role of schools might soon become outdated. But schools also provide a function of investing in social infrastructure, allowing children and the broader community to build connections and organize social events.

How long will we continue to use traditional diplomas as measurements of human infrastructure? In a world that is constantly changing, people will need to constantly be trained to remain up to date. It is not unreasonable to assume that future educational systems will become more blended with other organizations, such as the workforce, and people's human infrastructure will be evaluated on their recent professional activities and their micro-credentials.

15.3 Health futures

The life expectancy in the early 1800s around the world was around 30 years, and this level has more than doubled since. The countries with lowest life expectancy can be found in sub-saharan Africa, which is still higher than 50 years, and some countries, like Japan, the life expectancy is well above 80 years. The reason for the sharp rise of life expectancy is the widespread use of sanitation, vaccination for various infectious diseases and improvements in medicine.

The focus of healthcare has been moved from only treating disease to increasing quality of life. This also comes at a price. Good health care is not accessible for all in many countries. Those with more financial resources will be able to access better health care. To improve the quality of life one will be able to prevent diseases by regularly taking blood samples to test for various common diseases, taking supplements, analyzing one's DNA and tracking one's activities on smart watches. Those who cannot afford those preventive measures, may take action when symptoms of diseases are experienced. Those with lower incomes typically have lower life expectancies since their lifestyles increase risk via higher levels of obesity (due to cheap processed food), more stress, and less sleep. The opioid crisis in the USA illustrates this problem where the life expectancy of poor white men has been declining in recent years.

The COVID-19 pandemic demonstrated that inequality in access to medication is a global phenomenon. When vaccines became available, rich countries obtained the majority of the vaccines for their own citizens, but the pandemic will only be eradicated if there is a good vaccination rate at the global level. We also have seen the inequality of impacts within countries between groups of people. Essential workers, typically those who do the hard labor as cleaners, drivers, factory workers, had to go to work and be more at risk of infection of diseases. They could not afford to skip work. Contrast this to white collar workers who often could work remotely from home and did not experience the risk of those with worse health insurance.

The increased amount of health information which could provide better estimates of particular health risks, also generates some risks. Could this information be used by insurance companies or employers to adjust rates or labor contracts? We already see companies with genetic information from those interested in their heritage collaborating with police to detect criminals from DNA at crime scenes. Who else will be able to use this information? How private will health status information remain? We may be able to match potential donors and recipients, but how much may less powerful individuals be coerced to donate their organs. We already see this happening with blood donations, as well as organ harvesting for kidneys in various countries in the global south.

15.4 How we live and work in the future

Humanity increasingly lives in cities, and the urbanization trend is expected to continue. By 2050, more than two-thirds of the world's human population will live in urban areas. Nevertheless, there are different potential futures to imagine how we may live and work in the future. On the one hand, we have sprawling cities of largely single family houses in car dependent cities, like Phoenix, Arizona. On the other hand, we have high density cities with compact apartments and decent public transportation, which are features of the majority of new cities around the world. This latter trend is likely to have a lower ecological footprint (less building material, lower emissions for transportation, more opportunities for recycling).

At the time of writing this chapter there is a housing crisis around the world. One of the key factors is the increasing ownership of residential houses by investment companies. After the 2008 housing crisis, those companies started buying up housing stock. Since those companies are focused on maximizing return on investments, rents are increasing rapidly in many cities around the world. Those companies have no incentive to provide affordable housing, so the affordable housing stock decreases and leads to an increase in homelessness.

The COVID-19 pandemic led to a change in the housing market too. Workers who can work from home, started looking for bigger houses in desired locations. Furthermore, supply chain problems led to a delay in home building. As a consequence, some areas, like Florida and Arizona within the USA, experienced overheated housing markets with annual rent increases of more than 20%. And due to the 2008 housing crisis, there are more restrictions for who may get a mortgage, increasing the power of investment companies who can pay cash. The increasing dominance of the financial sector in the housing market is a global trend impacting affordability (Marcuse and Madden, 2016; Rolnick, 2019).

We already have addressed how the nature of work and how we work may change due to different types of skills needed as well as the opportunity to work remotely. Moreover, due to the increasing development of productive artificial intelligence, many tasks (including white collar tasks) will become automated. This may lead to shorter working hours, but may also affect social infrastructure since having a professional life can be an important part of somebody's identity. Who makes decisions about what tasks will become automated and which jobs remain active? Will those individuals steering the process of automatization become much more wealthy and powerful? In the movie WALL-E [xxx], obese blobs, representing humans, were binge watching and eating in space ships. Will this be the destiny of humanity, or will people be able to do work outside the efficient automated globalized supply chain?

15.5 Earth-space sustainability

Increasingly we are dependent on satellite infrastructure, that is moving high above us at and at high speed. Communication networks, weather measurements, the GPS coordinates of your car and your destination on your map app, are all critically dependent on well functioning infrastructure in space. However, the space around the low Earth's orbit is unregulated. With decreasing costs of sending satellites to space, we see an unregulated increase of governmental and private organizations active in investing in space infrastructure.

Space is not empty, especially the low Earth's orbit. There is an increasing problem of space debris, which consists of (parts of) defunct satellites. Due to the so-called Kessler Syndrome, there is a real risk that low Earth orbit will become inaccessible. The Kessler Syndrome refers to a situation where the density of objects in low Earth orbit is high enough such that collisions between objects could cause a cascade in which each collisions generates space debris that increases the likelihood of further collisions.

It isn't a far stretch of the imagination to see that space debris is analogous to air pollution. For a sustainable future on Earth, we will have to address the space debris problem because we rely increasingly heavily on space infrastructure to conduct global measurements of the state of the environment and for communication via satellites to implement solutions (imagine a Tesla that cannot be connected to the internet).

There are other factors that connect our desired transition to a sustainable future with outer space. There is a rapid increase in demand for lithium, nickel, cobalt and other rare minerals, needed for the creation of batteries, smart phones, and the like. Those rare minerals are located in a limited number of locations, and those limited resources are being exploited rapidly. To be able to make a transition to a sustainable future, we need to improve the capacity for recycling those minerals or get those minerals from somewhere else. This is one of the reasons behind the current space race. Some of those rare minerals can be found in asteroids, the Moon and Mars, and the space industry is developing knowledge and technology to extract these minerals. Like the low Earth orbit, there is no enforceable regulation for outer space mining. But that leads to questions about fairness. Which actors will be able and allowed to extract rare minerals from space, and how will the benefits be distributed among actors on Earth? Will there be regulations on how space mining takes place, especially since other actors may use those celestial objects for research purposes?

15.6 Critical reflections

In the exploration of potential future trends of four types of infrastructure, we see that conflict about access is a key shared concern. Although technologies become available which could provide better access to educational opportunities, health care, living sustainably off the grid using rare minerals from space, how are those opportunities being distributed? Will this be a small elite while the majority will not have access to those affordances and live to extract resources for the happy few. That would not be a global sustainable future. But if we do not allow individuals taking risks to advance exploration, will we derive those technological innovations? Will national governments step in?

15.7 Make yourself think

1. What are utopia and dystopian futures you can envision for the next 50 years?

15.8 References

Marcuse, P., and D. Madden (2016) *In defense of housing, The politics of crises*, Verso

Rolnick, R. (2019) *Urban Warfare: Housing under the Empire of Finance*, Verso

PART VII

PART 7: HOW TO APPLY COUPLED INFRASTRUCTURE SYSTEMS THINKING IN PRACTICE?

CHAPTER 16

A Protocol for Practice

Key Concepts

In this chapter, we will:

- Provide a number of examples how to use the lessons learned into practice
- Discuss how to avoid blueprint thinking and stimulate diagnostic inquiry.

16.1 Introduction

There is no simple solution to the many challenges we face to transition towards a sustainable future. There is no silver bullet, a blueprint or a magic pill. The reality is that it will be a struggle, like governance has always been a struggle. We have to try new solutions and initiatives. Some may succeed, and many will fail.

What can we do? How can we be prepared? We can learn from various other domains of expertise that cope with difficult problems. Let's have a look at medicine. Physicians are trained to have a protocol of systematic testing to diagnose a health problem: what is the most common disease, what is the most severe disease possible with those symptoms. A good physician will not give you a standard pill based on ideology but will do a series of tests. Based on those tests certain diseases will be eliminated, and at a certain point a diagnosis is made. The physician might not be sure, but based on their experience may have confidence on how best to treat the patient. The physician informs the patient about the diagnosis and the patient can consent to the treatment. The patient can ask for a second opinion and may get different advice with another physician. When symptoms disappear and it seems that the patient has recovered, we do not go back to verify it was a correct diagnosis, and the patient was properly treated.

For governance we need to improve our diagnostic capacity and a proper diagnosis should lead to proposed interventions, not ideology. In this chapter we will provide some approaches that could help with the diagnosis of the problem, and the crafting of solutions. The selection of approaches is based on the insights from the material we have presented in this book:

- Humans are cooperative beings who tend to cooperate with people they trust.
- Humans have a long history of living in small communities.
- Communities are able to create effective institutional arrangements that are aligned with the social and biophysical context.
- But it is not the institutional arrangements themselves that make those rules effective. It is because people have been involved in the crafting of those rules themselves, accept the process and implement monitoring and enforcement.
- In a globalizing world, those small scale community findings can partly be scaled up. Key is to nurture participatory processes.

16.2 Participatory approaches

Workshops with stakeholders

When working with stakeholders within a community on a governance challenge of a coupled infrastructure system, one could organize meetings with the stakeholders. Those meetings have different formats. It would be helpful to first scope the action situation of the governance challenge and try to identify the relevant stakeholders from policy documents or by meeting relevant stakeholders individually and try to understand the action situation.

Try to understand the incentive structures facing the different stakeholders. It is natural to focus on one or a few stakeholders, perhaps the one most in line with your desired outcomes, but that will not be helpful for the process. There are likely good reasons why some stakeholders make decisions that are harmful for other stakeholders, and if you want to help to improve the governance solution, you need to understand the logic for the various actors in the action situations.

Some stakeholders might not be very organized via (elected) representatives. In these circumstances you may organize focus group discussions with a number of community members from the same stakeholder group. This may help to understand the diversity of perspectives within the same stakeholder group. Perhaps not all residents in a neighborhood experience the same water availability challenges, or have different information about a campaign the government has rolled out.

If you organize workshops for a diverse set of stakeholders, you may use a professional mediator. At least make sure that there is someone who can keep the meeting on schedule, and moderate potential conflicts between stakeholders. Organize activities, like ice breaker questions, that enable people to socialize. Later in this chapter we discuss activities like scenario planning, game play and model exploration. Workshops are themselves action situations about which there is much research on factors that lead to more effective collaboration and more effective deliberative processes.

It is not unusual if some stakeholder groups do not show up. This could be done for strategic reasons (they could dismiss the outcomes of the workshops since they were not present). You could follow up with those stakeholder groups by individual meetings, to debrief about what happened during the workshops.

How to use design principles

Ostrom proposed a list of design principles, qualitative patterns typically seen in long-term successful communities managing shared resources and shared infrastructures. Ostrom regretted later that she used the term "design principles" instead of "hypotheses". The reason for this is that many organizations have been using the design principles as a blueprint for what institutional arrangements need to be implemented or a checklist that needs to be followed to ensure good governance. This is an incorrect approach to using insights on the design principles. In fact, going back to case study analysis, we find that there are case studies that are coded to be successful without having met all design principles (Baggio et al., 2016; Barnett et al., 2016). This is possible because of the social and biophysical context that make some design principles irrelevant (for example, isolated resources which have natural protection from costly access).

So how should we use the design principles? In chapter 6 we already discussed a number of questions that could be asked by the community based on the design principles. This will help to get a better understanding of the relevant action situations they participate in. Asking communities to address these questions will help identify underlying processes and diagnose the collective action problem.

Scenario development

One activity that could be done in a stakeholder workshop setting is scenario development. What are 3 to 5 possible scenarios of the future of the coupled infrastructure system of interest? What is a scenario where the future is a continuation of the past, and what is an alternative pathway to very different outcomes? To be productive, the workshop participants should explore what those scenarios mean or imply for the different stakeholders. Who are the winners and losers? What bottlenecks do these scenarios need to overcome? Scenarios are not predictions. The goal is not to be correct, but to explore alternative futures which may reveal potential conflicts between stakeholders or new opportunities no one had thought of before.

To prepare for a scenario development exercise, it is helpful to have historical trends of the past available in a visually engaging way. This may help identify the boundaries of expected and potential changes. Participants may develop better insight into the conflicts and alignments of activities they do or desire. Closing coal plants is a good idea to reduce CO2 emissions, but what future is there for those people in those coal plant industries? A reduction of biodiversity may seem a distant problem, but how will one cope with a lack of pollinators impacting agricultural production?

Using games

Games are commonly used as an activity to provide an active learning environment to understand complex concepts. Games can be individual or with multiple players. Since we focus on workshops with stakeholders, we focus on multi-player games. Games can address general issues, or are tailored to the specific case of the community. A widely used example of a general game is Fishbanks in which players represent fishers and need to make decisions which boats to buy and where to fish. A common outcome is that without coordination, this leads to overfishing of the shared resources. It is actually a good learning experience for a group to end up in a failure of resource governance. Debriefing after the game what happened and why individuals may reveal different strategies and assumptions.

In recent years, games have started becoming used as intervention tools. For example, in rural India, NGOs are using a groundwater game with communities who are experiencing extreme ground water depletion. The NGOs have been working with those communities for a number of years and had engineers coming to villages to discuss technical solutions, or teach water budgeting, but that had no effect on the day to day experience of the villages. But using the groundwater game led to a measurable change in engagement in the communities to find solutions.

In the groundwater game, players make a decision each round which crop to plant. One crop leads to more monetary outcomes, but also uses more water, while the other crop is less profitable and uses less water. If the players all use the water intensive crop this will lead to a complete depletion of the shared resource in 5 rounds. Better cooperation and coordination, where not all players use the water intensive crops each round, lead to better long term outcomes. After the game, there is a community wide debriefing where the results are discussed, and connections are made with water governance in the village. The moderators of the game do not impose a particular solution, but empower villagers with insights and an exploration tool to craft their own solutions.

Using models

Another common tools used in stakeholder workshops are computational models, such as system dynamics and agent-based models, that could be used to explore quantitatively the consequences of different scenarios. Models are by definition simplifications of reality and will thus not include all the details of the coupled infrastructure system under consideration. This can lead sometimes to tensions since some stakeholders may not trust a model if it does not include certain details. It is therefore important to have stakeholders involved in the design of the model from the start, so that they know what the model is about and why certain assumptions were made.

Once a model is designed and implemented, it can be used to implement different scenarios. This provides quantitative outcomes of the qualitative scenarios and can help provide additional ways to visualize future scenarios. One will get confronted with choices to be made, e.g. what is the expected level of investment in solar energy, and the results of those choices. Those scenarios provide a way to concretely visualize the different outcomes stakeholders may envision in their minds.

16.3 Critical reflections

To implement the lessons learned in this course in practice, a key takeaway is to work with stakeholders, typically at the local level. There is no guarantee for success, and the process can be slow and can be a struggle. It will be important to provide activities for stakeholders to empower them with skills to craft and implement their own solutions. Don't impose solutions, which are likely not to be adopted because they are imposed, but develop activities to co-create solutions.

We focus our lessons learned on local level activities. Even when we want to address global issues like climate change, we will have to make changes at the local level.

16.4 Make yourself think

1. What would be a problem in your community you would like to help to address? Who are the main stakeholders, and what are the main bottlenecks to solve this problem?

16.5 References

Barnett, A.J., J.A. Baggio, H.C. Shin, D.J. Yu, I. Perez-Ibarra, C.A. Rubiños, U. Brady, E. Ratajczyk, N. Rollins, R. Aggarwal, J.M. Anderies, M.A. Janssen (2016), An iterative approach to case study analysis: insights from qualitative analysis of quantitative inconsistencies. *International Journal of the Commons* 10(2): 467-494.

Meinzen-Dick, R., M.A. Janssen, S. Kandikuppa, R. Chaturvedi, K. Rao, and S. Theis (2018), Playing Games to Save Water: Collective Action Games for Groundwater Management in India, *World Development* 107: 40-53.

CHAPTER 17

Challenges Ahead

Key Concepts

In this chapter, we will:

- See a summary of the lessons learned from this book
- Learn why we still cannot solve all collective action problems
- Get exposed to the big challenges we are still facing in governing coupled infrastructure systems

17.1 What have we learned?

This book has provided an introduction to the study of institutions and governance in general and of governance of the coupled infrastructure systems in particular. Coupled infrastructure systems face the problem of under-provision, under-maintenance and over-extraction of the affordances. Despite the difficult challenges associated with governing these systems, we see successful performance of many shared resources especially in small-scale communities. We need to extend the lessons learned from these successes to better understand the general properties of different approaches to successfully governing the coupled infrastructure systems.

Elinor Ostrom developed a coherent theoretical framework that enables scholars to clearly articulate how institutional arrangements can facilitate successful governance of the commons (shared resources and infrastructures) . By institutions we refer to the prescriptions that humans use to organize all forms of repetitive and structured interactions. The prescriptions are rules and norms. They apply not only to common-pool resources such as groundwater, but also to other types of social dilemma situations like traffic, outer space, digital commons, sports, and health care.

Rules can be written laws, or agreed upon and commonly understood verbal rules in a community. Norms do not include explicit consequences if forbidden activities are performed or requirements are not met. Even though it may not be explicit, not following social norms may have negative consequences since people may decide to avoid interacting with people who have bad reputations.

A key concept in studying institutions is the action arena. An action arena consists of people as participants and an action situation in which they participate. When people interact in an action situation, they make decisions based on the choice rules associated with the position they occupy in that action situation. In a given action situation, people may hold different positions and therefore may not be able to make the same decisions, or have the same information. The interactions of the participants lead to outcomes that can be evaluated.

Figure 17.1 shows the schematic representation of the Institutional Analysis and Development (IAD) framework and highlights the key components necessary for studying how institutions structure action situations. The IAD framework emphasizes the fact that action situations are influenced by broader contextual variables. The biophysical conditions—whether you live in a desert or a rainforest—affect rules and norms concerning how to build houses and how to organize health care (e.g., due to different diseases that are prevalent in a given area). The attributes of a community such as the age and income distributions, education, and kin-relationships, affect which kind of interactions one can expect in action situations.

The rules-in-use are one of the key foci of the IAD framework. Rules on paper are important, but if those rules are not known, understood, and accepted by participants in the action situation, they will not effectively guide behavior. In studying the governance of the commons, we are interested in which rules people actually use, how they monitor rule compliance, sanction rule infractions, and how contextual variables impact how the rules function.

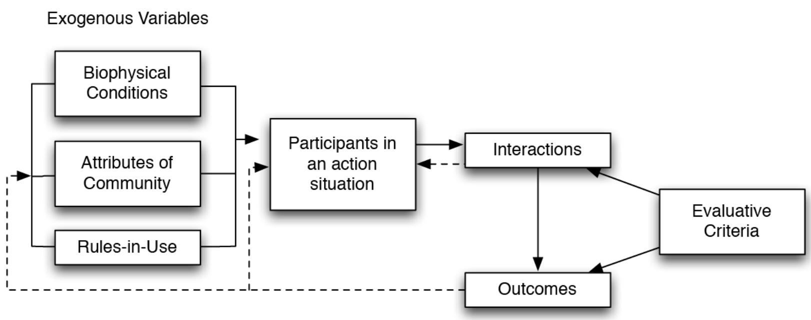

Figure 17.1: A framework for institutional analysis (adapted from Ostrom, 2005).

We have illustrated the application of the framework through several examples. The framework is just that—a framework. Frameworks are an articulation of key elements that should be considered when trying to understand the impact of institutional arrangements on human behavior and social interactions. The framework provides a set of concepts and language that enables scholars to communicate effectively about the key working parts of an action situation. Thus, if a student has developed a working knowledge of the IAD framework, they should be able to translate observations of social phenomena into the language of the IAD

Framework and action situations. This process of translating phenomena into a formal language enables us to compare different cases and uncover regularities.

It is important to understand that the IAD framework is not a theory or a model. It does not suggest a hypothesis about how different parts of action situations relate to outcomes. Theory relevant to understanding social phenomena is an additional layer related to how people make decisions in different action situations.

Much of the discussion and the majority of examples focused on a particular context that we call coupled infrastructure systems. In those situations there are incentives for individuals to free ride on the cooperative actions of others. We experience social dilemmas in many action situations in our daily lives. For example, who is doing all the work in a group project, how do we pay for the highways we use, how do we make sure there is healthcare available when we need it, who writes the articles on Wikipedia, are our bridges being inspected for safety, who reduces their energy use to help reduce pollution?

How do we organize incentives such that we reduce free-riding in problems associated with issues we care about? One option is to use coercion. If people have a tendency to free-ride on the cooperative behavior of others, then privatization of common-pool resources and public goods is an option. The reasoning is that individuals will make better decisions regarding the use of private goods. We illustrated the problems associated with common-pool resources with an example in which multiple people share a meadow. Each individual has an incentive to add animals to the meadow and when everyone in the group does so, this will lead to overgrazing. If, on the other hand, everybody owns a part of the meadow, everybody will take care of their own property and won't damage others' property. Another policy might be to tax the use of resources, so that people will not overuse common-pool resources.

Both of these economic instruments (privatization and taxation) are used in managing shared resources in practice. However, these instruments face several practical limitations and are not the only options available. There are many examples of self-governance, meaning that the users and producers of the commons are crafting, implementing and maintaining the institutional arrangements themselves. Based on these institutional arrangements, communities can successfully govern the commons without privatization or taxation from an outside governmental body.

The challenge that the tools developed in this book are meant to address is to understand what kind of institutional arrangements are successful in which circumstances. A coercive approach is not necessarily a productive approach. Coercion may demotivate participants. Providing monetary incentives may also not always be beneficial. An illustrative example is a study by Gneezy and Rustichini (2000) on daycare centers. Parents often come late to collect their kids from daycare. To reduce the number of people who are late, an experiment was performed that imposed a monetary penalty when parents were late. Surprisingly, parents came late more often. Why should anyone complain when they have paid for it? Parents who were willing to pay the price could come late without feeling guilty. When the daycare centers wanted to revert back to the original situation and remove the penalty, the number of parents who came late remained high. A behavior that is a moral obligation (coming on time to collect your child) became an economic transaction (paying a fee). This is a risk of using economic incentives to stimulate behavioral change; it may have unintended, difficult-to-reverse consequences. That is, economic behavior may 'crowd out' moral behavior.

The study of successful institutional arrangements shows that it is important that participants in action situations are involved in the creation of rules, that there are low-cost conflict-resolution mechanisms, and that there are clear rules about who and when people can use the commons. Effective institutional mechanisms stimulate personal interactions that facilitate trust relationships and allow participants to build reputations. When people edit the English text of Wikipedia articles, they gain respect and a good reputation in the community that may enable them to occupy a special role in the community. When a tennis player, who just lost a match, shakes the hand of the opponent, it reinforces the respectful relationship they have with each other.

The emerging picture of effective institutional arrangements is that in order to be successful , it is important that people can develop trust relationships, gain a reputation, experiment with new arrangements, tolerate mistakes people make, and have commonly understood rules-in-use. Most of these insights have been derived from studies of communities who share common-pool resources. If we know so much about successes, why are there still so many problems?

17.2 Why are there still so many problems in governance?

In this book, we have discussed insights relating to the ability of communities to solve collective action problems. If we know so much about what leads to effective institutional arrangements, why are there still so many problems? What prevents us from successfully governing shared infrastructure systems?

More than a billion people around the world do not have sanitation or access to clean water. Many species go extinct each year and human activities cause long-term disruption to biogeochemical cycles in nature. Many of us waste hours each week in traffic jams and complain about the performance of elected officials.

Knowing what leads to better institutional arrangements will not solve all these problems. What are the main challenges? What are the open questions in our understanding of institutional arrangements that require further research? In the following paragraphs, we attempt to list some of the most important challenges. This list is not exhaustive but, rather, represents only a starting point.

One of the big challenges in our modern society is the scale of the problems we face. We are no longer living in small communities where we know exactly what everybody is doing. We may not even know who our neighbors are. In an increasingly urbanized world, we interact with many people who are strangers to us. Even so, there is still an incredible level of cooperation in most modern economies. A moment's reflection should give the reader a sense of astonishment at the fact that hundreds of millions of people can effectively coordinate their behavior every day. How do we do this? Institutions are a big part of the story. We are able to signal to each other our reputation and trustworthiness because of the uniforms we wear (in the position of police officer, you must wear an official uniform), the tattoos we have, the certificates we have earned (positions defined by boundary rules) and the gossip that is spreading about us. It is not uncommon for us to give a stranger our credit card (backed by an enormous stock of institutions and organizations) to make a payment. We are accustomed to conditionally trusting strangers.

Nevertheless, the larger scale of our interaction spheres increases the possibility that we may lack the appropriate information to make good decisions. Think about people accepting the terms of home loans that they cannot understand. Think about large institutional investors who purchase investments for which they cannot assess the risks. The financial crisis of 2008 demonstrated just how calamitous and how much suffering such information failures may generate. There is also the possibility of misunderstandings about our intentions, motivations and the meaning of rules. An important condition of well-functioning institutional arrangements is that rules are commonly understood.

Being in larger groups makes it more difficult for individuals to be involved in rule crafting. In the position of a U.S. citizen of eighteen years or older, you may vote but you may also feel that your vote is insignificant. You may not be able to have an impact on the outcomes at the national level, but you still can participate in local governance issues whether this is through an elected office, community service project, or a volunteer activity for your children's school. Individual actions in the community add up. Because the impact of such activities is difficult to measure, the incentives to take them are weak. This is one of the fundamental problems of society—the under-provision of public goods.

Further, larger groups will make it easier to be invisible as a free-rider. You can be one of the many who do not volunteer. Larger groups make it more likely that there are different opinions and more disagreement among the participants. Disagreement makes it easier not to act, even though we know we should.

How can we stimulate cooperation in large populations? Can we apply the insights from this book to an urbanized and globalized world? New technologies may provide solutions. Many of us have a mobile phone with us; a small computer that can register where we are and can be used to take photos and exchange information with friends in social networks. Can we use these devices to improve the information we have about each other in order to improve trust in relationships and monitor the actions of each other? How might this impact personal privacy? How we may be able to use the crowd to govern the crowd without impinging on privacy is an important, open question.

Another big challenge is that new problems always emerge. With every new technology there are benefits but there also come new problems. There was no cyber bullying before the Internet. It is more difficult to bully someone in person than virtually. There was no illegal downloading before digital recording. To illegally obtain a music recording 50 years ago, it was necessary to walk into a record store and walk out with a vinyl disc! Again, before the Internet, stealing was a more personal affair—you had to actually see the victim. Now it has become impersonal. New problems also emerge due to new insights from science. Improved technology allows better measurements and enables new discoveries, such as the emergence of the hole in the ozone layer. Our understanding of chlorofluorocarbons enabled us to determine that they were responsible. Reducing chlorofluorocarbons was fairly easy—the problem was clear, measurable, and well understood. The solution was technologically feasible and economical. This is in stark contrast to climate change, which poses a much more difficult collective action problem. If and when we develop global governance arrangements to deal with climate change, what will be the next problem to emerge? Will human society ever have enough time to solve its existing set of social dilemmas before being presented with another new problem? Or, put in another way, will humans ever learn to craft institutions and governance structures fast enough to address new challenges?

History suggests there are some reasons to be hopeful—e.g., the Montreal Protocol, which deals with chlorofluorocarbons—but the challenges are many. Globalization will bring with it global-scale problems. These will require global-level solutions. This will require cooperation between people from many different cultures. New mixtures of populations may require generations to develop commonly understood well-functioning regulations, slowing our capacity to respond. Further, because solving social problems is difficult and complex, people tend to stick with institutions that have worked in the past. The No Child Left Behind program is a good example. The tried and true solutions (based on the Protestant work ethic) of trying to create incentives for more discipline and harder work through higher standards and more measurement simply does not work for a public good like education today. Why? The social context is completely different and "education" is complex. In order to learn more material more quickly as the No Child Left Behind Act demanded, children need mentoring. In the past when parents had the time to mentor, No Child Left Behind may have been a great success (at least by its own measure of improved standardized test scores). At present, when in many households both parents work and have little time or energy to mentor their children, higher standards and more testing will have little effect (the No Child Left Behind Act was replaced in 2015 by the Every Student Succeeds Act due to bipartisan criticisms). Old solutions do not translate well to new situations and simple panaceas will fail. Rather, we must perform small scale experiments to get experience with new institutional arrangements in new contexts. Because such experiments are costly and require patience, developing effective institutions will require considerable collective will on the part of society.

The third challenge we face is that it is often not in everybody's interest to solve a problem. Different people have different positions and interests. A problem for one participant can be an opportunity for another. Hence not everybody has an incentive to solve a problem. Problems don't exist in a vacuum, there is already a social and ecological context for every problem we face. If the poor and unemployed don't receive health care, it is not a direct benefit for those who have health care to pay for and share their health care benefits. The status quo, although not perfect, might be beneficial to many participants as compared to an alternative.

Finally, sometimes constitutional choice rules make it difficult to change a regulation. The European Union now consists of more than 25 nations. The EU employs an aggregation rule by which decisions are made by a unanimity vote. In a unanimity vote everybody needs to vote in favor in order for a proposal to be accepted. If the group is relatively small and people are sufficiently aligned in terms of their understanding and preferences, this will work. But in large groups, one individual country can take negotiations hostage to receive benefits for voting in favor of a motion.

17.3 Closing

In closing we can say that there have been significant developments over the past 50 years in our understanding of institutional arrangements and the way they structure social interactions. This book provides ways to study and analyze institutions. After reading this book, we hope you will view the problems we face everyday and the very diverse ways we are solving collective action problems through a new lens and in a different light.

Different disciplines contribute to our understanding of human behavior in the context of complex social and ecological systems. Unfortunately, we cannot provide a blueprint for how to solve all the problems we experience. Experimentation at the small scale and finding mechanisms to connect successful solutions to larger scales are key. Although we cannot provide simple solutions to complex problems, we have provided you with a powerful set of tools to make more informed decisions and recognize the importance of your own role in society.

17.4 Critical reflections

The rules and norms that govern human interactions can be studied with the framework that is presented in this book. The framework can be applied to many different topics including sustainability, health care, sports, education and the digital commons. Despite our increasing understanding of institutions and lessons regarding the conditions of successful collective action, there are still many failures.

Major challenges exist in governance in modern society since the scale at which we interact with others is much larger than it has ever been in human history. This makes lessons about success from studies of small groups difficult to apply. Furthermore, we experience misunderstandings if we don't speak the same language or live in different social and ecological contexts. Finally, new problems are constantly emerging due to rapid environmental and technological change.

17.5 Make yourself think

1. How can you make a difference in addressing major problems in society?

2. Ask older family members how they made arrangements for going out (in a time before mobile phones and texting). Do you see changes in rules and norms?

3. What do you see as the most challenging topic of governance in the future?

17.6 References

Gneezy, U. and A. Rustichini (2000). A fine is a price. *Journal of Legal Studies* 29(1): 1–17.

Ostrom, E. (2005) *Understanding institutional diversity*. Princeton, NJ: Princeton University Press.

Credits

Cover Photo: Photo by Quang Nguyen Vinh: https://www.pexels.com/photo/aerial-footgae-of-scenic-rice-fields-6711530/

Images appear courtesy of the owner, except in cases where author is unknown (after extensive research) or permission was repeatedly requested without response. If you believe a copyright has been violated, please contact CBIE at cbie@asu.edu.

Chapter 1

Figure 1.1: Accessed on May 6, 2013 from the Garret Hardin Society Website at http://www.garretthardinsociety.org/gh/gh_pictures.html. Photographer unknown.
Figure 1.2. Relative biomass estimates from the beginning of industrialized fishing (Myers, R.A. and B. Worm (2003) Rapid worldwide depletion of predatory fish communities, Nature 423: 280-283.)
Figure 1.3: First published in William R. Shepherd, Historical Atlas, New York, Henry Holt and Company, 1923. Public domain.
Figure 1.4: Reprinted courtesy of Indiana University.

Chapter 3

Figure 3.1: Photo credit Andrew*. Licensed under Creative

Commons.
Figure 3.2: Photo credit Midloa. Licensed under Creative Commons.
Figure 3.3: Cartoon owned by Kirk Anderson. Permission requested.
Figure 3.4: Photo credit Center for Climate and Energy Solutions. Permission pending.

Chapter 4

Figure 4.1: Image licensed from Shutterstock.
Figure 4.2: Image licensed from Shutterstock.
Figure 4.3: Photo credit the Vincent and Elinor Ostrom Workshop in Political Theory and Policy Analysis.
Figure 4.4: Adapted from Ostrom, E. (2005) Understanding Institutional Diversity, Princeton University Press.
Figure 4.5: Adapted from Ostrom, E. (2005) Understanding Institutional Diversity, Princeton University Press.
Figure 4.6: Original to this manuscript.

Chapter 5

Figure 5.1: Adapted from Ostrom, E. (2005) Understanding Institutional Diversity, Princeton University Press.
Figure 5.2: Photographer unknown.
Figure 5.3: Image licensed from Shutterstock.
Figure 5.4a: Photo credit Eckhard Pecher. Licensed under Creative Commons.
Figure 5.4b: Photo credit Steiner. Licensed under Creative Commons.
Figure 5.4c: Image licensed from Shutterstock.
Figure 5.4d: Image licensed from Shutterstock.
Figure 5.5a: Photo credit Pete Souza. Licensed under Creative Commons.
Figure 5.5b: Photo credit Steve Petteway, Collection of the Supreme

Court of the United States. This work is in the public domain.
Figure 5.5c: Photo owned by American Idol, Inc. Permission requested.
Figure 5.5d: Photo credit United States Department of Interior (photographer unknown). This work is in the public domain.
Figure 5.6: Photo credit United States Coast Guard (photographer unknown). This work is in the public domain.
Figure 5.7: Image licensed from Shutterstock.
Figure 5.8: Image licensed from Shutterstock.
Figure 5.9a: Photo used courtesy of TrueStory Films.
Figure 5.9b: Photo used courtesy of Ampersand Advisory Group.
Figure 5.9c: Photo used courtesy BWOG at Columbia University.
Figure 5.9d: Photo used courtesy of Arizona State Board of Regents.

Chapter 6

Figure 6.1: Image used courtesy of Hawk Eye Innovations.

Chapter 7

Figure 7.1: Image from Lake Scientist, “Eutrophication” by Jermy Mack. Accessed January 6, 2016 http://www.lakescientist.com/lake-facts/water-quality/. Permission pending.
Figure 7.2: Original to this manuscript.
Figure 7.3: Original to this manuscript
Figure 7.4: Photo credit Simon Fraser University Public Affairs and Media Relations. Licensed under Creative Commons.
Figure 7.6: Photo credit Malcolm Paterson, CSIRO. Licensed under Creative Commons.
Figure 7.7: Photo credit Jovel. Licensed under Creative Commons.

Chapter 8

Figure 8.1: Original to this manuscript

Figure 8.2: Original to this manuscript
Figure 8.3: Original to this manuscript
Figure 8.4: Original to this manuscript

Chapter 9

Figure 9.1: Figure 1 of doi: 10.18352/ijc.651. Licensed under CC BY 4.0
Figure 9.2: Photo credit Kelly Garbato. Licensed under Creative Commons

Chapter 10

Figure 10.1: Photo used courtesy of About.com.
Figure 10.2: Photo credit U.S. Department of Interior Reclamation. This work is in the public domain.
Figure 10.3: Map owned by NedWater. Permission requested.
Figure 10.4: Photo credit Lucas Hirschegger. Licensed under Creative Commons.
Figure 10.5: Photo credit Sono Adriano. Licensed under Creative Commons.

Chapter 11

Figure 11.1: Original to this manuscript

Chapter 12

Figure 12.1: Patten White, a fisherman out of York, Maine. Photo used courtesy of Gulf of Maine Lobster Foundation.
Figure 12.2: Photo used courtesy of the AVC Lobster Science Centre.
Figure 12.3: Graphs created based on data from Maine.gov

"Historical Summary of Maine Lobster Fisheries." http://www.maine.gov/dmr/rm/lobster/lobdata.htm.
Figure 12.4: Figure 1 of FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Rome. https://doi.org/10.4060/ca9229en published under CC BY-NC-SA 3.0 IGO
Figure 12.5: Figure 1 9 of FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Rome. https://doi.org/10.4060/ca9229en published under CC BY-NC-SA 3.0 IGO
Figure 12.6: Photo used courtesy Sajid Pareeth.
Figure 12.7. Figure 2 of FAO and UNEP. 2020. The State of the World's Forests 2020. Forests, biodiversity and people. Rome. DOI: https://doi.org/10.4060/ca8642en published under CC BY-NC-SA 3.0 IGO
Figure 12.8: Photo used courtesy of Torbel Moosalp Tourimus.
Figure 12.9: Photo used courtesy of Gita Mehra.
Figure 12.10: Original to this manuscript

Figure 14

Figure 14.1: Original to this manuscript.
Figure 14.2: Original to this manuscript.

Chapter 17

Figure 17.1: Adapted from Ostrom, E. (2005) Understanding Institutional Diversity, Princeton University Press.