Received: 18 September 2021 | Revised: 12 October 2021 | Accepted: 12 October 2021

DOI: 10.1111/1758-5899.13030

POLICY INSIGHTS

A Safe Governance Space for Humanity: Necessary Conditions for the Governance of Global Catastrophic Risks

Len Fisher[1] | **Anders Sandberg**[2]

[1]School of Physics, University of Bristol, Bristol, UK

[2]Future of Humanity Institute, Oxford Martin School, University of Oxford, Oxford, UK

Correspondence
Len Fisher, University of Bristol Physics, Tyndall Ave., Bristol BS8 1TL, UK.
Email: len.fisher@bristol.ac.uk

Funding information
H2020 European Research Council, Grant/Award Number: 669751

Abstract

The world faces a multiplicity of global catastrophic risks (GCRs), whose functionality as individual and collective complex adaptive networks (CANs) poses unique problems for governance in a world that itself comprises an intricately interlinked set of CANs. Here we examine necessary conditions for new approaches to governance that consider the known properties of CANs—especially that small changes in one part of the system can cascade and amplify throughout the system and that the system as a whole can also undergo rapid, dramatic, and often unpredictable change with little or no warning.

Many governance schemes have been proposed for the management of global catastrophic risks,[1] defined by Bostrom and Ćirković (2008) as situations that 'have the potential to inflict serious damage to human well-being on a global scale'. We argue here that most of these schemes suffer from a fatal logical flaw, in that they begin with a favoured system of governance and attempt to apply it to the world situation, rather than examining the world situation and asking what system of governance might be most appropriate. Here we analyse some of the major schemes that have been proposed, and ask how they stack up against the criteria required for governance in the face of real-world complexity.

Our argument is developed in four steps:

1. A brief review of global catastrophic risks (GCRs) and their governance
2. Conceptual framing of our social-economic-ecological world and the threats that endanger it as complex adaptive networks (CANs)
3. Analysis of the necessary conditions for the effective governance of GCRs as CANs
4. Evaluation of different proposed forms of governance in terms of those necessary conditions.

1 | PRINCIPLES FOR GOVERNANCE OF GLOBAL CATASTROPHIC RISKS

Bostrom & Ćirković's definition of global catastrophic risks (GCRs) states that they must be 'serious', but without defining this term. Here we adopt a criterion suggested by the authors themselves, that a GCR is serious if its consequences are likely adversely to affect tens of millions of people, or to cost trillions of (US) dollars.

Many current or looming events fit these conditions. We have culled a (non-exhaustive) list of those that are believed by many authors to be among the most important from the World Economic Forum *Global Risks Report* (World Economic Forum, 2020), the Global Challenges Foundation *Global Catastrophic Risks 2020* (Global Challenges Foundation, 2020), the Stockholm Resilience Centre review *Planetary Boundaries*: *Exploring the Safe Operating Space for Humanity* (Rockström et al., (2009), the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES, 2019), the Intergovernmantal Panel on Climate Change 6th Assessment Report (IPCC, 2021), and Toby Ord *The Precipice* (Ord, 2020). Our criteria for inclusion are that the risk has been identified by multiple sources as being among the most important and that claims for its catastrophic nature are based on hard evidence.

We will demonstrate that much of the power of each threat derives from its being a component of a complex network, whose other members include the individuals, communities, and environments that are under threat. The individual threats are also linked with each other to form an overarching network whose governance must be considered as a whole.

Our non-exhaustive list comprises:

Internal threats

1. Climate change
2. Loss of biodiversity
3. Degrading environment and resource depletion
4. Food insecurity
5. Pandemics
6. Population increase and urban expansion
7. Collapse of international governance
8. Unaligned artificial intelligence
9. Cyber risks
10. Increasing polarisation of societies
11. Rising disparity of income and wealth
12. Weapons of mass destruction
13. Great power war
14. Genocidal totalitarianism
15. Runaway technological disasters

External threats

16. Asteroid impact
17. Supervolcanic eruptions
18. Geomagnetic storms generated by solar superflares

1.1 | Governance principles for GCRs

Any successful governance scheme for GCRs must take into account their variability in scope, severity, and probability (Avin et al., 2018). There are strong arguments (Ord, 2020) for giving high priority to existential risks, even those with relatively low probability. As an aid to prioritisation, Bostrom (2013) has proposed a 'rule of thumb' *maxipok* principle: *Maximise the probability of an 'OK outcome', where an OK outcome is any outcome that avoids existential catastrophe*.

Bostrom points out that this principle, although superficially similar to the well-known *maximin* principle ('choose the action that has the best worst-case outcome'), is in fact quite different in outcome. The maxipok principle promotes relevant action, whereas the maximin principle is open to the interpretation that, in the face of existential risk, 'we ought all to start partying is if there were no tomorrow'.

The maximin principle nevertheless has some merit for lesser, but still catastrophic, risks, so long as there is enough information for the best worst-case outcome to be reliably assessed (e.g. Bognar, 2011; Sunstein, 2019). If this is not the case, then the *precautionary principle* comes into play. The principle has been formulated in a number of different ways (references in Clarke, 2005) and may be exemplified by the closing Ministerial Declaration from the United Nations Economic Conference for Europe in 1990, which states that 'When there are threats of serious or irreversible damage, lack of full scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation' (quoted in Sunstein, 2007).

The precautionary principle has been the subject of extensive philosophical and political debate (Read & O'Riordan, 2017). Failure to apply it at the start of the Covid-19 pandemic may have been responsible for many excess deaths (Basili, 2019), but its application later in the pandemic, when the dangers of the AstraZeneca vaccine came into question, may also have resulted in excess deaths (Faranda et al., 2021). Clarke (2005) also points out that the precautionary principle, as commonly formulated, leads to a paradox. It suggests, for example, holding back on 'risky' research in some areas. But what if that research provides the only route to an eventual solution?

Sunstein (2007) has suggested a stronger form of the principle, in the form of the *Catastrophic Harm Precautionary Principle*: 'When risks have catastrophic worst-case scenarios, it makes sense to pay special attention to those risks, even when existing information does not enable regulators to make a reliable judgement about the probability that the worst-case scenarios will occur'.

One way of paying special attention to catastrophic risks is what Turchin (2018) calls the 'Plan A, Plan B' model. In this dual approach, Plan B is 'a backup option, implemented if Plan A fails. In the case of global risks, Plan A is intended to prevent a catastrophe and Plan B to survive it ….' Turchin claims that this model has 'shown its effectiveness in planning actions in unpredictable environments'. Other models that make similar claims are those based on resilience (Folke et al., 2010), sustainability (Burch et al., 2019), and the primacy of human rights (Voeneky, 2019).

A similar, but more subtle, scheme has been proposed by Cotton-Barratt et al., (2020) as a 'Defence in Depth' against human extinction. In this scheme, three sequential layers of protection provide a defensive structure, in the manner of the concentric defences of a mediaeval castle (Faulkner, 1963). The layers here are Prevention, Response, and Resilience, with the inner layer of resilience especially acting to prevent global catastrophes from becoming extinction catastrophes.

All of these schemes, and others that have been suggested in the very large literature on global catastrophic risks (cf. Baum & Barrett, 2018; Baum & Handoh, 2014; Galaz, 2019) and global systemic risks (Centeno et al., 2015), come with question marks as to *when* they should be implemented and to *how* they should be implemented. Sunstein (2007), for example, admits that the Catastrophic Harm Precautionary Principle is 'lamentably vague' in these regards. It does not 'specify the threshold information that would trigger the principle; the role of costs; and how regulators should incorporate whatever information exists about the probability of catastrophe'.

Faber (2011) offers a specific framework in response to these questions. According to this framework, schemes for the management of catastrophic risks must fulfil the following ten practical requirements:

1. facilitate modeling of the considered system such that all relevant events leading to losses may be represented together with their interdependencies
2. consistently account for the level of available knowledge as well as natural variability
3. facilitate decision making at a scale of system representation necessary to support the decisions in question
4. quantify risks in a marginal as well as a non-marginal sense; i.e., be able to represent the effect of losses due to a given event on economic growth and the living conditions for future generations
5. specifically address decision making in the situations before, during and after hazard events
6. facilitate standardised procedures for systems representations in risk assessments
7. account for information which might become available in the future and facilitate that options for future decisions are included in the decision optimisation
8. facilitate for consistent risk aggregation whereby it is ensured that the results of independently performed risk assessments can be applied to assess and manage the risk in larger context-portfolios
9. facilitate decision optimisation and the assessment of the acceptability of decisions
10. enhance risk communication and risk management documentation

These general principles for risk management apply to all types of system. We now show how they emerge naturally as general principles for the governance of complex adaptive systems.

2 | CONCEPTUAL FRAMING OF THE WORLD AND GCRS AS COMPLEX ADAPTIVE NETWORKS

The conceptual framing of our global socio-(economic)–ecological system as a complex adaptive network, in which the components interact in nonlinear ways, with many positive and negative feedback loops, was initiated in the 1990s (Pohl, 1999). It has since been put on a firm footing (Levin, 2019; Levin et al., 2013; Ostrom, 2009; Sayama et al., 2013; Schweitzer et al., 2009). The typical features of such a system (Chan, 2001; Helbing, 2013; Pattberg & Widerberg, 2019; Sayama et al., 2013) are:

- Connectivity (the system forms a network).
- Self-organisation and strong correlations dominate the system behaviour, and elements can co-evolve, based on their interactions with other elements and the environment.
- Distributed control (no single centralised control mechanism, so that opportunities for external or top–down control are very limited).
- Sensitive dependence on initial conditions (a small change in one part of the system can lead to large [often unpredictable] changes in other parts). When change *does* happen, the system might show numerous different behaviours (multiple equilibria), depending on the respective initial conditions.
- Emergent order: The behaviour of the system cannot be understood or predicted just by understanding the behaviour of the individual elements (Miller & Page, 2007).

Our socio-economic–ecological world displays all of these features (Levin, 2019; Levin et al., 2013; Ostrom, 2009; Pohl, 1999; Sayama et al., 2013; Schweitzer et al., 2009). Its individual members (people, societies, ecosystems, economies, plants, animals, oceans, atmosphere, etc.) interact either directly or indirectly, and change over time as a result of these interactions. There is no central control of these interactions. A small change in one part of the system (collapse of a bank or the eating of a bat) can lead to dramatic, system-wide changes (financial collapse, pandemic). It is usually impossible to predict the long-term effects of the behaviour of the individual members of the system.

The governance system itself can be a complex system in its own right (e.g. international law [Kim and Mackey, 2014]), and is also a part of the larger complex adaptive system. In terms outlined by George Soros (2013) and placed into the context of complex adaptive systems by Eric Beinhocker (2013), it is *fallible* and *reflexive*. *Fallible*, because the complexity of the world that we are trying to govern exceeds our capacity to understand it. *Reflexive*, because the governance system is an active participant in the system that it is trying to govern. Thus, any governance actions are liable to feed back and affect the governance system itself. According to Beinhocker, such a reflexive system has two additional elements that distinguish it from a normal dynamic feedback system:

- *Internal model updating*: The internal decision model of the agents (governance systems) is not fixed, but can itself change in response to interactions between the agent and its environment (the system to be governed).
- *Complexity*: The system has *interactive complexity* owing to multiple interactions between heterogeneous agents, and *dynamic complexity* owing to nonlinearity in feedbacks in the system.

2.1 | GCRs as CANs

Global catastrophic risks themselves form an interconnected network (Fisher, 2019) that has all of the characteristics of a complex adaptive network (Levin et al., 2013). Global warming, for example, is connected to food security, with longer growing seasons meaning that pests can increasingly survive between seasons. Our evolving choice of food, on the other hand, may affect global warming (Wilett et al., 2019). Food insecurity can even drive revolution and war (Lagi et al., 2011), which affect food supplies in their turn.

Each threat has an internal structure which makes it a complex network (global warming, for example, involves many interlinked chemical, physical, and social processes, with multiple feedback loops). Each network also has most or all of the characteristics of a complex *adaptive* network (Table 1). The assembly of networks also forms a super-complex adaptive network, whose governance must be considered as a whole.

A relatively clear-cut example is provided by the pandemic spreading of the Covid-19 virus, which produces CAS and CAN dynamics. As shown in Figure 1, there are several positive feedback loops producing accelerating change and sensitivity to initial conditions, but also inhibitory feedback allowing for bistability and oscillation. Control is distributed among numerous actors who update their behaviour based on their partial understanding of the system. Strength of interaction can change, different subsystems overlap, and external factors can feed into the dynamics unexpectedly.

Beyond this simple model, Covid-19 has had obvious outside knock-on effects such as the cancellation of sports tournaments, closure of restaurants, restriction of travel, social isolation, and loss of income for many small businesses (Haleem et al., 2020). But there are many less obvious connections, including reduced carbon dioxide emissions (Anjum, 2020), an increase in endangered sea turtle nesting and hatchling survival as beaches remain clear of people and rubbish (Luscombe, 2020), and interlinked disruptions to the global economy that could even lead to the reversal of globalisation and large consequent shifts in the economic power base (Baldwin & di Mauro, 2020).

The Covid-19 pandemic has also generated considerable mistrust in governance structures across the world, whose behaviour has had to change in response (Garrett, 2020). Early overconfidence in some cases was later used as evidence of lack of knowledge among authorities, leading to maladaptive public or institutional responses, such as reluctance to use masks, vaccinate, or increase testing capacity. The consequences of these responses fed back into the systems in the form of greater infection rates, leading to further cycles of more vigorous action (such as compulsory mask wearing), and stronger public responses to these actions (such as public demonstrations). These are examples of reflexivity in action, and reinforce our point that the governance of GCRs cannot be considered in isolation, but only in the overall context of governance of CANs, and especially an awareness that ongoing feedback loops are always liable to offer a potential for instabilities.

3 | NECESSARY CONDITIONS FOR THE GOVERNANCE OF GCRS

3.1 | Unsuitability of current governance structures

As the examples above reveal, GCRs constitute a unique challenge to governance. Klinke (2014) argues that 'the key peculiarities of global risks—complexity, scientific uncertainty, and socio-political ambiguity—are … generic features' and that 'there is a lack of a broader societal and political consensus of how to handle this kind of insecurity'. Silja Voeneky (2019) offers many concrete examples, from artificial intelligence to gene editing, and points out that 'Thus far, no international treaty on existential and global catastrophic risks and scientific research exists' and that, in general, 'international treaty law is not sufficient to govern these research areas'.

Pegram and Kreienkamp (2019) argue that the major problem is that legacy governance structures, such as the UN Security Council or the General Agreement on Tariffs and Trade, are designed for the administration of *complicated* problems, which 'may have many components, but the relationships between the components are fixed and clearly defined' so that 'a rules-based governing framework is appropriate to establish order and control' because 'cause and effect relationships are linear such that … we can identify a clear cause for each observed effect and predict system-level outcomes of each change'.

Complicated problems, they say, are however quite different from *complex* problems, where 'The relationship between cause and effect is nonlinear and effects are usually the result of several interacting causes. Due to feedback loops, we cannot establish clear

TABLE 1 Examples of how different GCR threats have CAN features embedded

	Connectivity	Self-organisation	Distributed control	Sensitive dependence	Emergent order
Climate change	Complex feedbacks between atmosphere, geosphere, biosphere, anthroposphere	Strong couplings, biosphere and economy respond to climate events and policy	Multilateral international political/economic system	Chaotic weather/disaster responses; system tipping points	Yes
Loss of biodiversity	Ecosystem nutrient web; ecosystem services; economy	Ecosystems and economies change as a response (adaptation, mitigation, restoration); transnational cooperation and spillovers	Species, farmers, industry, governments	Keystone species; biodiversity hotspots; multiple equilibria (e.g. kelp forest/urchin barren)	Yes
Degrading environment and resource depletion	International economy; global supply chains	Economic price responses; technological substitution and adaptation	Industry, governments	Alternative economic or technology paths; innovation	Yes
Food insecurity	Connected energy-agriculture-distribution system; economy as a whole	Price shocks; farmer and societal adaptation	Crops, pests, farmers, industry, governments	Climate and pest drivers; choice of agricultural system; import/export policy choices	Yes
Pandemics	International transport network; social interaction network	Ongoing research and experience changes response; infodemics; literal evolution of pathogens	Individual and societal decisions, international organisations	Exponential pathogen growth; policy decisions; successful or unsuccessful containment	Yes
Population increase and urban expansion	Networked demographic factors (urban economics; health systems; education); culture	Demographic and urban network effects (e.g. costs of child-rearing, urban economies of scale); cultural shifts	Individual and local decision-making. Rare cases of top–down control (interacting with individual choices in complex ways, e.g. China one-child policy, planned cities)	-	Yes
Collapse of international governance	International governance norms, treaties, laws, and relationships	Cascade effects; formation or dissolution of international organisations or alliances	Sovereign states	International crisis events; formation/breakup of alliance constellations	Yes
Unaligned artificial intelligence	(Varies depending on scenario)	Self-improving technology; instrumental goal convergence	Software, programmers, companies, governments	Intelligence amplifies probability of desired goal states from low-probability states	?
Cyber risks	Internet; economics of cyber actors	Technological, legal, and economic responses	Software, programmers, companies, governments	Exploit detection; technological and security regime choices; liability and insurance rules	Yes
Increasing polarisation of societies	Social media networks; social networks	Sociological group dynamics; online and offline community construction	-	-	Yes
Rising disparity of income and wealth	Economic network	Rich-get-richer-dynamics; redistribution policies	-	-	Yes

TABLE 1 (Continued)

	Connectivity	Self-organisation	Distributed control	Sensitive dependence	Emergent order
Weapons of mass destruction	(International alliance network a driver)	Arms race dynamics	(Sovereign states and UN)	Technological discovery and availability; different control regimes	?
Great power war	(International alliance network a driver)	Strategic and tactical interactions; responses in all affected parts of society	(Sovereign states)	Inciting events causing escalation; randomness of war	Yes
Genocidal totalitarianism	Hierarchical social networks	-	-	-	?
Runaway technological disasters	Varies	Varies	Varies	Varies	Varies
Asteroid impact	-	Disaster response	-	Timing and size determines location and consequences (e.g. land-based fires, urban disaster, or tsunami)	Yes (in human disaster response)
Supervolcanic eruptions	Environment; food system	Disaster response	-	-	Yes (in human disaster response)
Geomagnetic storms generated by solar superflares	Power grid networks	Disaster response	-	-	Yes (in human disaster response)

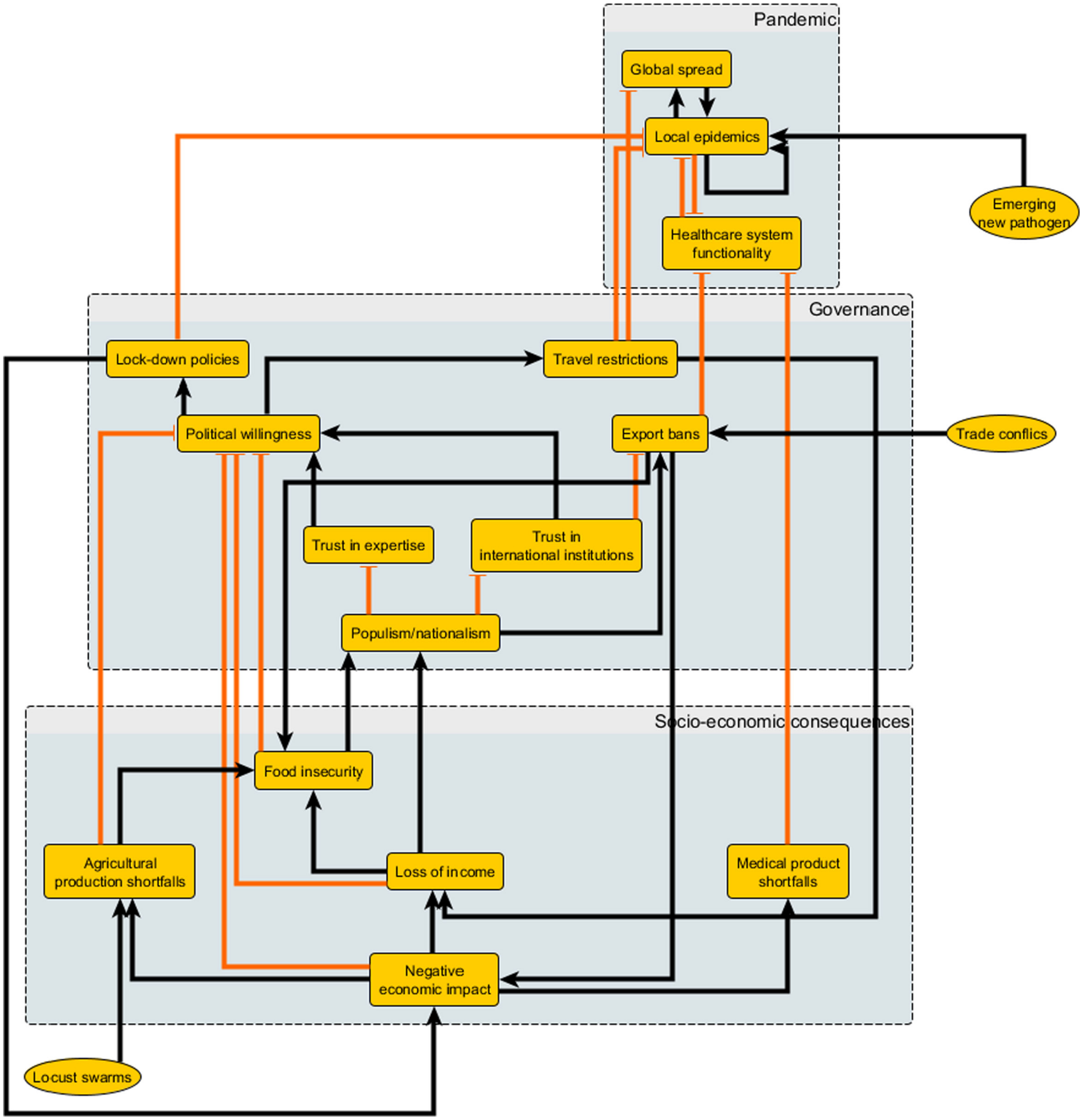

FIGURE 1 Simple CAN feedback model of part the Covid-19 pandemic system. Black lines indicate amplifying impact; red lines indicate inhibiting impact. This figure suppresses the spatial and organisational dimensions: Most factors are actually clusters of linked but separate (sub/inter)national factors.

cause-and-effect relationships or predict system-level outcomes'.

Andy Haldane, then chief economist of the Bank of England, made this point in a speech delivered to the Peterson Institute for International Economics (Haldane, 2017). He demonstrated that the global financial system behaves as a complex adaptive system, and that 'Complex systems exhibit tipping points, with small changes in parameter values capable of moving the system from stability to collapse. In complex webs, the failure of two identical-looking banks can have very different implications for financial system stability. *The radical uncertainty in such complex webs generates emergent behaviour which can be near-impossible to predict, model, and estimate*' [our emphasis].

Haldane went on to argue that traditional governance systems, which are based on prediction, modelling, and estimating, are ill-suited to the governance of the world's financial networks, and that a new approach must be sought. The same argument applies to GCRs.

Adriana Abdenur (2020), writing for the Global Challenges Foundation, argues that 'rather than inventing new governance mechanisms from scratch, the most effective and legitimate route for dealing with unknown (or little understood) risks is to strengthen the existing global governance system'. We believe that this approach, unfortunately still being used by many governments and international organisations such as the UN (to which Abdenur is an adviser), is ill conceived in principle and dangerous in practice.

3.2 | A fresh start: Key conditions for effective governance of global catastrophic risks

We argue that human society *does* need new governance mechanisms, better suited to handling the catastrophic risks that it now faces. We examine here the necessary conditions for the governance of such risks in the light of their behaviour as CANs, and then analyse the types of governance system best adapted to implementing those principles.

Our list derives from our considerations of GCRs as CANs. We have identified five necessary conditions for their governance. These may not be sufficient, and indeed there may be more, but these five at least are necessary for effective governance.

3.2.1 | Recognition

Successful governance must consist in maximising the chances of the best outcomes while preparing for the worst. An effective governance system must be 'epistemically humble' about what it can predict and control. Unfortunately, human nature seeks certainty (Kruglanski & Orehek, 2012), which means that incentives in governance have generally favoured avoiding uncertainty, and that politicians and other decision-makers have tended to overclaim their degree of control. The feedback following inevitable failure is another example of both fallibility and reflexivity in governance.

The first and obvious requirement for the effective governance of global catastrophic risks is **recognition** that the traditional goals of certainty and control are not generally achievable (Makridakis & Taleb, 2009). In particular, the risks involved are not usually susceptible to traditional methods of top–down governance, the governance system itself forms part of the network (Kooiman, 2003), and the governance system may even be a threat to stability on its own account (Keohane, 2001).

This is the opposite of the traditional concept of 'legibility'—the approach of viewing a system to be governed in simplistic, orderly terms that make it governable (Scott, 1999). In real life, this still-common approach (reflected in the common political demand to provide explanations that can fit on a single sheet of paper) (1) looks at a complex and confusing reality; (2) fails to understand the subtleties of how the complex reality works; (3) attributes that failure to the irrationality of the system being looked at; (4) comes up with an idealised version of how it *ought* to look; and (5) uses authoritarian power to impose that vision, demolishing the old reality (Rao, 2010). Scott provides many real life examples; the reader can no doubt furnish more of this very common approach to governance, which is exemplified by the history of changing approaches to mask wearing during the Covid-19 pandemic, with politicians frequently imposing simplistic 'solutions' on what is a confused and complex reality (McConnell & Stark, 2021).

3.2.2 | Flexibility and speed

Because CANs can undergo rapid, irreversible, dramatic change with little or no warning, effective governance requires **flexible, rapid decision-making processes** that can respond to and cope with such changes.

Ashby's Law of Requisite Variety (Ashby, 1958) suggests that this can only be achieved if the governance system has more potential variety than the system to be governed. Peters et al. (2019) argue that this need not be the case, and point to simple strategies such as that of Balinese rice farmers (copy your most successful neighbours) that have enabled them to survive the vicissitudes of politics and war over centuries. Gigerenzer and Todd (2001) have provided evidence of the success of such simple ('heuristic') approaches. Perhaps Ashby's Law should be replaced by the not-quite-equivalent 'The only way to control your destiny is to be more flexible than your environment' (Dawson, 2012). Requisite variety is just one way to achieve such flexibility, but a more effective way may be to concentrate on just a few key issues or decision points where change can be implemented rapidly.

Rate factors are certainly important in many cases, especially when one part of the system cannot keep up with the rate of change in another part and loses the previous relation to it. One example is soil carbon-temperature feedback, where rapid warming causes CO_2 release, and possibly the collapse of thermohaline circulation in the deep ocean (Ashwin et al., 2012). In governance itself, there are numerous examples when governance does not or cannot keep up with change or overshoots change, as with the governance of climate change (Victor, 2011), and the resistance to 'lock-down' measures in some parts of the United States during the Covid-19 pandemic (Pellis et al., 2020; Sevastopulo & Shubber, 2020), A useful illustrative example is offered by Simon Levin (2019). 'Many corals and barnacles', he says 'have evolved rigid structures that resist strong flows, whereas the bull kelp bend with the flows. In our societies, as in the marine environment, rigid design and robust components may work best over the short term; but a flexible adaptive component, either bending with the flow or involving replaceable components, can prolong persistence. The right balance between them varies from organism to organism, and from strategy to strategy'.

Rate factors become important in a different way when considering the speed at which computer-aided decision-making can take place. 'Speedups appear to pose a serious challenge to human ability to control technological processes due to growing gaps of speed between computation and control ("cybernetic gaps") and challenges to setting the goals they are optimising for due to gaps of speed between computation and the human world ("ethical gaps"), in turn posing a profound challenge to governance systems that are themselves to some extent hybrid human computational systems suffering internal speed gaps' (Sandberg, 2019).

3.2.3 | Integrated monitoring and action

Successful application of Ashby's Law (or any simpler version) requires the **ability to monitor** the ongoing behaviour of the network and its interactions and to **act** on this information. Clearly not everything can be known, but it is important at least to capture key features that can serve as a guide to action.

For example, if we can predict that something (e.g. atmospheric CO_2 concentrations) will have an effect (climate change), then we can focus governance on that something. As an extreme example, if all contacts of infected persons in a pandemic can be traced, they can be isolated and the spread of the disease brought under control (Keeling et al., 2020).

Monitoring the structure of the network itself can also help with effective governance. It may help to avoid tipping points (known technically as *critical transitions*) through guiding changes in the organisation of the connections, as could have been the case with the global financial crisis of 2008 (May & Haldane, 2011). Even under conditions of deep uncertainty, monitoring can still be valuable in setting limits on the number and type of scenarios that need to be considered (Walker et al., 2010).

Reflexivity may appear to be a fundamental limit to monitoring, causing an infinite regress of considering the consequences of monitoring. However, many existing engineering systems accurately take into account their own predictions using, for example, adaptive control theory and Bellman's equations (Bellman, 1961). This is possible because they typically do not aspire to perfection, merely a high level of practical optimality. The reflexivity problem is by no means easy, but it is not unsolvable if one is willing to work with approximations.

Some of the deepest uncertainties can occur when stochastic internal variability triggers a shift in the state of a system. There may be a complete lack of warning (Lenton, 2013), and actions during rapidly changing situations (such as the occurrence of a new pandemic) must be taken 'on the hoof'. Integrated monitoring and action is especially important during such scenarios.

Sometimes, however, there can be warning signs. Bifurcation tipping points, for example, are often preceded by *critical slowing down* (Scheffer et al., 2009), where the system becomes more and more sluggish in its response to small perturbations and disruptions. It is important to monitor and respond to such warning signs before a 'runaway' situation develops. This can require substantial changes in governance culture. As the history of actions to cope with climate change has demonstrated (Harrison & Geyer, 2019), it can be difficult to persuade policymakers to take warning signals seriously until it is too late. Also, the interpretation of some early-warning signs may be subject to the prosecutor's fallacy—'conditionally selecting systems known to experience a transition of some sort and failing to account for the bias that this introduces' (Boettiger & Hastings, 2012).

Another change in culture concerns care in the use of metrics. Once an indicator is made into a policy target, it can lose the information content that qualifies it to play its role as an indicator (Newton, 2011). This effect (known as Goodhart's Law) is particularly relevant to the governance of CANs, because indicators and the system reciprocally affect each other (Manheim, 2016, 2018). Therein lies the problem, because 'Complex systems can only be managed using metrics, and once the metrics are put in place, everyone is being incentivised to follow the system's logic, to the exclusion of the original goals. If you're not careful with your metrics, you're not careful with your decisions. And you can't be careful enough' (Manheim, 2018). A prime example is the failure of the algorithm for modifying UK examination results in 2020 (Hao, 2020).

These various *caveats*, however, are not arguments against the use of integrated monitoring and action as a support for effective governance. They illustrate, rather, the importance of using the information gained in a precise and accurate manner.

3.2.4 | Cooperation and coordination

It hardly needs saying that achieving the necessary monitoring and action requires *cooperation* and *coordination* at individual, group, and international levels. The principles underlying effective cooperation have been the subject of numerous studies, with action often being sadly restricted by Underdal's 'Law of the Least

Ambitious Programme' (Victor, 2006), which says that action tends to be restricted by the least enthusiastic party.

Cooperation and coordination are nevertheless necessary for the governance of GCRs, because flexibility and speed are generally unachievable without them. They are especially important in three key areas:

1. taking actions that change the system to meet goals (e.g. reducing greenhouse gas emissions to mitigate climate change (Mason et al., 2017; Mattoo & Subramanian, 2013; Victor, 2016);
2. taking actions that reduce uncertainty, both in practical terms (e.g. government guarantees, insurance [Louaas & Picard, 2020]) and in terms of community perceptions (Kuhlemann, 2019; Wachinger et al., 2013);
3. steering the system away from tipping points (Galaz et al., 2016; for example, reducing the reproduction number R to below 1 so as to stop the spread of a pandemic (Nouvellet et al., 2021).

3.2.5 | Resilience and preparedness

Finally, effective governance of global systemic risks needs to recognise that unexpected or unpredictable systemic change is always on the cards, and that dealing with such change requires **preparedness** for situations when change becomes inevitable.

When it comes to complex adaptive systems, effective preparedness for sudden change involves **investment** in resilience, which may mean investment in restoring the status quo and/or investment in adapting to new situations (Carpenter et al., 2012; Fisher, 2015).

An example of the former is resilience planning for global catastrophic biological threats such as pandemics, biological weapons, and synthetic biological risks. According to Luby and Arthur (2019), resilience planning should occur at multiple levels and take several forms, including having distributed systems (e.g. urban gardens and urban farms) to provide essential food, water, and power, because these are far less susceptible to cataclysmic point failure than completely centralised systems.

Implicit in Luby and Arthur's proposal is the idea that resilience should involve protection of the current system and an eventual return to normality. This may not always be possible, however, or even desirable (cf. Kareiva and Fuller, 2016), and resilience may need to involve the capacity to adapt and transform (Carpenter et al., 2012).

ALLFED (The Alliance to Feed the Earth in Disasters) has considered a number of options with regard to the provision of food in the event of a natural disaster, such as a massive volcanic explosion that fills the atmosphere with dust and blocks out the sunlight necessary for normal plant growth. Stockpiling, microbial electrosynthesis, scaling of greenhouse crop production to low sunlight scenarios, and the use of microbial protein are just some of the scenarios under consideration (Baum et al., 2015).

Importantly, and especially because the most serious GCRs are so unpredictable, the investment in either case must be made ahead of time. Persuading those in power of this necessity is, perhaps, the most difficult problem of all.

4 | POTENTIAL SYSTEMS OF GOVERNANCE

It is clear that most, if not all, current governance systems do not and cannot meet the necessary criteria as outlined above. The reasons for this have been spelled out by a number of authors (e.g. Duit & Galaz, 2008; Young, 2017), and especially papers in Galaz, 2019). Here we examine some of the major alternative governance systems that have been proposed and ask how they stack up against our five conditions (see Table 2).

4.1 | Close fits to necessary conditions

We find that three of the proposed sets of governance principles (Control vs. Emergence, Adaptive Policies for Handling Deep Uncertainty, and Dynamic Adaptive Policy Pathways) fulfil all five of our necessary conditions, while two others (Resilience Thinking and Sensitive Intervention Points) come very close. Here we examine them in greater detail.

4.1.1 | Balance between positive and negative feedback; control versus emergence (Choi et al., 2001).

Thomas Choi and his colleagues point out that supply chain networks are often complex adaptive networks that 'emerge', rather than resulting from purposeful design by a single entity. The problems of their management/governance are thus similar in principle to those of other complex adaptive networks, including GCRs, which can similarly emerge from a combination of circumstances, rather than a single identifiable cause.

The major problem identified by Choi et al. is selecting an appropriate balance between control and emergence. 'The emergent patterns in a supply network', they argue, 'can much better be managed through positive feedback, which allows for autonomous action. [But] allowing too much emergence can undermine managerial predictability and work routines [while] imposing too much control detracts from innovation and flexibility'.

TABLE 2 Selected proposed governance systems assessed in terms of our five necessary conditions

Proposal	Recognition	Flexibility and speed	Integrated monitoring and action	Cooperation and coordination	Resilience and preparedness
Act local; think global (Clemens, 2013)		+		+ (in part)	
Dispersed authority (Brosig, 2019)		+			
Multiple plausible futures (Maier et al., 2016)	+	+	+		+
Scenario planning *via* ensemble forecasting (Lempert, 2002)	1(?)		+	+	
Resilience thinking (Berkes, 2007; Folke, 2019; Folke et al., 2010)	+	+	+	+ (but not far enough)	+
Balance between positive and negative feedback; control v emergence (Choi et al., 2001)	+	+	+	+	+
Adaptive management (Allen et al., 2011)	(Depends on situation)				
Reframing decision theory for CAS (Bankes (2002))	+?	+?	+?	+?	+?
Adaptive policies for handling deep uncertainty (Walker et al., 2010)	+	+	+ (implicit)	+	+
Decision theory plus threshold approach (Polasky et al., 2011)	+		+		+
Dynamic Adaptive Policy Pathways (Haasnoot et al., 2013; Kwakkel et al., 2016)	+	+	+	+	+
Orchestrating Interactions Between Institutions (Haas 2019)				+	
Catalytic Probes (Harrison and Geyer, 2019)			+		
Sensitive Intervention Points (Farmer et al., 2018)	+	+?	+	+	+

This general balance between control and emergence could provide a foundation for the governance of GCRs and is compatible with our five necessary conditions. Those in power must *recognise* that perfect certainty and control are not achievable. Continuous monitoring and consequent action are necessary to maintain the dynamic balance between control and emergence, as is flexible, rapid decision-making. Cooperation between planners and those who are responsible for implementing plans is essential. And allowance must be made for the possibility of unexpected situations.

4.1.2 | Adaptive policies for handling deep uncertainty (Walker et al., 2010).

'Deep uncertainty' is defined as 'The condition in which analysts do not know or the parties to a decision cannot agree upon (1) the appropriate models to describe interactions among a system's variables, (2) the probability distributions to represent uncertainty about key parameters in the models, and/or (3) how to value the desirability of alternative outcomes' (Lempert et al., 2003).

The history of most, if not all, GCRs reveals that they fit this description. Policymakers have a choice of how to respond to it. Apart from burying their heads in the sand, or maintaining a belief in an overarching dogma and/or an ability to control, there appear to be three sensible (not necessarily exclusive) options (see Leusink & Zanting, 2009):

- Resistance: Plan for worst possible case or future situation.
- Resilience: Whatever happens, make sure you can recover quickly.
- Adaptation: Prepare to change the policy in case conditions worsen.

Adaptive policies provide the flexibility required by our necessary conditions. As discussed by Walker et al., they may be purposeful (planned adaptation, autonomous adaptation) or timed (anticipatory adaptation, reactive adaptation). In both cases, adaptive policies fit with our five necessary conditions. They recognise that perfect certainty and control are not achievable. By their very nature, they require integrated monitoring and action to allow flexible, rapid decision-making, and cooperation and coordination to implement those decisions over appropriate timescales. And they are able to incorporate investment in resilience and preparedness.

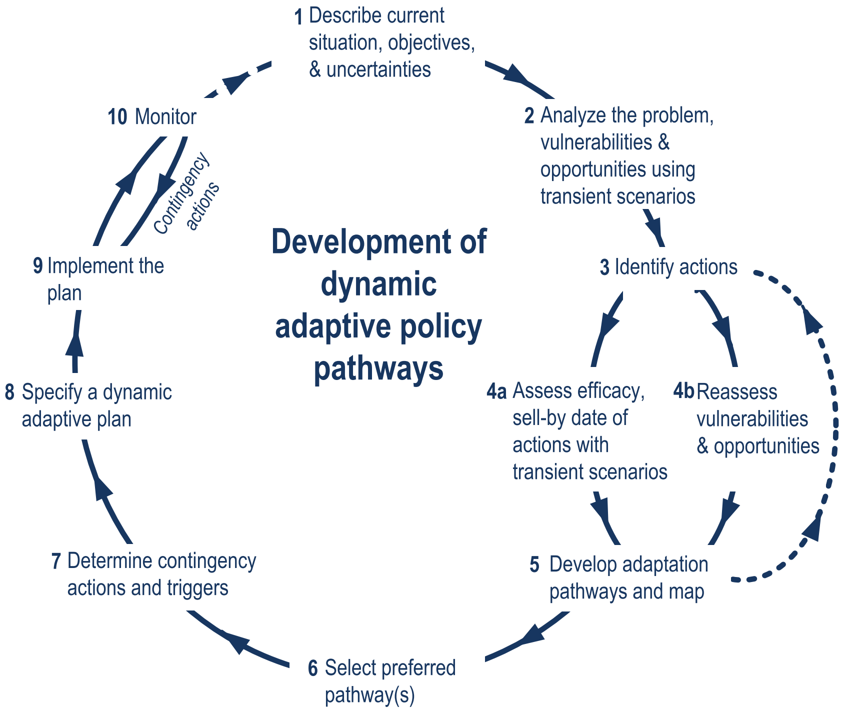

FIGURE 2 Development of dynamic adaptive policy pathways (from Haasnoot et al., 2013, with permission).

4.1.3 | Dynamic adaptive policy pathways (Haasnoot et al., 2013; Kwakkel et al., 2016).

Dynamic Adaptive Policy Pathways are a refinement of Walker et al.'s adaptive policies, incorporating the idea of a flexible strategic vision. They are based on the concept that, in light of deep uncertainties about the future, one needs to design dynamic adaptive plans. Such plans contain a strategic vision of the future, commit to short-term actions, and establish a framework to guide future actions. [They are] a fusion of adaptive policymaking and adaptation tipping points (Figure 2).

As the above diagram shows, they incorporate all of our necessary conditions, some directly (recognition [1, 2], monitoring and action [10], resilience and preparedness [7], and flexible, rapid decision-making [4a, 4b]), with cooperation and coordination being necessary for effective implementation of the whole process.

4.1.4 | Resilience thinking (Berkes, 2007; Folke, 2019; Folke et al., 2010)

Investment in resilience is one of our key conditions, but some authors believe that it can be taken further to form the foundation for governance of social–ecological systems. Here we examine whether this approach might also be appropriate to the governance of global catastrophic risks.

The underlying concept in resilience thinking is that of transformability across multiple scales. Resilience in this context (Folke et al., 2010) is 'the capacity of a SES (or any CAN [LRF & AS]) to continually change and adapt yet remain within critical thresholds. Adaptability is part of resilience. It represents the capacity to adjust responses to changing external drivers and internal processes and thereby allow for development along the current trajectory (stability domain). Transformability is the capacity to cross thresholds into new development trajectories. Transformational change at smaller scales enables resilience at larger scales. The capacity to transform at smaller scales draws on resilience from multiple scales, making use of crises as windows of opportunity for novelty and innovation, and recombining sources of experience and knowledge to navigate social–ecological transitions'.

Governance in this context consists of finding 'ways to foster resilience of smaller, more manageable SESs that contribute to Earth System resilience and to explore options for deliberate transformation of SESs that threaten Earth System resilience'.

A number of strategies have been proposed for enhancing resilience in complex adaptive systems (e.g. Crépin, 2019; Duit, 2015; Sellberg et al., 2018). These include fostering ecological, economic, and cultural diversity; planning for changes that are likely to occur; fostering learning; and communicating the societal consequences of recent changes. These strategies, and the basic concept, certainly fit our conditions (1) and (5). Remaining within stability domains also requires that conditions (2) and (3) be met. It is not so clear whether resilience thinking requires cooperation and coordination (one may envisage situations where built-in resilience through law or custom does not particularly require cooperation); nevertheless, we may consider resilience thinking to be a serious option for the governance of GCRs.

4.1.5 | Sensitive intervention points (Farmer et al., 2018)

Sensitive intervention points are points (in time, function, or place) where 'an intervention kicks or shifts the system so that the initial change is amplified by feedback effects that deliver outsized impact'.

Clearly, the use of SIPs for governance requires that our conditions (2)–(4) be met. Monitoring and subsequent action are obviously essential, as is flexible, rapid decision-making and cooperation and coordination on timescales compatible with the changes to be induced.

It is possible, however, to visualise a governance system whose leaders believe in the possibility of top–down control and predictability of outcomes, but who could nevertheless use SIPs as a tool for governance. Without the recognition of GCRs as CANs, however, the effectiveness of the interventions would be a matter of luck, and interventions could even backfire (as with the introduction of cane toads for pest control in Australian cane fields). Our condition (1), then, is not strictly necessary, but becomes highly desirable.

Governance solely by the use of SIPs does not strictly require investment in resilience and preparedness either (our condition [5]), but such investment is highly desirable on more general grounds.

Overall, SIPs offer a very useful tool that fits our conditions (2)–(4), but which may best be used to facilitate other approaches to the governance of GCRs, particularly in the implementation of dynamic adaptive pathways.

5 | CONCLUSIONS

We have established necessary and enabling conditions for the governance of global catastrophic risk and have examined a broad set of policy proposals in the light of these conditions. We find that **Adaptive Policies for Handling Deep Uncertainty**, as proposed by Walker et al., (2010), provide the most promising approach, with a **Balance Between Positive and Negative Feedbacks**, **Dynamic Adaptive Policy Pathways**, **Resilience Thinking**, and the use of **Sensitive Intervention Points** providing suitable enabling tools.

We are not aware of any existing governance system that fulfils these conditions, and argue that a totally new approach to the governance of global catastrophic risk is required. This must be based on the recognition of the nature of GCRs as CANs and of the known properties of CANs—especially, that they possess emergent properties that are more than the sum of their parts, and that they are liable to sudden, unpredicted (and often unpredictable), system-wide change.

We add here that there is one further practical question. This is that *enabling conditions* must be found which will facilitate transition to the new form of governance. These conditions are *processes* that must be possible within any governance system that fulfils the five necessary conditions.

Three processes are particularly important:

- The incorporation of 'bridging organisations' to connect governance levels and spatial and temporal scales (Folke, 2019)
- The evocation and maintenance of trust (Prieser & Woermann, 2019)
- Complexity leadership (Nooteboom & Teismann, 2019)

We will discuss these processes in detail, and whether they need to be modified for societies with different cultural values (Ruck et al., 2020), in a subsequent paper.

ACKNOWLEDGEMENTS

LF acknowledges with gratitude the continued support of the Head and staff of the School of Physics, University of Bristol, UK. For this research and manuscript preparation, AS received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 669751). This paper is also a contribution to the work of the Oxford Martin Programme on the Post-Carbon Transition.

DATA AVAILABILITY STATEMENT

Data sharing not applicable—no new data generated.

ENDNOTE

1. The American psychologist Frank Knight (1921) drew a distinction between *risk* ('decision situations in which probabilities are available to guide choice') and *uncertainty* ('decision situations in which information is too imprecise to be summarised by probabilities'; Runde, 1998). The risks that are encompassed in the phrase 'global catastrophic risks' might better be described in Knightian terms as 'uncertainties', because often we have no means of assessing their probabilities, or whether there are additional scenarios that we have not considered, or even been able to consider. The phrase 'global catastrophic risks' is, however, now firmly embedded in the literature, and we will stay with it, clarifying where necessary any ambiguity with the Knightian meaning.

REFERENCES

Abdenur, A.E. (2020) *Global governance and global catastrophic risks: Is the United Nations ready for the anthropocene?* Global Challenges Foundation Report.

Anjum, N.A. (2020) Good in the worst: COVID-19 restrictions and ease in global air pollution. Preprints 2020, 2020040069 https://doi.org/10.20944/preprints202004.0069.v1

Ashby, W.R. (1958) Requisite variety and its implications for the control of complex systems. *Cybernetica*, 1, 83–99.

Ashwin, P., Wieczorek, S., Vitolo, R. & Cox, P. (2012) Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society*, 370, 1166–1184.

Avin, S., Wintle, J.W., Weitzdörfer, J., Heigeartaigh, S.Ö., Sutherland, W.J. & Rees, M.J. (2018) Classifying global catastrophic risks. *Futures*, 107, 20–26.

Baldwin, R. & di Mauro, B.W. (Eds.) (2020) *Economics in the Time of COVID-19*. London: CEPR Press.

Bankes, S.C. (2002) Tools and techniques for developing policies for complex and uncertain systems. *Proceedings of the National Academy of Sciences of the USA*, 99, 7263–7266.

Basili, M. (2019) The CoVid-18 pandemic outbreak: The precautionary principle and institutional failures. *Mercato Concorrenza Regole, Società Editrice Il Mulino*, 3, 475–483.

Baum, S.D. & Barrett, A.M. (2018) Global Catastrophes: The Most Extreme Risks. In: Bier, V. (Ed.) *Risk in Extreme Environments: Preparing, Avoiding, Mitigating, and Managing*. New York, NY: Routledge, pp. 174–184.

Baum, S.D., Dekenberger, D.C., Pearce, J.M., Robock, A. & Winkler, R. (2015) Resilience to global food supply catastrophes. *Environment, Systems and Decisions*, 35, 301–313.

Baum, S.D. & Handoh, I.C. (2014) Integrating the planetary boundaries and global catastrophic risk paradigms. *Ecological Economics*, 107, 13–21.

Beinhocker, E.D. (2013) Reflexivity, complexity, and the nature of social science. *Journal of Economic Methodology*, 20, 330–342.

Bellman, R.E. (1961) *Adaptive control processes: a guided tour*. Princeton: Princeton University Press (republished in Princeton Legacy Library series (https://doi.org/10.1515/9781400874668))

Berkes, F. (2007) Understanding uncertainty and reducing vulnerability: lessons from resilience thinking. *Natural Hazards*, 41, 283–295.

Boettiger, C. & Hastings, S. (2012) Early warning signs and the prosecutor's fallacy. *Proceedings of the Royal Society*, B279, 4734–4739.

Bognar, G. (2011) Can the maximin principle serve as a basis for climate change policy? *The Monist*, 94, 329–348.

Bostrom, N. (2013) Existential risk prevention as global priority. *Global Policy*, 4, 15–31.

Bostrom, N. & Ćirković, M.M. (2008) *Global catastrophic risks*. Oxford: Oxford University Press.

Brosig, M. (2019) Restricted complexity: a middle path between post-modern complexity theory and positivist mainstream IR, In: Orsini, A., Le Prestre, P., Haas, P.M., Brosig, M., Pattberg, P., Widerberg, O., Gomez-Mera, L. et al. (Eds.) *Complex systems and international governance. International Studies Review*, pp. 1015–1019. https://doi.org/10.1093/isr/viz005

Burch, S., Gupta, A., Inoue, C.Y.A., Kalfagianni, A., Persson, Å., Gerlak, A.K. et al. (2019) New directions in earth system governance research. *Earth System Governance*, 1, 100006.

Carpenter, S.R., Arrow, K.J., Barrett, S., Biggs, R., Brock, W.A., Crépin, A.-S. et al. (2012) General resilience to cope with extreme events. *Sustainability*, 4, 3248–3259. https://doi.org/10.3390/su4123248.

Centeno, M.A., Nag, M., Patterson, T.S., Shaver, A. & Windawi, A.J. (2015) The emergence of global systemic risk. *Annual Review of Sociology*, 41(1), 65–85.

Chan, S. (2001) *Complex adaptive systems*, ESD.83 Research Seminar in Engineering Systems. Available from: https://stuff.mit.edu/afs/athena/course/other/esd.83/OldFiles/www/notebook/Complex%20Adaptive%20Systems.pdf

Choi, T.Y., Dooley, K.J. & Rungtusanatham, M. (2001) Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, 19, 351–366.

Clarke, S. (2005) Future technologies, dystopic futures and the precautionary principle. *Ethics and Information Technology*, 7, 121–126.

Clemens, W.C. (2013) *Complexity Science and World Affairs*. New York: SUNY Press.

Cotton-Barratt, O., Daniel, M. & Sandberg, A. (2020) Defence in Depth Against Human Extinction: Prevention, response, Resilience, and Why They All Matter. *Global Policy*, 11, 271–282.

Crépin, A.-S. (2019) Complexity, resilience and economics. In: Galaz, V. (Ed.) *Global challenges, governance, and complexity: applications and frontiers*. Cheltenham: Edward Elgar Publishing Ltd., pp. 166–x.

Dawson, R. (2012) The law of requisite variety: Why flexibility and adaptability are essential for success. *Blog Post* https://rossdawson.com/blog/the-law-of-requisite-variety-why-flexibility-and-adaptability-are-essential-for-success/

Duit, A. (2015) Resilience thinking; lessons for public administration. *Public Administration*, 94, 364–380. https://doi.org/10.1111/padm.12182

Duit, A. & Galaz, V. (2008) Governance and complexity – emerging issues for governance theory. *Governance: an International Journal of Policy, Administration and Institutions*, 21, 311–335.

Faber, M.H. (2011) On the governance of global and catastrophic risks. *International Journal of Risk Assessment and Management*, 15(5/6), 400. https://doi.org/10.1504/IJRAM.2011.043698

Faranda, D., Alberti, T., Arutkin, M., Lembo, V. & Lucarini, V. (2021) Interrupting vaccination policies can greatly spread SARS-CoV-2 and enhance mortality from COVID-19 disease: The AstraZeneca case for France and Italy, *Chaos* 31, 41105. https://doi.org/10.1063/5.0050887

Farmer, J.D., Hepburn, C., Ives, M.C., Hale, T., Wetzer, T., Meay, P. et al. (2018) Sensitive intervention points in the transition to the post-carbon society. *Science*, 364, 132–134.

Faulkner, P.A. (1963) Castle planning in the fourteenth century. *Archaeological Journal*, 120, 215–235.

Fisher, L. (2015) More than 70 ways to show resilience. *Nature*, 518, 35.

Fisher, L. (2019) Global challenges: personal reflections on the Stockholm "New Shape" competition. *Journal & Proceedings of the Royal Society of New South Wales*, 151, 213-231.

Folke, C. (2019) Governing for emergence in social-ecological systems. In: Galaz, V. (Ed.) *Global challenges, governance, and complexity: applications and frontiers*. Cheltenham: Edward Elgar Publishing Ltd., pp. 24–37.

Folke, C., Carpenter, S.R., Walker, B., Scheffer, M., Chapin, T. & Rockström, J. (2010) Resilience thinking: integrating resilience, adaptability and transformability. *Ecology and Society*, 15, 20–28.

Galaz, V. (2019) *Global challenges, governance, and complexity: applications and frontiers*. Cheltenham: Edward Elgar Publishing Ltd.

Galaz, V., Österblom, H., Bodin, Ö. & Crona, B. (2016) Global networks and global change-inducing tipping points. *International Environmental Agreements*, 16, 189–221.

Garrett, L. (2020) COVID-19: the medium is the message. *The Lancet*, 395, 942–943.

Gigerenzer, G., Todd, P.M. & the ABC Research Group (2001) *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.

Global Challenges Foundation. (2020) *Global Catastrophic Risks 2020*. https://globalchallenges.org/wp-content/uploads/Global-Catastrophic-Risks-2020-Annual-Report.pdf

Haas, P. (2019) IR Encounters Complexity. In: Orsini, A., Le Prestre, P., Haas, P.M., Brosig, M., Pattberg, P. & Widerberg, O. et al. (Eds.) *Complex Systems and International Governance*, International Studies Review, 1–30. https://doi.org/10.1093/isr/viz005

Haasnoot, M., Kwakkel, J., Walker, W.E. & ter Maat, J. (2013) Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, 23, 485–498.

Haldane, A. (2017) *Rethinking financial stability*. Speech delivered to the conference Rethinking Macroeconomic Policy IV, Peterson Institute for International Economics, Washington, DC, USA. https://www.bankofengland.co.uk/speech/2017/rethinking-financial-stability

Haleem, A., Javaid, M. & Vaishya, R. (2020) Effects of COVID 19 pandemic in daily life. *Current Medicine Research and Practice*, 10(2), 78–79. https://doi.org/10.1016/j.cmrp.2020.03.011

Hao, K. (2020) The UK exam debacle reminds us that algorithms can't fix broken systems. *MIT Technology Review*. https://www.technologyreview.com/2020/08/20/1007502/uk-exam-algorithm-cant-fix-broken-system/

Harrison, N. & Geyer, R. (2019) The complexity of governance of climate change. In: Orsini, A., Le Prestre, P., Haas, P.M., Brosig, M., Pattberg, P., Widerberg, O., et al. (Eds.) *Complex systems and international governance*. International Studies Review, pp. 1–30. https://doi.org/10.1093/isr/viz005

Helbing, D. (2013) Globally networked risks and how to respond. *Nature*, 497, 51–59.

IPBES. (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. https://www.ipbes.net/system/tdf/spm_global_unedited_advance.pdf?file=1&type=node&id=35245

IPCC (2021) *Climate Change 2021: The Physical Science Basis*. https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/

Kareiva, P. & Fuller, E. (2016) Beyond resilience: how to better prepare for the profound disruption of the Anthropocene. *Global Policy*, 7, 107–118.

Keeling, M.J., Hollingsworth, T.D. & Read, J.M. (2020) The Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). *MedRxiv Preprint*. https://doi.org/10.1101/2020.02.14.20023036

Keohane, R.O. (2001) Governance in a partially globalized world. *American Political Science Review*, 95, 1–13.

Kim, R.E. & Mackey, B. (2014) International environmental law as a complex adaptive system. *International Environmental Agreements*, 14, 5–24.

Klinke, A. (2014) Postnational discourse, deliberation, and participation toward global risk governance. *Review of International Studies*, 40, 247–275.

Knight, F.H. (1921) *Risk, uncertainty and profit*. Chicago: University of Chicago Press.

Kooiman, J. (2003) *Governing as governance*. London: SAGE Publications.

Kruglanski, A.W. & Orehek, E. (2012) The need for certainty as a psychological nexus for individuals and society. In: Hogg, M.A. & Blaylock, D.L. (Eds.) *Extremism and the psychology of uncertainty*. London: Blackwell.

Kuhlemann, K. (2019) Complexity, creeping normalcy and conceit: sexy and unsexy risks. *Foresight*, 21, 35–52.

Kwakkel, J., Haasnoot, M. & Walker, W. (2016) Coping with the wickedness of public policy problems: Approaches for decision-making under deep uncertainty. *Journal of Water Resources Planning and Management*, 142, 1–5. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000626

Lagi, M., Bertrand, K.Z. & Bar-Yam, Y. (2011) The food crises and political instability in North Africa and the Middle Eeast. rXiv:1108.2455v1 [physics.soc-ph] 11 Aug 2011.

Lempert, R.J. (2002) A new decision sciences for complex systems. *Proceedings of the National Academy of Sciences of the USA*, 99, 7309–7313.

Lempert, R., Popper, S. & Bankes, S. (2003) *Shaping the next one hundred years: New methods for long-term policy analysis*. MR-1626-RPC, The RAND Pardee Center.

Lenton, T.M. (2013) What early warning signs are there for environmental shocks? *Environmental Science & Policy*, 27S, S60–S75.

Leusink, A. & Zanting, H.A. (2009) Towards a trade-off framework for climate proofing the Netherlands, with experience from 4 case studies. Available from: http://edepot.wur.nl/15291 (in Dutch)

Levin, S. (2019) The architecture of robustness. In: Galaz, V. (Ed.), *Global challenges, governance, and complexity: applications and frontiers*. Cheltenham: Edward Elgar Publishing Ltd, p. 18.

Levin, S., Xepapadeas, T., Crépin, A.-S., Norberg, J., DeZeew, A., Folke, C. et al. (2013) Social-ecological systems as complex adaptive systems: modeling and policy implications. *Environment & Development Economics*, 18, 111–132.

Louaas, A. & Picard, P. (2020) Optimal insurance coverage of low-probability catastrophic risks, hal-02875534. Available from: https://hal.archives-ouvertes.fr/hal-02875534/file/Insurability%20of%20low%20probability%20catastrophic%20risks.pdf

Luby, S. & Arthur, R. (2019) Risk and response to biological catastrophe in lower income countries. In: *Global catastrophic biological risks*, Berlin/Heidelberg: Springer. pp. 85–105.

Luscombe, R. (2020) Florida: endangered sea turtles thrive thanks to Covid-19 restrictions. *The Guardian,* 19 April. https://www.theguardian.com/us-news/2020/apr/19/florida-leatherback-turtles-coronavirus-beaches?utm_source=dlvr.it&utm_medium=twitter

Maier, H.R., Guillaume, J.H.A., van Delden, H., Riddell, G.A., Haasnoot, M. & Kwakkel, J.H. (2016) An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling and Software*, 81, 154–164.

Makridakis, S. & Taleb, N. (2009) Decision making and planning under low levels of predictability. *International Journal of Forecasting*, 25, 716–733.

Manheim, D. (2016) Goodhart's law and why measurement is hard. *Ribbonfarm.* https://www.ribbonfarm.com/2016/06/09/goodharts-law-and-why-measurement-is-hard/.

Manheim, D. (2018) *Building less flawed metrics*. MPRA Paper No. 90649, posted 21 Dec 2018. Available from: https://mpra.ub.uni-muenchen.de/98288/

Mason, C.F., Polasky, S. & Tarui, N. (2017) Cooperation on climate-change mitigation. *European Economic Review*, 99, 43–55.

Mattoo, A. & Subramanian, A. (2013) *A "Greenprint" for international cooperation on climate change*. World Bank Development Research Group Policy Research Working Paper 6440. Available from: https://openknowledge.worldbank.org/bitstream/handle/10986/15581/wps6440.pdf;sequence=1

May, R.M. & Haldane, A.G. (2011) Systemic risk in banking ecosystems. *Nature*, 469, 351–355.

McConnell, A. & Stark, A. (2021) Understanding policy responses to COVID-19; the stars haven't fallen from the sky for scholars of public policy. *Journal of European Public Policy*, 28, 1115–1130.

Miller, J.H. & Page, S. (2007) *Complex adaptive systems*. Princeton & Oxford: Princeton University Press.

Newton, A.C. (2011) Implications of Goodhart's Law for monitoring global biodiversity loss. *Conservation Letters*, 4, 264–268.

Nooteboom, S. & Teisman, G. (2019) Leadership and complexity: can individuals make differences in complex systems? In: Galaz, V. (Ed.) *Global challenges, governance, and complexity: applications and frontiers*. Cheltenham: Edward Elgar Publishing Ltd., pp. 239–273.

Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K.E.C., Baguelin, M., Bhatt, S. et al. (2021) Reduction in mobility and COVID-19 transmission. *Nature Communications*, 12, 1090. https://doi.org/10.1038/s41467-021-21358-2

Ord, T. (2020) *The precipice*. Oxford: Oxford University Press.

Ostrom, E. (2009) A general framework for nalysing sustainability of social-ecological systems. *Science*, 325, 419–422.

Pattberg, P. & Widerberg, O. (2019) Studying global governance as a complex system: a network perspective. In: Orsini, A., Le Prestre, P., Haas, P.M., Brosig, M., Pattberg, P. et al. (Eds.), *Complex systems and international governance. International studies review*, pp. 1–30. https://doi.org/10.1093/isr/viz005

Pegram, T. & Kreienkamp, J. (2019) *Governing complexity: design principles for improving the governance of global catastrophic risks*. Global Governance Institute Research Policy Brief. London: University College London.

Pellis, L., Scarabel, F., Stage, H.B., Overton, C.E., Chappell, L.H.K., Lythgoe, K.A., Fearon, A., Bennett, E., Curran-Sebastien, J., Das, R., Fyles, M., Lewkowicz, H., Pang, X., Vekaria, B., Webb, L., House, T. & Hall, I. (2020) Challenges in control of Covid-19: Short doubling time and long delay to effect of interventions. arXiv:2004.00117v1 [q-bio.PE] 31 Mar

Peters, B.G., Pierre, J., Galaz, V. (2019) Simple solutions for complexity? In: Galaz, V. (Ed.), *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar Publishing Ltd., pp. 188–210.

Pohl, A. (1999) Some notions of complex adaptive systems and their relationship to our world, InterSymp-99: Advances in Collaborative Decision-Support Systems for Design, Planning and Execution (Baden-Baden, Germany, Aug 2–7), 9–24.

Polasky, S., Carpenter, S.R., Folke, C. & Keeler, B. (2011) Decision-making under great uncertainty: environmental management in an era of global change. *Trends in Ecology and Evolution*, 26, 398–404.

Prieser, R., Woermann, M. (2019) Complexity, philosophy and ethics. In: Galaz, V. (Ed.), *Global challenges, governance, and complexity: applications and frontiers*. Cheltenham: Edward Elgar Publishing Ltd., pp. 38–62.

Rao, V. (2010) A big little idea called legibility. Available from: https://www.ribbonfarm.com/2010/07/26/a-big-little-idea-called-legibility/

Read, R. & O'Riordan, T. (2017) The precautionary principle under fire. *Environment: Science and Policy for Sustainable Development*, 59, 4–15.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F.S.I.I.I., Lambin, E. et al. (2009) Planetary boundaries: exploring the safe operating space for humanity. *Ecology and Society*, 14(2), https://doi.org/10.5751/ES-03180-140232

Ruck, D.J., Matthews, L.J., Kyritsis, T., Atkinson, Q.D. & Bentley, R.A. (2020) The cultural foundations of modern democracy. *Nature Human Behaviour*, 4, 265–269.

Runde, J. (1998) Clarifying Frank Knight's discussion of the meaning of risk and uncertainty. *Cambridge Journal of Economics*, 22, 539–546.

Sandberg, A. (2019) There is plenty of time at the bottom: The economics, risk and ethics of time compression. *Foresight*, 21, 84–99. https://doi.org/10.1108/FS-04-2018-0044

Sayama, H., Pestov, I., Schmidt, J., Bush, B.J., Wong, C., Yamanoi, J. et al. (2013) Modeling complex systems with adaptive networks. *Computers and Mathematics with Applications*, 65, 1645–1664.

Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V. et al. (2009) Early-warning signs for critical transitions. *Nature*, 461, 53–59.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A. & White, D.R. (2009) Economic networks; the new challenges. *Science*, 325, 422–425.

Scott, J. (1999) *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale: Yale University Press.

Sellberg, M.M., Ryan, P., Borgström, S.T., Norström, A.V. & Peterson, G.D. (2018) From resilience thinking to resilience planning: Lessons from practice. *Journal of Environmental Management*, 217, 906–918.

Sevastopulo, D. & Shubber, K. (2020) Trump cheers as anti-lockdown protests spread. *Financial Times*, Available from: https://www.ft.com/content/c8f6f413-39c4-47ce-b1ff-0e02969cb612

Soros, G. (2013) Fallibility, reflexivity, and the human uncertainty principle. *Journal of Economic Methodology*, 20, 309–329.

Sunstein, C.R. (2007) The catastrophic harm precautionary principle, *Issues in Legal Scholarship Symposium: Catastrophic Risks: Prevention, Compensation and Recovery*.

Sunstein, C.R. (2019) Maximin. *Harvard Law School Discussion Paper* 1020. Available from: https://ssrn.com/abstract=3476250

Turchin, A. (2018) Approaches to the prevention of global catastrophic risks. *Human Prospect*, 7(2), 53–65.

Victor, D. (2006) Toward effective international cooperation on climate change: Numbers, interests and institutions. *Global Environmental Politics*, 6, 90–103.

Victor, D. (2011) *Global warming gridlock: creating more effective strategies for protecting the planet*. Cambridge: Cambridge University Press.

Victor, D.G. (2016) What the framework convention on climate change teaches us about cooperation in climate change. *Politics and Governance*, 4, 133–141.

Voeneky, S. (2019) Human rights and legitimate governance of existential and global catastrophic risks. Available from: https://ssrn.com/abstract=3363552

Wachinger, G., Renn, O., Begg, C. & Kuhlicke, C. (2013) The Risk perception paradox–Implications for governance and communication of natural hazards. *Risk Analysis*, 33, 1049–1065.

Walker, W.E., Marchau, V.A.W.J. & Swanson, D. (2010) Addressing deep uncertainty using adaptive policies. *Technological Forecasting & Social Change*.

Wilett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S. et al. (2019) Food in the Anthropocene: the EAT-Lancet commission on healthy diets from sustainable food systems. *The Lancet Commissions*, 393, 447–492.

World Economic Forum (2020) *The Global Risks Report*, 15th edition. Available from: http://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf

Young, O. (2017) *Governing complex systems. Social capital for the anthropocene*. Cambridge, MA: MIT Press.

AUTHOR BIOGRAPHIES

Len Fisher is an Honorary Senior Research Fellow at the University of Bristol, UK, He is a researcher, author, and broadcaster with a background in fundamental physics and its application to biological questions. He is now primarily concerned with using and promoting science to understand and help resolve global risks of all kinds.

Anders Sandberg is Senior Research Fellow at the Future of Humanity Institute at University of Oxford. He is also Research Associate at the Oxford Uehiro Centre for Practical Ethics and the Institute for Futures Studies in Stockholm. His background is computational neuroscience, neuroethics, futures studies, and existential risk research.

How to cite this article: Fisher, L. & Sandberg, A. (2022) A Safe Governance Space for Humanity: Necessary Conditions for the Governance of Global Catastrophic Risks. *Global Policy*, 13, 792–807. Available from: https://doi.org/10.1111/1758-5899.13030