MAIN ARTICLE

Challenges for sustainability: misperceptions and misleading advice

Erling Moxnes*

Abstract

To ensure sustainable development, governments depend on informed decision-makers including the electorates. Previous studies show evidence of widespread and systematic misperceptions, voter ignorance, and reliance on inappropriate cognitive heuristics. The wait and see heuristic is one such trusted heuristic that is used repeatedly and with minimal effort. When applied to the management of dynamically complex renewable resources, it leads to overinvestment and resource depletion. On the optimistic side, a laboratory experiment finds that proactive expert advice counteracts the wait and see heuristic and contributes to sustainability. However, when expert advice is challenged by misleading advice that supports the wait and see heuristic, most of the positive effect of the expert advice disappears. The experiment generalizes to a large number of sustainability problems. Different from fake news that can be corrected by simple facts, misperceptions and misleading advice call for extraordinary information policies and deliberative democracy.

Syst. Dyn. Rev. **39**, 185–206 (2023)

Introduction

People's preferences for sustainability in terms of stability (Moxnes, 2014) call for proper management of renewable resources to prevent overutilization and collapse. Still, a large number of renewable resources have been overutilized causing much harm to those who depend on these resources. Much evidence points to misperceptions of dynamic systems as a major cause of mismanagement in general (Sterman, 2011) and of renewable resources in particular (Moxnes, 1998a, 1998b, 2004; Sterman and Sweeney, 2007; Sterman, 2008; Moxnes and Saysel, 2009; Guy *et al*., 2013). Of particular concern is the wait and see heuristic, which is used to simplify complex problem solving. Using this heuristic, decision-makers rely on outcome feedback to take actions. In doing so, they ignore essential system-knowledge about the time and effort needed to correct problems.

Previous laboratory experiments have provided subjects with essential system-knowledge. The results for renewable resource management show only limited improvements in performance (Moxnes, 1998a, 1998b; Rouwette *et al*., 2004; Sterman and Sweeney, 2007; Moxnes and Jensen, 2009;

Department of Geography, University of Bergen, Bergen, Norway

* Correspondence to:Erling Moxnes, System Dynamics Group, Department of Geography, University of Bergen, Bergen, Norway. E-mail: erling.moxnes@uib.no

Accepted by James Duggan, Received 9 January 2023; Revised 12 April 2023; Accepted 19 April 2023

System Dynamics Review
System Dynamics Review vol 39, No 3 (July/September 2023): 185–206
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/sdr.1733

Moxnes and Saysel, 2009). Apparently, most subjects do not have the needed background to make use of such knowledge.

An alternative is to provide expert advice about what decisions to make. To mimic public debates, a new renewable resource experiment has been designed where both proper expert advice and misleading advice are provided. The existence of misleading advice is well-known (Oreskes and Conway, 2010) and intentionally misleading advice is well documented (Supran *et al*., 2023). In the new experiment, misleading advice builds on and supports the wait and see heuristic. Supported by this heuristic, misleading policy advice is likely to be harder to correct than fake news, which can often be corrected by simple facts.

The results of the experiment may generalize to many dynamically similar renewable resources. While expert advice is found to be very effective in isolation, misleading advice counteracts most of its effect. This paper concludes by discussing information policies to reduce the effects of misperceptions and misleading advice, and to strengthen electorate participation, information sharing, transparency, and use of available tools, i.e. deliberative democracy (Dryzek *et al*., 2019).

A generic model of renewable resource systems

To set the stage for the experiment, the diagram in Figure 1 presents a generic model of a renewable resource system. While it may look overly simplified, it captures the essence of the management problem. The lower stock represents the limiting *natural resource*, and the upper stock denotes the *capacity to utilize* this resource. Using a well-established reindeer experiment (Moxnes, 2004), the limiting resource is lichen. Reindeer represent the capacity to utilize the resource. In addition to the well-known biological cause and effect relationships, the diagram illustrates the wait and see policy by a thickened arrow from Lichen to Slaughtering.

The complexity of management lies in the stock nature of lichen and reindeer. The herd size can increase for many years before scarcity of the perennial lichen is perceived to become a problem. When scarcity becomes evident, the wait and see heuristic suggests that the herd size should be stabilized at the current high level, to see if lichen stabilizes as well. However, lichen continues to decrease as long as grazing exceeds the growth of lichen. Then the wait and see heuristic suggests further gradual (and insufficient) reductions in grazing. A frustrating situation occurs where fewer reindeer appear to cause lichen to decrease as well. This is contrary to what is expected from an erroneous mental model where reductions in the number of reindeer are expected to increase lichen instantaneously (Moxnes, 2000). What is needed is a drastic and early reduction in the herd size to bring

DOI: 10.1002/sdr

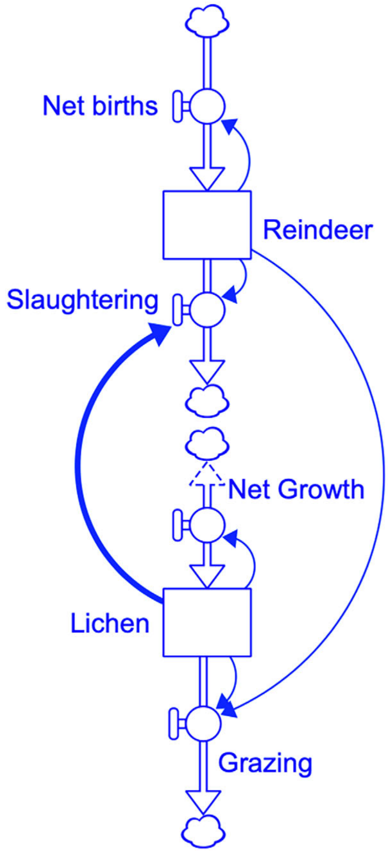

Fig. 1. Stock and flow diagram showing how reindeer and lichen pastures interact [Color figure can be viewed at wileyonlinelibrary.com]

grazing below the growth rate of lichen. That is the advice of the expert in the experiment.

Laboratory experiments and real data show that slaughtering does not increase before lichen is perceived to becoming scarce (Moxnes, 2004). Figure 2 shows an example of mismanagement for the Snøhetta area in Norway. After Hardin's (1968) article "The Tragedy of the Commons" it has been commonplace to blame resource depletion on the commons problem. However, there was no commons problem in this case. Total slaughtering was already regulated by the Department of Agriculture. Hence, there was only a management problem, namely for the Department to set yearly hunting quotas (slaughtering).

In 1952, locals called for an assessment of the lichen conditions and asked for larger hunting quotas (Jordhøy, 2001). However, the Department wanted to increase the herd size and allowed for only a small increase in legal hunting, contrary to the advice from biologists. The herd size continued to grow

DOI: 10.1002/sdr

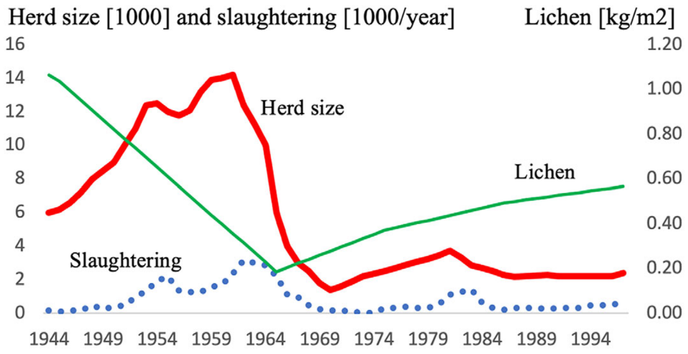

Fig. 2. Slaughtering and number of reindeer (Jordhøy, 2001). Lichen is interpolated between imprecise and infrequent data points (Moxnes *et al*., 2003) [Color figure can be viewed at wileyonlinelibrary.com]

until 1961. After that, the herd size fell rapidly due to increased hunting quotas, outmigration, deaths, and illegal hunting. Hunting licenses increased much too late and too little to prevent the collapse of reindeer and lichen. In many localities, the lichen never recovered over the 50-year period shown in the graph. Reported slaughtering never reached as high as the maximum sustainable yield (MSY) at well-managed lichen pastures.

Scheffer (1951) documents a similar example of overshoot and collapse after reindeer were introduced to the island of St. Paul in Alaska in 1911. Again, there was no commons problem, only one herd. Grazing by the reindeer increased to three times the MSY before lichen was fully depleted and the herd collapsed. Again, slaughtering was increased too little and too late to prevent the collapse. Scheffer (1951, p. 360) quotes the American Society of Mammologists as urging the need, *prior* to introducing reindeer into a new area, to study thoroughly “the problems of integrating lichen ecology, reindeer biology, and native culture—serious problems that have not been solved to date on any workable scale on the North American continent.”

A reindeer experiment with expert advice

The experiment by Moxnes (2004) replicates the observations from Snøhetta and St. Paul. The very same experiment is well suited to test the effects of different advice for two main reasons. First, real data and the original experiment show that there is a need for expert advice. Second, the available subject pool has no prior information about reindeer and lichen biology, reindeer management, and existing policy proposals. Lack of prior knowledge is important in order to investigate the uncontaminated effects of expert advice and of misleading advice.

The central hypothesis is that people behave according to a wait and see heuristic and delay actions until they see a need to act. The design of the experiment rules out five alternative hypotheses for unsustainable resource

DOI: 10.1002/sdr

management. First, subjects' potential lack of care about the future is overruled by an explicit goal of reaching and maintaining the harvesting of the MSY as quickly as possible. Second, the commons problem is removed by giving each subject private property rights. Third, the possibility of "hierarchical individualists" opposing governmental interference (Kahan *et al.*, 2012) is also prevented by assuming private property rights. Fourth, potentially detrimental effects of information delays are removed by providing yearly updates of lichen data. Fifth, there is no randomness; information is perfect with no need to delay actions while filtering received data over time.

The decisions that subjects make in the experiment are likely to reveal policy preferences better than questionnaires and polls. Answers to polls may be biased towards what is expected or what is politically correct rather than reflecting subjects' true feelings. Therefore, decisions may give a better indication of the electorate's voting behavior. This is supported by Ansolabehere and Jones (2010, p. 596) who find that voters have preferences over important bills and have fairly accurate beliefs about their legislators' roll-call votes. They conclude that "the extent to which a constituent agrees with the policy positions of the member of Congress strongly affects the constituent's approval rating of the member and likelihood of voting for the member." Hence, if subjects' decisions change as a result of advice, voting behavior is also likely to be influenced.

The experiment (Moxnes, 2004) makes use of a discrete-time, one-stock simulator with the following stock-equation for the lichen, measured by the average height of the plants:

$$\begin{aligned} L_{t+1} &= L_t + G_t - H_t, \\ L_0 &= 24.4\,\mathrm{mm}, \end{aligned} \tag{1}$$

where time t moves in steps of one year and where L_0 is the initial average height of lichen. The yearly net growth of the lichen G_t is given by a surplus growth curve:

$$G_t = M\left\{1 - \left[\frac{L_t - C/2}{C/2}\right]^2\right\}, \tag{2}$$

where the carrying capacity is $C = 60$ mm. The MSY, the maximum yearly average growth of the lichen, is $M = 5$ mm/year. The growth is zero when $L_t = 0$ and when $L_t = C$. In between these extremes, the maximum occurs when $L_t = C/2$. Reindeer grazing per year,

$$H_t = hR_t \tag{3}$$

DOI: 10.1002/sdr

is given by the yearly grazing per animal, $h = 0.004$ mm/reindeer/year, and by the yearly number of reindeer R_t decided by the subjects. The number of reindeer is not controlled by slaughtering as indicated in Figure 1. As demonstrated earlier (Moxnes, 2004), while allowing R_t to be set freely simplifies the management problem, this still leads to mismanagement.

The starting point for the experiment is similar to the 1952 conditions in the Snøhetta district shown in Figure 2. This was the year when local hunters started to worry about the lichen conditions. The subjects get information about the historical development over the preceding 15 years. This announced historical development is similar to the historical development in the Snøhetta district before 1952, except that the announced development is adjusted to be fully consistent with the model. Subjects get precise information about the fixed grazing per reindeer (h) and they get a verbal description of the function for growth $G_t(L_t)$. Each year over a 15-year time horizon, participants receive yearly information about the amount of lichen L_t at the beginning of the year and then make decisions about the number of reindeer R_t for the rest of the year. Subjects do not get information about the exact maximum sustainable herd size (R^*) and the corresponding lichen level (L^*): $R^* = M/h = 1250$ reindeer, and $L^* = C/2 = 30\,\text{mm}$. These general instructions are identical to those of the original experiment's treatment 1 (Moxnes, 2004).

While the provided historical data are sufficient to estimate R^* and L^*, the subjects were not expected to do this (and none seem to have done so). However, a prior understanding of the dynamic system is sufficient to get close to the goal: to reach the highest possible herd size that can last forever without depleting the lichen pasture, and to do this as quickly as possible. In each round of the experiment, the subject obtaining the best result was promised and awarded a symbolic prize (Bolle, 1990).

The simulator was programmed using the Macro function in Excel and subjects received treatment information on paper. For their own sake and as a backup, subjects were told to fill in yearly decisions and other information in an empty table as they made progress.

The reindeer system generalizes to other two-stock resources

The model in Figure 1 and the experiment generalize to other natural resources and a host of other stock-management systems (see Table 1 in Sterman, 1989). For instance, in Figure 1, replace the upper stock by fishing capacity and the lower stock with fish. Then the model explains over-investment in fishing capacity and depletion of fish resources. Over-investment and over-fishing have been the rule rather than the exception in real fisheries (Jackson *et al.*, 2001; Schrank, 2003). Figure 3 shows the same depletion dynamics for a herring fishery as for lichen in Figure 2. Catch reflects capacity in the period when the size of the fish stock is not limiting

DOI: 10.1002/sdr

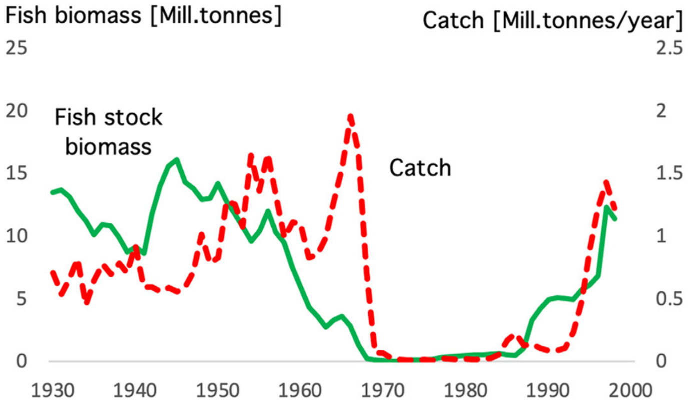

Fig. 3. Development of fish stock and catch (reflecting capacity) for Norwegian Spring Spawning Herring (Toresen and Østvedt, 2000) [Color figure can be viewed at wileyonlinelibrary.com]

the catch. These dynamics have been replicated in a fishery experiment with property rights (Moxnes, 1998a).

Replace the upper stock with the capacity to use water and the lower stock with the amount of water in an underground water reservoir. Consistent with this system structure, depletion of aquifers has become a problem globally (Konikow and Kendy, 2005).

Replace the upper stock by the capacity to emit greenhouse gases (GHGs) and the lower stock by the amount of GHGs in the atmosphere. Experiments show that even highly educated subjects rely on pattern matching (Sterman and Sweeney, 2007), instantaneous mental models (Moxnes and Saysel, 2009), and on explicitly stated wait and see policies (Guy *et al.*, 2013). The experiments are consistent with observations of opposition to proactive policies to limit climate change (Leiserowitz, 2007; Dabla-Norris *et al.*, 2023).

Information treatments

To explore the effects of advice, the subjects are supplied with narratives. By testing how the use of narratives "relates to policy outcomes" (Jones and McBeth, 2010, p. 345), the experiment represents a novel quantitative test of the effect of narratives.

The experiment departs from previous experiments (Moxnes, 1998a, 2004; Sterman and Sweeney, 2007; Sterman, 2011; Guy *et al.*, 2013) where subjects get information only about the system (theory) and the historical development (data). As already mentioned, such information has been found to have limited effects on outcomes. Here, subjects get advice about what *actions* to take, in addition to the general system information and historical data. The advice does not say what the MSY is. The only quantitative advice is about

DOI: 10.1002/sdr

how much to reduce the number of reindeer in the first year. This advice is supposed to mimic simple advice that could be given in news programs, and which is not updated when the problem is no longer "breaking news."

The experiment has three categories of information treatments:

1. General instructions with no advice
2. General instructions with expert advice or misleading advice
3. General instructions with expert advice combined with misleading advice

None of the pieces of advice are intended to be representative of any particular group of advisors. Rather the narratives are meant to capture what distinguishes different types of advice. Each of the narratives is followed by a piece of quantitative advice for the first year's herd size R_0. The expert advice is given in two versions, either as an analyst's or an activist's narrative. They represent two broad categories of advice and it seems important not to miss out on any of them. Also, it seems interesting for change agents to learn about the effects of the two types of expert advice. Both types of advice are followed by the very same advice of reducing R_0 to zero. Even if this is an item of advice only for the first year, it signals that a drastic reduction in reindeer is needed immediately. The misleading advice supports the wait and see heuristic and recommends stabilizing R_0 at the level that was reached at the end of the historical period.

Misleading (announced as Expert no. 1 when combined with expert advice)
The previous owner has built up the herd size carefully over the historic period. The data show that the herd size has been increased to the current level without serious problems for the lichen and for the reindeer. Thus, drastic reductions in the herd size are not called for and would only bring the herd much below the sustainable level. In light of the considerable uncertainty about what adjustments are needed, a very careful trial-and-error approach should be applied when making adjustments in the herd size.
Concrete first advice: In the first few years the herd size should be kept at 1800 animals, then one should observe the development before possible further adjustments to reach the maximum sustainable herd are considered.

Analyst (announced as Expert no. 2 when combined with misleading advice)
Since the amount of lichen has decreased steadily under the previous owner, removal by reindeer grazing must have been greater than the natural growth of lichen in all previous years. Thus, in year zero in the historic period, the lichen growth rate must have been smaller than what is needed to feed 1150 animals. To increase the amount of lichen to a level that gives the maximum lichen growth, grazing must be reduced below the current growth rate for a while, then increased towards the growth rate again to stabilize the amount of lichen around the maximum sustainable level.

DOI: 10.1002/sdr

Concrete first advice: Reduce the herd size to zero in the first year, then gradually increase the herd size as the amount of lichen grows towards the level that yields the maximum lichen growth.

Activist (announced as Expert no. 2 when combined with misleading advice)
The previous owner has followed an irresponsible policy of overgrazing. This is yet another example of the overutilization of renewable resources that has been observed so often around the world. The reason for the overgrazing is that the owner has been greedy and that the government lacks a firm policy to regulate the use of natural, renewable resources. The only sensible thing to do now is to reduce the herd size drastically, both to protect the natural environment and to ensure the sustainable operation of the reindeer business.
Concrete first advice: Reduce the herd size to zero in the first year, then gradually increase the herd size as the amount of lichen grows towards the level that yields the maximum lichen growth.

Subjects

Most sessions of the experiment were carried out in the introductory class to one and the same master program in System Dynamics over many years. Hardly any of these students come from Norway or other countries with reindeer herding. Thus, they have no prior knowledge about reindeer management, have not been influenced by public debates, are not aware of existing policy advice, and have no knowledge of anybody's position on the issues. Data for four subjects were removed because, when asked, they reported knowledge of the original reindeer experiment (Moxnes, 2004). To expand the number of subjects with the combined expert and misleading advice, two sessions took place at a military academy. All these subjects were from Norway, and some knowledge about reindeer policies cannot be ruled out.

Three groups were defined related to different backgrounds when it comes to understanding of dynamic systems:

1. *Unskilled*: Master students at the University of Bergen on their first day of the study program.
2. *Trainee*: Master students at the University of Bergen after their first week, in which they learn about stock and flow dynamics.
3. *Military*: Students at the Norwegian Military Academy with varied backgrounds taking a course in logistics at the academy in Bergen (available only in 2004).

A plan was made for the experiment where different treatments were spread over future classes before the exact class sizes were known. Each

Table 1. Overview of experimental treatments and subjects, for 191 subjects total

Advice	Groups	*N*	Years
No Expert	Unskilled	33	2002 & 2003[a]
Analyst	Unskilled	14	2017
Activist	Unskilled	17	2017
Misleading	Unskilled	17	2017
Analyst & Misleading	Unskilled	13	2004 & 2005
Activist & Misleading	Unskilled	12	2004 & 2005
Analyst & Misleading	Trainee	22	2006 & 2007
Activist & Misleading	Trainee	21	2006 & 2007
Analyst & Misleading	Military	21	2004
Activist & Misleading	Military	21	2004

[a] The No Expert data are copied from the original reindeer experiment carried out in 2002 & 2003 using the same subject pool (Moxnes, 2004).

subject participated only once such that the analysis relies on between-subject comparisons. Table 1 gives an overview of treatments, subject groups, number of subjects, and years.

The following protocol for randomization was used. To simplify the teaching and the administration of the experiment, no class received all six of the different treatments shown in Table 1. Analyst and activist advice combined with misleading advice were split randomly between subjects within classes. None of these treatments were carried out in only one student cohort. The class of 2017 was split randomly into three treatments. These last data were added to find the isolated and individual effects of expert and misleading advice, similar to the no expert advice in the original experiment.

Statistical tests

The tests focus is on the lichen only since there is a unique relationship between reindeer (the decision) and lichen (the outcome). Table 2 gives an overview of the tests, with the obtained statistics. The table shows median and average statistics over the last 14 years of the experiment, thus omitting the fixed initial value of lichen. Only in one case do median and average statistics lead to different inferences. Statistical tests are Student's *t*-tests for comparisons to the optimal development and Welch's *t*-tests for comparisons of means of independent samples with equal or unequal variances. Cohen's d is represented by Glass' Δ for comparisons to the optimal development and by Hedges' g when comparing means of independent samples.

DOI: 10.1002/sdr

Table 2. Summary of statistical tests with average and median treatment effects over the years 1–15

Group	Testing (*N*)	Average *p*-value	Median *p*-value	Median Cohen's *d*
Unskilled	No Expert (33) vs. Optimal	≈ 0	≈ 0	2.19
	Experts (31) vs. Optimal	0.62	0.64	0.08
	Experts (31) vs. No Expert (33)	≈ 0	≈ 0	1.50
	Experts & Misleading (25) vs. Experts (31)	0.002	0.001	0.88
	Experts & Misleading (25) vs. Optimal	≈ 0	≈ 0	0.94
	Experts & Misleading (25) vs. No Expert (33)	0.15	0.05	0.57
	Misleading (17) vs. No Expert (33)	0.52	0.46	0.24
	Misleading (17) vs. Experts (31)	≈ 0	≈ 0	1.15
	Analyst (14) vs. Activist (17)	0.68	0.66	0.16
	Analyst & Mis. (13) vs. Activist & Mis. (12)	0.61	0.62	0.20
Trainees	Experts & Misleading (43) vs. Optimal	0.001	0.001	0.54
	Analyst & Mis. (22) vs. Activist & Mis. (21)	0.49	0.50	0.21
Military	Experts & Misleading (42) vs. Optimal	≈ 0	≈ 0	1.02
	Analyst & Mis. (21) vs. Activist & Mis. (21)	0.04	0.04	0.68
Testing	Treatments (*N*)			
Tr. vs. Un.	An. & Mis. (22 Tr.) and An. & Mis. (13 Un.)	0.07	0.07	0.65
Tr. vs. Mi.	An. & Mis. (22 Tr.) and An .& Mis. (21 Mi.)	0.003	0.002	1.04
Un. vs. Mi.	An. & Mis. (13 Un.) and An. & Mis. (21 Mi.)	0.39	0.37	0.35

Note: Sample sizes are shown in parentheses.
Abbreviations: An., Analyst; Mis., Misleading; Mi., Military; Tr., Trainee; Un., Unskilled.

Experimental results with no advice

Figure 4 shows the optimal reindeer and lichen developments for the simulator. Initially, the stock of lichen $L_0 = 24.4$ mm is below $L^* = 30$ mm giving the MSY. The optimal policy is to reduce the number of reindeer from the high historical level to zero, consistent with a target escapement policy (Reed, 1979). As lichen grows towards L^*, reindeer grazing should be increased to the MSY, $R^* = 1250$. Obviously, the optimal policy is much different from a wait and see policy. Not shown in the graph, a good result can also be obtained by reducing grazing to less than optimal, as long as grazing is well below growth in the early years. If so, the MSY will be reached at a later point in time.

Figure 4 also shows the yearly median results for the group of *Unskilled* with no advice. The number of reindeer follows a typical wait and see

DOI: 10.1002/sdr

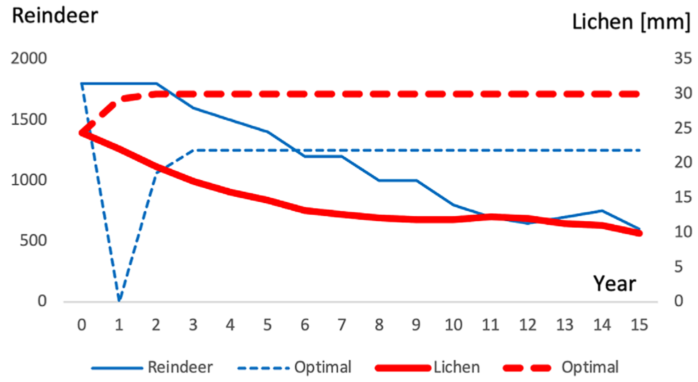

Fig. 4. Optimal (dashed) and yearly median (solid) results for the number of reindeer (thin blue lines) and for lichen (thick red lines). The results are for the group of *Unskilled* with *No Expert* advice, $N = 33$ [Color figure can be viewed at wileyonlinelibrary.com]

development. The number is kept constant for two years. Then, as the amount of lichen is seen to decrease, subjects start to gradually reduce the number of reindeer. Repeated small reductions are consistent with a belief in an inverse and instantaneous cause and effect relationship between reindeer and lichen. However, to the subjects' big surprise, as long as grazing exceeds lichen growth, lichen decreases in pace with the number of reindeer. Only 3 out of 33 subjects manage to increase the lichen to or above the level giving the MSY. For the vast majority, there is hardly any evidence of learning. Allowing for different parameters for aggressiveness, one and the same wait and see policy can explain nearly all decisions over the entire period (Moxnes, 2004).

Why rely on a wait and see policy?

By its design the experiment precludes rational reasons for using a wait and see strategy. Subjects do not benefit from shying away from addressing an unpleasant problem (Walgrave and Dejaeghere, 2017). They do not need repeated observations to figure out the strategies of competitors. Information is exact, readily available, and there is no need to filter randomness. The stated goal precludes a possible preference for discounting the future.

So, what are the less rational reasons for relying on a wait and see policy? According to Kahneman (2011), System 1 decision-making relies on heuristics, is simple and emotional, reflects recent experiences, and produces biases. For example, System 1 thinking leads to biased risk perception. Kunreuther (1996) finds that people often wait to see if a natural disaster occurs before they buy insurance. This modus operandi is not likely to be reserved for uncertain random events. Uncertainty could be just as prevalent if one is not able to anticipate predictable policy consequences.

Kahneman's System 2 decision-making involves cost-benefit tradeoffs, recognizes complex system interactions, and has a focus on the long term. Since

DOI: 10.1002/sdr

natural resource systems are dynamically complex, it seems overly optimistic to assume that majorities of the electorates, politicians, and influencers are fully capable of performing the needed analysis by themselves. Nor are they likely to be capable of assessing the quality of the analyses made by experts.

In a paper on reinforcement learning with references to the Rescorla-Wagner model of animal learning, Niv (2009, p. 141) writes that “learning occurs *only when events violate expectations*.” In a two-stock reindeer experiment by Moxnes (1998b), 92% of the subjects reported that they were surprised by the behavior of lichen. However, in spite of repeated and unexpected outcome feedback, the results did not show evidence of learning in terms of changed policies.

Learning from violations of expectations can be seen in light of two different learning models, “model-free” and “model-based” reinforcement learning. Using magnetic resonance imaging, Gläscher, Daw et al. (2010) found neural signatures of both these types of learning. Thus, their study suggests that lack of suitable neural networks is not an explanation for limited learning. Still, the existence of networks may not say much about the capacity for learning. As a minimum, trained scientists do demonstrate an ability to utilize their neural networks for model-based learning.

Model-free learning does not leave many options for learning other than probing actions and waiting to see whether actions lead to improvement or not. Similar to the heuristics described by Tversky and Kahneman (1974), the wait and see heuristic presents itself automatically. While it requires little deliberation, it could cause the problem one should avoid. So, if the potential costs of mismanagement are large, is it likely that people upgrade to model-based learning?

In model-based “state prediction error learning,” a cognitive map or model of the system at hand is needed. The model is used to make predictions that can be compared to observations, i.e. the scientific method. For this learning talent to be useful, two conditions must be met: people must be able to identify system interactions and they must be able to predict the consequences of actions.

Previous research has found that people tend to simplify reality into isolated, instantaneous cause and effect relationships (Moxnes, 1998b; Sweeney and Sterman, 2000; Cronin *et al.*, 2009). If so, lichen is not represented properly as a stock with an inflow of net growth and an outflow of grazing. Rather lichen is believed to vary inversely and instantaneously with changes in the number of grazing reindeer. When this latter model fails to predict the lichen development, subjects are challenged to formulate a stock and flow representation for lichen. This task is just as complex as it was for Newton to formulate his first law of motion. That was a task that Aristotle and many others failed before Newton (DiSessa, 1982).

DOI: 10.1002/sdr

Making predictions of model behavior is also complicated. Modelers know that predicting the behavior of nonlinear, dynamic models requires simulations. Consistent with simplified instantaneous cause and effect models, people expect to see pattern matching (correlations) between salient causes and effects (Sterman and Sweeney, 2002; Cronin *et al*., 2009).

With an instantaneous model in mind, state-prediction error learning implies that this theory should be rejected when lichen decreases in spite of reductions in reindeer grazing. However, subjects tend to save their theory by auxiliary hypotheses allowing for fictitious information delays or ideas expressed in comments such as “Few reindeer eat much” and “has lichen been permanently injured?” (Moxnes, 1998b). All this speaks of the complexity of the reindeer management task and the need for guidance.

The wait and see heuristic is not only observed among inexperienced students, it is also observed among experienced decision-makers such as stakeholders, regulators, journalists, and researchers (Moxnes, 1998a, 2000). According to a former Vice Chairman of the US Federal Reserve (Blinder, 1997, p. 10): “I cannot tell you how many times, both at the Federal Reserve and at meetings with foreign central bankers, discussions of future policy were cut short with phrases like ‘let’s see what happens’. … Unfortunately, this bit of received central banking wisdom is not at all wise.”

Tversky and Kahneman (1974) found that people make judgements with a representativeness heuristic where they rely on observations and neglect prior information. This heuristic may also bias choices between methods of analysis. For instance, basing recommendations on a method of “skeptical empiricism” may justify a wait and see strategy to gather more data. If more data is prioritized over prior systems information, that may preclude a System 2 comparison of the *expected value* of more data to the *expected cost* of delaying actions. The complexity of such a comparison (Moxnes, 2003) explains why more data is typically prioritized.

Effects of expert and misleading advice

Figure 5 shows average lichen developments for the different information treatments for the group of *Unskilled*. Table 2 shows the related statistical tests comparing treatment effects.

No Expert advice leads to significantly less lichen than *Optimal* with a very large median Cohen’s *d* of 2.19, implying the difference is consequential. Pooled *Experts* advice (analyst and activist) leads to lichen levels that are not significantly different from *Optimal* with a very low Cohen’s *d* of 0.08. Pooled *Experts* advice also leads to significantly larger lichen levels than *No Expert* advice, with a large Cohen’s *d* of 1.50.

By itself, *Misleading* advice, which supports the wait and see heuristic, leads to average lichen levels that are not significantly different from the

DOI: 10.1002/sdr

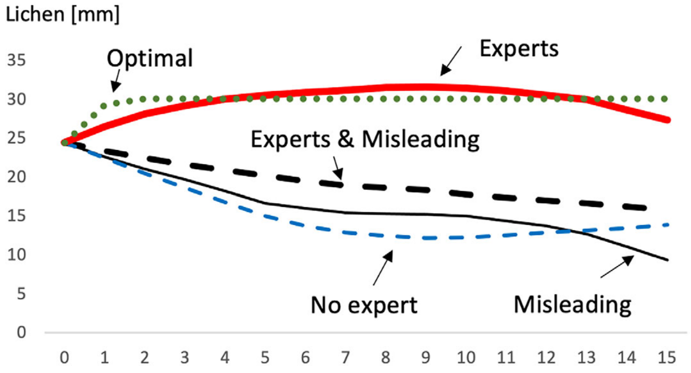

Fig. 5. Optimal lichen development and average lichen developments for *Unskilled* with: No *Expert* advice, pooled *Experts* advice (*Analyst* and *Activist*), *Misleading* advice, and combined *Experts & Misleading* advice [Color figure can be viewed at wileyonlinelibrary.com]

case with *No Expert* advice, with a small median Cohen's *d* of 0.24. This indicates that an explicitly stated wait and see strategy resonates well with the subjects and leads to no significant change of actions. Similar to *No Expert* advice, *Misleading* advice leads to significantly poorer results than the pooled *Experts* advice with a large Cohen's *d* of 1.15.

Combined *Expert & Misleading* advice (*Analyst* combined with *Misleading* advice pooled with *Activist* combined with *Misleading* advice) leads to lichen levels significantly lower than for the case with pooled *Experts* advice, with a large Cohen's *d* of 0.88. Lichen is also lower than optimal with a large Cohen's *d* of 0.94. On average, *Experts & Misleading* advice is not significantly different from *No Expert* advice, with a medium Cohen's *d* of 0.57. By combining *Misleading* advice with *Expert* advice, on average 73 percent of the gain obtained by *Expert* advice is lost. Hence, the experiment suggests that misleading advice is a major problem for democratic decision-making.

Combined *Experts & Misleading* advice leads to lichen levels that are significantly lower than optimal not only for *Unskilled* but for all three subject groups, with Cohen's *d* ranging from 0.54 for *Trainee* to 1.02 for *Military*.

Comparing *Analyst* and *Activist* advice, both combined with *Misleading* advice, there is a significant difference for *Military* only, with a Cohen's *d* of 0.68. For *Military*, the *Activist* wording leads to a better result than the *Analyst* advice. This could be because *Military* got the least help from the *Analyst* advice or because they were more alert to the *Activist* advice.

Figure 6 shows effects of combined *Analyst & Misleading* expert advice for the three subject groups. *Trainee* performs significantly better than *Military* with a large Cohen's *d* of 1.04, and marginally better than *Unskilled* with a medium Cohen's *d* of 0.65. *Unskilled* is not significantly better than *Military*. Hence, one week of general training in understanding stock and flow relationships increases the effect of the *Analyst* advice and helps counteract misleading advice. For the combined *Activist* & *Misleading* advice,

DOI: 10.1002/sdr

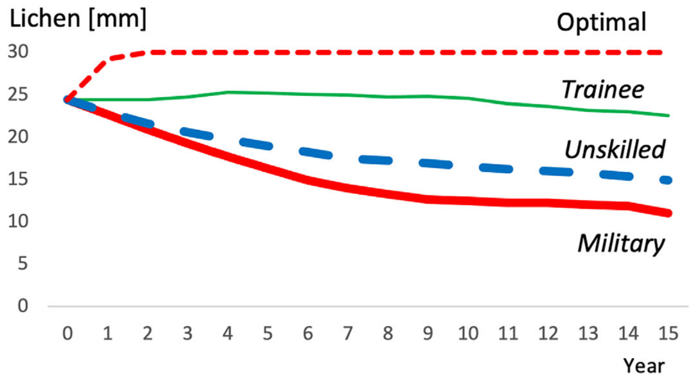

Fig. 6. Lichen development for *Analyst & Misleading* advice for the groups of *Trainee*, *Unskilled*, and *Military* [Color figure can be viewed at wileyonlinelibrary.com]

there are no significant differences between the three subject groups. Hence, *Trainee*'s stock and flow training does not increase the benefits of the *Activist* advice.

The positive effect of training is similar to observed effects of educational background in earlier studies (Moxnes and Saysel, 2009; Guy *et al.*, 2014). These results seem contrary to Kahan *et al.* (2012), who found no effect of general science literacy and numeracy on climate risk perception. However, their findings apply to risk perception and not to policy-making. They also consider general knowledge rather than problem-oriented knowledge. Hence these results of Kahan *et al.*'s study are not likely to apply to policy-making.

Dealing with misperceptions and misleading advice

The experiment reveals disturbing effects of misperceptions and of misleading advice on reindeer management, the management of many other renewable resources, and overall sustainability. This may seem consistent with the statement by Lau and Redlawsk (2001, p. 951) that "The widespread ignorance of the general public … is one of the best documented facts in all of the social sciences." However, the problem of misleading advice is more complex than the problem of ignorance and fake news, which can be counteracted by simple facts. Misperceptions of policies are supported and bolstered by complexity, unconscious heuristics, and misleading advice that in the experiment removed 73 percent of the effect of the *expert* advice. This calls for extraordinary policies to educate change-makers, decision-makers, and electorates and to counter misleading advice. This adds to previous calls for "elites and advocacy groups" to correct misperceptions about climate change (Brulle *et al.*, 2012; Cook *et al.*, 2017).

As a beacon of hope, even in the two treatments with *No Expert* and *Misleading* advice, a vast majority starts to reduce the number of reindeer long

DOI: 10.1002/sdr

before it is too late to reach the MSY. This is consistent with people's talent for pattern recognition. When observing an ongoing decline in lichen, model-free learning produces no alternative to simply assuming that the observed phenomenon will repeat itself. Consequently, dire long-term predictions provoke emotions such as frustration, fear, and risk perception. Recall the demand by Snøhetta locals for lichen assessments already in 1952, about 13 years before lichen reached its low point (see Figure 2). Hence, observed patterns of decline may allow for earlier actions than sudden and unexpected events like flooding (Kunreuther, 1996).

However, frustration is of limited help if decision-makers do not understand what it takes to halt the decline in lichen. In the experiment, hardly any of the early adjustments were sufficient to stop the decline in lichen in time. Hence, the wait and see strategy becomes a recurring trial-and-error strategy. One real world counterpart is the more than fifty years it has taken to improve fishery policies to deal with the commons problem (Moxnes, 2010). Since observed trends motivate policy change, the experiment suggests that information policies are primarily needed to correct policies.

How can people and in particular decision-makers learn to act on system knowledge rather than relying on trial-and-error strategies of the wait and see type? At a detailed level, experiments show that partial system knowledge of the core feedback structure leads to better performance (Gary and Wood, 2016). For instance, information about in- and outflows leads to better policies than observations of stock developments (Moxnes, 1998a, 1998b; Moxnes and Saysel, 2009). The present experiment shows that a minimum of general education in stock and flow behavior (*Trainees*), enhances the effect of *Analyst* expert advice.

Available observations that are contrary to expectations provide learning opportunities (Niv, 2009). For example, in Figures 2 and 4 there are periods with "unexpected" positive correlations between the number of reindeer and the amount of lichen. Another example is the unexpected negative correlation between global CO_2 emissions and the amount of CO_2 in the atmosphere in the early 1980s (Moxnes and Saysel, 2009). However, these learning opportunities require systems knowledge to be utilized to their full potential.

When systems are complex, people seek guidance in experience rather than in theory. Diffusion of technologies is helped by observing positive experiences by others, by carefully trying out new technologies in familiar settings, and by compatibility with current ways of doing things. Diffusion is hindered by complexity (theory) (Rogers, 1995). These findings are also valid for public policy innovations (Mintrom, 1997). The dynamic similarity between reindeer, fishery, and climate management implies that climate policy-making could benefit from reindeer and fishery experiences. This opens up for larger sets of data to learn from in comparative studies. If there

DOI: 10.1002/sdr

are no relevant experiences to learn from, theory and expert advice become particularly important.

Since Phineas Gage's brain damage, it has been known that emotions are important for transforming knowledge into actions (Damasio *et al.*, 1994). How can this insight be used by media to contribute to beneficial policies? Böhm and Pfister (2017) distinguish between consequence-based and morality-based emotions. They find stronger support for public action when problems are perceived to be caused by humans rather than by nature. Thus, media may be correct in identifying and blaming those responsible for problems. However, the experiment shows that there is no significant difference between the effects of the blaming *Activist* and the explaining *Analyst* expert for the groups of *Unskilled* and *Trainees*. Hence, media should not be shy of presenting *Analyst* type of advice. While the experiment does not measure the polarization and societal conflict that follows from different types of advice, it seems reasonable to expect that blaming leads to more polarization than *Analyst* advice.

The *Misleading* advice praises the historical management (no blaming), is optimistic about the future, and warns against drastic reductions in the number of reindeer. According to a summary article by Schneider *et al.* (2021), optimism and hopefulness are positive emotions that motivate actions. This optimism is likely to have strengthened the effect of the *Misleading* advice. If correct, expert advice should also be formulated to convey optimism and hopefulness. After all, there is a very positive message in the *Analyst* advice because it promises a better future than the *Misleading* advice. To illustrate the potential, Figure 8 in Dabla-Norris *et al.* (2023) identifies several side-effects of climate policies that people rate as positive.

What institutions could help counteract misleading advice? By definition, democracies cannot abandon the freedom of speech. However, general legislation punishes those who give misleading advice about Ponzi schemes and tobacco smoking. For sustainability, this is of limited help if one has to wait to gather evidence that misleading advice is indeed misleading. In contrast, mismanagement can be limited by general formulations in international treaties and national constitutions. A recent example is the Urgenda (2019) climate case against the Dutch Government, where the state lost in the supreme court with references to international human rights law.

New institutions may be developed to support deliberative democracy (Dryzek *et al.*, 2019). For captive audiences of limited size, learning can be accelerated by the use of "Group model building" (Vennix, 1996; Andersen *et al.*, 1997) and by the related "Participatory modeling" (Stave, 2010; Etienne *et al.*, 2011; Hovmand, 2014). With these methods, problems are analyzed together by representatives for the electorate, stakeholders, changemakers, politicians, and experts with different backgrounds. All participants get a chance to express and test their ideas about system interactions. According to a survey by Rouwette *et al.* (2002), participants gain a sense of

DOI: 10.1002/sdr

ownership of the analysis and they report positively about the results in terms of insights, shared language, and commitments to revised policies. New institutions for deliberative democracy are needed to ensure that governments involve citizens and listen to their policy recommendations. Dryzek *et al*. (2019) give examples of successful experiences.

A light version of the above is Fishkin's (1991) deliberative polls where results of ordinary polls are compared to polls of citizens that have received information to provide more knowledge-based answers. Significant shifts in opinions have been observed (Luskin *et al*., 2002). Also see Dabla-Norris *et al*. (2023), who find that providing information about the effectiveness of carbon pricing policies increases public support. Fishkin speculates that media will dramatize such shifts in opinions. This seems important in light of Kahan et al.'s (2012, p. 734) conclusion that a person would "be best off if he formed risk perceptions that minimized any danger of estrangement from his community." Hence, by publishing knowledge-based polls that favor proactive policies, deliberative polls may reduce the danger of estrangement. However, if expert advice is blended with misleading advice, Fishkin's deliberative polls may fail their purpose. Therefore, deliberative polls may work better if citizens take part in the above participatory activities or in role playing events (Rooney-Varga *et al*., 2018).

Acknowledgements

I would like to thank the editor and three anonymous reviewers as well as Etiënne Rouwette, and Birgit Kopainsky for most helpful comments.

Biography

Erling Moxnes is Professor Emeritus in System Dynamics at the University of Bergen. His research focuses on sustainability and understanding human decision-making. He has contributed methods to find near-optimal policies in stochastic, non-linear, System Dynamics models with measurement error. He also uses optimization to analyze model behavior and the efficacy of policies. He is currently engaged in building MOOCs to teach resources management and system dynamics.

References

Andersen DF, Richardson GP, Vennix JAM. 1997. Group model building: adding more science to the craft. *System Dynamics Review* **13**(2): 187–201.

DOI: 10.1002/sdr

Ansolabehere S, Jones PE. 2010. Constituents' responses to congressional roll-call voting. *American Journal of Political Science* **54**(3): 583–597.

Blinder AS. 1997. Distinguished lecture on economics in government: what central bankers could learn from academics–and vice versa. *The Journal of Economic Perspectives* **11**(2): 3–19.

Böhm G, Pfister H-R. 2017. The perceiver's social role and a risk's causal structure as determinants of environmental risk evaluation. *Journal of Risk Research* **20**(6): 732–759.

Bolle F. 1990. High reward experiments without high expenditure for the experimenter? *Journal of Economic Psychology* **11**: 157–167.

Brulle RJ, Carmichael J, Jenkins JC. 2012. Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the US, 2002–2010. *Climatic Change* **114**(2): 169–188.

Cook J, Lewandowsky S, Ecker UK. 2017. Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence. *PLoS One* **12**(5): e0175799.

Cronin MA, Gonzalez C, Sterman JD. 2009. Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes* **108**(1): 116–130.

Dabla-Norris E, Helbling T, Khalid S, Khan H, Magistretti G, Sollaci A, Srinivasan K. 2023. Public perceptions of climate mitigation policies: evidence from cross-country surveys. *Staff Discussion Notes* **2023**(2): 33.

Damasio H, Grabowski T, Frank R, Galaburda AM, Damasio AR. 1994. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* **264**(5162): 1102–1105.

DiSessa AA. 1982. Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science* **6**(1): 37–75.

Dryzek JS, Bächtiger A, Chambers S, Cohen J, Druckman JN, Felicetti A, Fishkin JS, Farrell DM, Fung A, Gutmann A. 2019. The crisis of democracy and the science of deliberation. *Science* **363**(6432): 1144–1146.

Etienne M, Du Toit DR, Pollard S. 2011. ARDI: a co-construction method for participatory modeling in natural resources management. *Ecology and Society* **16**(1). Available online: http://www.ecologyandsociety.org/vol16/iss1/art44/

Fishkin JS. 1991. *Democracy and deliberation: new directions for democratic reform, 217*. Yale University Press: New Haven, CT.

Gary MS, Wood RE. 2016. Unpacking mental models through laboratory experiment. *System Dynamics Review* **32**(2): 101–129.

Gläscher J, Daw N, Dayan P, O'Doherty JP. 2010. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**(4): 585–595.

Guy S, Kashima Y, Walker I, O'Neill S. 2013. Comparing the atmosphere to a bathtub: effectiveness of analogy for reasoning about accumulation. *Climatic Change* **121**(4): 579–594.

Guy S, Kashima Y, Walker I, O'Neill S. 2014. Investigating the effects of knowledge and ideology on climate change beliefs. *European Journal of Social Psychology* **44**(5): 421–429.

Hardin G. 1968. The tragedy of the commons. *Science* **162**(13 December): 1243–1248.

Hovmand PS. 2014. Group model building and community-based system dynamics process. In *Community Based System Dynamics*. Springer: New York, NY; 17–30.

Jackson JB, Kirby MX, Berger WH, Bjorndal KA, Botsford LW, Bourque BJ, Bradbury RH, Cooke R, Erlandson J, Estes JA. 2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science* **293**(5530): 629–637.

Jones MD, McBeth MK. 2010. A narrative policy framework: clear enough to be wrong? *Policy Studies Journal* **38**(2): 329–353.

Jordhøy P 2001. Snøhettareinen. Snøhetta forlag a.s.

Kahan DM, Peters E, Wittlin M, Slovic P, Ouellette LL, Braman D, Mandel G. 2012. The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change* **2**(10): 732–735.

Kahneman D. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux: New York.

Konikow LF, Kendy E. 2005. Groundwater depletion: a global problem. *Hydrogeology Journal* **13**(1): 317–320.

Kunreuther H. 1996. Mitigating disaster losses through insurance. *Journal of Risk and Uncertainty* **12**(2-3): 171–187.

Lau RR, Redlawsk DP. 2001. Advantages and disadvantages of cognitive heuristics in political decision making. *American Journal of Political Science* **45**((4)): 951–971.

Leiserowitz A. 2007. International public opinion, perception, and understanding of global climate change. *Human Development Report* **2008**: 1–40.

Luskin RC, Fishkin JS, Jowell R. 2002. Considered opinions: deliberative polling in Britain. *British Journal of Political Science* **32**(3): 455–487.

Mintrom M. 1997. Policy entrepreneurs and the diffusion of innovation. *American Journal of Political Science* **41**(3): 738–770.

Moxnes E. 1998a. Not only the tragedy of the commons, misperceptions of bioeconomics. *Management Science* **44**(9): 1234–1248.

Moxnes E. 1998b. Overexploitation of renewable resources: the role of misperceptions. *Journal of Economic Behavior & Organization* **37**(1): 107–127.

Moxnes E. 2000. Not only the tragedy of the commons: misperceptions of feedback and policies for sustainable development. *System Dynamics Review* **16**(4): 325–348.

Moxnes E. 2003. Uncertain measurements of renewable resources: approximations, harvest policies, and value of accuracy. *Journal of Environmental Economics and Management* **45**(1): 85–108.

Moxnes E. 2004. Misperceptions of basic dynamics, the case of renewable resource management. *System Dynamics Review* **20**(2): 139–162.

Moxnes E. 2010. Complexities in fisheries management: misperceptions and communication. In *Handbook of Marine Fisheries Conservation and Management*, Hilborn R, Squires D, Williams M, Tait M, Grafton Q (eds). Oxford University Press: Oxford.

Moxnes E. 2014. Discounting, climate and sustainability. *Ecological Economics* **102**: 158–166.

Moxnes E, Danell Ö, Gaare E, Kumpula J. 2003. *A Decision Tool for Adaptation of Reindeer Herds to Rangelands: The User's Manual*. Tromsø: Norway.

Moxnes E, Jensen LC. 2009. Drunker than intended; misperceptions and information treatments. *Drug and Alcohol Dependence* **105**: 63–70.

Moxnes E, Saysel AK. 2009. Misperceptions of global climate change: information policies. *Climatic Change* **93**(1-2): 15–37.

Niv Y. 2009. Reinforcement learning in the brain. *Journal of Mathematical Psychology* **53**(3): 139–154.

Oreskes N, Conway EM. 2010. *Merchants of Doubt*. Bloomsbury Press: New York.

DOI: 10.1002/sdr

Reed WJ. 1979. Optimal escapement levels in stochastic and deterministic models. *Journal of Environmental Economics and Management* **6**: 350–363.

Rogers EM. 1995. *Diffusion of Innovations*, Fourth ed. The Free Press: New York.

Rooney-Varga JN, Sterman JD, Fracassi E, Franck T, Kapmeier F, Kurker V, Johnston E, Jones AP, Rath K. 2018. Combining role-play with interactive simulation to motivate informed climate action: evidence from the world climate simulation. *PLoS One* **13**(8): e0202877.

Rouwette EAJA, Größler A, Vennix JAM. 2004. Exploring influencing factors on rationality: a literature review of dynamic decision-making studies in system dynamics. *Systems Research and Behavioral Science* **21**(4): 351–370.

Rouwette EAJA, Vennix JAM, Mullekom TV. 2002. Group model building effectiveness: a review of assessment studies. *System Dynamics Review* **18**(1): 5–45.

Scheffer VB. 1951. The rise and fall of a Reindeer Herd. *The Scientific Monthly* **75**: 356–362.

Schneider CR, Zaval L, Markowitz EM. 2021. Positive emotions and climate change. *Current Opinion in Behavioral Sciences* **42**: 114–120.

Schrank WE. 2003. *Introducing Fisheries Subsidies*. FAO: Rome.

Stave K. 2010. Participatory system dynamics modeling for sustainable environmental management: observations from four cases. *Sustainability* **2**(9): 2762–2784.

Sterman JD. 1989. Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science* **35**(3): 321–339.

Sterman JD. 2008. Risk communication on climate: mental models and mass balance. *Science* **322**(5901): 532–533.

Sterman JD. 2011. Communicating climate change risks in a skeptical world. *Climatic Change* **108**: 811–826.

Sterman JD, Sweeney LB. 2002. Cloudy skies: assessing public understanding of global warming. *System Dynamics Review* **18**(2): 207–240.

Sterman JD, Sweeney LB. 2007. Understanding public complacency about climate change: adults' mental models of climate change violate conservation of matter. *Climatic Change* **80**(3-4): 213–238.

Supran G, Rahmstorf S, Oreskes N. 2023. Assessing ExxonMobil's global warming projections. *Science* **379**(6628): eabk0063.

Sweeney LB, Sterman JD. 2000. Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review* **16**(4): 249–286.

Toresen R, Østvedt OJ. 2000. Variation in abundance of Norwegian spring-spawning herring (Clupea harengus, Clupeidae) throughout the 20th century and the influence of climatic fluctuations. *Fish and Fisheries* **1**(3): 231–256.

Tversky A, Kahneman D. 1974. Judgment under uncertainty: heuristics and biases. *Science* **185**: 1124–1131.

Urgenda 2019. *ECLI:NL:HR:2019:2007*, 2019. Available online: https://www.urgenda.nl/wp-content/uploads/ENG-Dutch-Supreme-Court-Urgenda-v-Netherlands-20-12-2019.pdf.

Vennix JAM. 1996. *Group Model-Building: Facilitating Team Learning using System Dynamics*. Wiley: Chichester.

Walgrave S, Dejaeghere Y. 2017. Surviving information overload: how elite politicians select information. *Governance* **30**(2): 229–244.