European Journal for Philosophy of Science (2022) 12:68
https://doi.org/10.1007/s13194-022-00488-4

PAPER IN PHILOSOPHY OF SCIENCE IN PRACTICE

Values in climate modelling: testing the practical applicability of the Moral Imagination ideal

Karoliina Pulkkinen[1,2] · Sabine Undorf[3,4] · Frida A.-M. Bender[4]

Received: 16 March 2022 / Accepted: 12 September 2022 / Published online: 18 November 2022

Abstract

There is much debate on how social values should influence scientific research. However, the question of practical applicability of philosophers' normative proposals has received less attention. Here, we test the attainability of Matthew J. Brown's (2020) Moral Imagination ideal (MI ideal), which aims to help scientists to make warranted value-judgements through reflecting on goals, options, values, and stakeholders of research. Here, the tools of the MI ideal are applied to a climate modelling setting, where researchers are developing aerosol-cloud interaction (ACI) parametrizations in an Earth System Model with the broader goal of improving climate sensitivity estimation. After the identification of minor obstacles to applying the MI ideal, we propose two ways to increase its applicability. First, its tools should be accompanied with more concrete guidance for identifying how social values enter more technical decisions in scientific research. Second, since research projects can have multiple goals, examining the alignment between broader societal aims of research and more technical goals should be part of the tools of the MI ideal.

Keywords Values in science · Moral Imagination ideal · Climate modelling · Climate sensitivity

1 Introduction

The centre of the debate on values in science has shifted from discussing the viability of the ideal that science should be free from social, political, and ethical values towards what Holman and Wilholt (2022) title the new "demarcation problem" – the task of distinguishing between legitimate and illegitimate influence of values on

Extended author information available on the last page of the article research.[1] Dovetailing this development, authors have issued numerous recommendations on how values should influence scientific research (Anderson, 2004; Douglas, 2009; Intemann, 2015; Elliott, 2017; Steel, 2017; Brown, 2020; Lusk, 2020; Koskinen & Rolin, 2022). The large number of different normative proposals gives rise to the question of how to evaluate them. Albeit assessing the desirability of the ideals *qua* ideals is an important task of philosophy of science, (e.g., Douglas, 2009; Steel, 2017), here, the focus is on another mode of evaluation: assessing a proposal based on its practical attainability.

Evaluating the practical attainability of proposals to specific fields and tasks in science is especially needed if one rejects the idea that there will be one proposal that satisfies both the necessary and jointly sufficient criteria for legitimate influence of values in science in all cases. For example, Koskinen and Rolin (2022) argue that the problem of distinguishing between legitimate and illegitimate influence of values should not be framed as a quest for both necessary and jointly sufficient criteria: instead, having an "open-ended list of criteria" is useful even if "none of the criteria is a necessary condition of legitimate non-epistemic value influence in science" (2022, 191).

Taking a step towards a more pluralistic direction signalled by Koskinen and Rolin encourages centralising questions that have not previously been given as much attention in the debate. Such questions include: which disciplines, research projects, and phases of research are philosophers' recommendations most applicable to? How general or context-specific are the proposals? What are the main obstacles for putting a proposal in practice? Although such questions are more practice-oriented, bringing them more firmly into the discussion would aid the theoretical development of normative proposals, too. After all, exploring the applicability of proposals can reveal instances of "bad fit" between proposal and practice, which paves the way for further theoretical development.

The normative proposal that will be the subject of such an in-depth evaluation here is the Moral Imagination ideal introduced by Matthew J. Brown in his book *Science and Moral Imagination: A New Ideal for Values in Science* (2020). The Moral Imagination ideal (henceforth the MI ideal) encourages scientists to consider the legitimate stakeholders of the inquiry and imaginatively explore options to contingent, unforced choices that emerge during research. The practical attainability of the MI ideal is worth our consideration at least for two reasons. First, it contributes to the task of drawing a line between legitimate and illegitimate influence of values by introducing a substantial theory of values that "can help us distinguish the legitimate roles for values in science from those that lead to rigid and wishful thinking" (Brown, 2020, 20). Furthermore, what makes the MI ideal an especially suitable proposal to assess in terms of practical attainability is that Brown endorses and defends the idea that such attainability is a key measure of the success of normative proposals, including the MI ideal. In Brown's words, "[i]deas, even philosophical ones, prove their merit in their use; the proof of the pudding is in the eating, not in the formal correctness of the bak-

[1] Although a rigid distinction between epistemic and nonepistemic values (e.g., political, ethical, and other social values) are notoriously difficult to maintain (Rooney, 1992, 2017), we follow Ward (2021) and Rooney (2017) in holding that *some* distinction can be upheld.

ing. Likewise, the proof of an idea is in its success in guiding practice" (Brown, 2020, 215-216). To achieve this end, the MI ideal includes a tool of four guiding activities to help scientists to put the ideal in practice.

We explore the MI ideal's attainability by applying it to a climate modelling setting. After introducing the MI ideal in greater detail in Sect. 2, its central instrument for practical application – the four guiding activities – will be applied to a climate modelling task (Sect. 3.) The climate scientists of our team, experts on Aerosol-Cloud Interactions (ACI), have used the tools of the MI ideal and related them to their project which involves modifying a parametrization of ACI in an Earth System Model. After introducing the key features of their discussion, the applicability of the MI ideal and its tools is evaluated with two tests. The first test involves identifying more minor obstacles that point towards small amendments to the existing tools of the ideal (Sect. 4). The second test involves evaluating whether the tools aid in achieving the broader aims of the ideal – that is, whether they do guide scientists in making warranted value-judgments in research (Sect. 5). As a result, we identify two areas of further development of the MI ideal and its tools. First, there is a need for more substantial guidance for recognising how social values influence research. Second, since the application to the MI ideal to climate modelling reveals that often the goals of research form a network of broader societal aims and more specific technical ones, assessing the alignment of goals should be part and parcel of the MI ideal and its toolkit. Section 6 concludes.

2 The Moral Imagination ideal

Brown's *Science and Moral Imagination* introduces a new normative ideal for conducting warranted value-judgement in scientific research. The aim of the book is twofold: to help to evaluate decisions in scientific practice and to provide guidance on how value-judgments should be incorporated in research (Brown, 2020, 4).

Underlying the proposed ideal is a substantial theory of values. Building from the recognition that science and values influence each other, Brown's account of values views them as intimately connected with action. Values are derived "from many sources in human life, practice, and experience." Since they play many roles in our activities, they also come in many different types. A central distinction is that between unreflective or habitual value-judgement and reflective one, where reflective value-judgments are "a type of empirical inquiry into questions of what to do" (Brown, 2020, 20). An important step in cultivating such a reflective value-judgement is to recognize that research is full of contingent choices, where each contingency could "become an explicit choice" (Brown, 2020, 19). Resolving such choices cannot always be done on the basis of evidence, epistemic standards, and logic alone. Instead, many choices require reflecting on their foreseeable consequences, as "[a]ny such choice could have foreseeable consequences for what we value" (Brown, 2020, 19). For this reason,

> [T]o find these [foreseeable consequences] out for any particular case, we have to think about values, exercise moral imagination to determine the con-

> sequences of each option, and exercise value judgment as part of the choice. We cannot always foresee the consequences; the choices may sometimes be irrelevant to any values, but we cannot determine that ahead of time without looking at the details of the case. Thus scientists have a responsibility to make value judgments about scientific contingencies (Brown, 2020, 19).

Central in this process of making each contingency an explicit choice is moral imagination, which is needed for evaluating how different choices and actions might have implications for values and to stakeholders. Considering the potential consequences of research in this way calls for "exercising imagination via empathy, dramatic rehearsal, and creative problem solving," which are all part of "all evidence-based inquiry" (Brown, 2020, 21).

From the recognition that research involves many contingent choices that may be rendered explicit with the exercise of moral imagination, Brown summarises the MI ideal as follows:

> Scientists should recognize contingencies in their work as unforced choices, discover morally and epistemically salient aspects of the situation they are deciding, empathetically recognize and understand the legitimate stakeholders and their interests, imaginatively construct and explore possible options, and exercise fair and warranted value judgment in order to guide those decisions. (Brown, 2020, 186)

As noted in the introduction, for Brown, an important avenue for evaluating the strength of the proposal is to test its practical attainability. For this reason, the MI ideal comes with a tool of four guiding activities which are geared to help scientists in deliberating on contingencies that arise during research.[2] The four activities are the following:

1. Identify the goal or task at hand.
2. Identify and imaginatively multiply options for how to carry out the task.
3. Determine the standards and values that are relevant to the situation.[3]
4. Identify the legitimate stakeholders to consider and identify their interests.

The four activities are connected to the MI ideal in virtue providing the concrete actions that help to realise the recommendations spelled out by the ideal. In Brown's words, "[m]utual refinement and development of these four areas provide the materials necessary for acting on the ideal of moral imagination," where the four activi-

[2] It should be noted that Brown introduces other practical tools e.g., for helping to settle which values ought to be prioritized in cases of conflict (Brown, 2020, 153–154) and to "test" value judgments in research through "dramatic rehearsal" and tentative application (Brown, 2020, 161). However, the worksheet with the four guiding activities is a central strategy for putting the MI ideal to action. For that reason, the discussion in this article is going to target it in particular.

[3] The first activity is not represented explicitly in Brown's own summary of the MI ideal above, but is best seen as setting the scope for inquiry and thus helping to determine which options, stakeholders, and values are relevant for the activities 2–4.

ties need not proceed in a linear order (Brown, 2020, 186). These four activities are accompanied by a worksheet to help to put it to use (Brown, 2020, 270).

A specific task in climate modelling would provide a good test-case for the MI ideal for at least two reasons. First, it could benefit the climate modelling community, as there appears to be a need for more applicable normative proposals regarding values in climate modelling (Winsberg, 2012). One indicator of such a need can be found in the IPCC's recent Working Group I contribution to the Sixth Assessment Report (AR6). Although Chap. 1 of AR6 WGI explicitly notes that social values guide certain decisions in construction, assessment, and communication of climate change information, values are not brought up in the subsequent chapters of the report, with the exception of Chap. 10, Linking global to regional climate change, (IPCC, 2021, Ch.1, Executive Summary). This suggests that there is more work to be done in tightening the links between normative proposals on values and climate science, although some progress in this direction has already been achieved (Winsberg, 2012; Intemann, 2015; Parker & Risbey, 2015; Parker & Winsberg, 2018; Parker & Lusk, 2019). The second reason why climate modelling provides a good test-case for the MI ideal is that it was not one of the in-depth examples that Brown uses to illustrate the applicability of the MI ideal to practice. This suggests climate modelling would provide good testing grounds for the ideal, as it has not been designed with that in mind.[4] Thus, considering the practical applicability of the MI ideal to climate science would be a welcomed development.

3 Developing model-representation of aerosol-cloud interactions

In the following sections, we turn our attention to the question of whether the MI ideal can guide scientists in conducting more warranted value-judgements in climate modelling. To test the applicability of the MI approach, the two climate scientists of our team reflected on the four activities that help to put the ideal to practice. As their project concerns modifying the representation of ACI in an Earth System Model for the broader goal of characterising and reducing climate sensitivity estimates, some background information regarding ACI and climate sensitivity should be introduced first.

ACI remain one of the persistent uncertainties in climate science. Atmospheric aerosols are tiny solid and liquid particles mixed in the atmosphere and they can be of both natural and human origin (Fu, 2015; Lohmann, 2015). Apart from having direct

[4] However, Brown brings up climate modelling in relation to the "democratic objection" to values in science. According to this objection, value-laden decisions should be deferred from scientists to policymakers especially in the case of policy-relevant science; otherwise too much authority would be given to scientists (who are not democratically elected) in concerns related to policy making. In response, Brown argues that the technical nature of many decisions renders it both very difficult and impractical to defer value-laden decisions to policymakers. This is illustrated with an example from climate modelling, where the technicality of decisions prevents eliminating scientists' value-judgments and replacing them with the input of policymakers. Such a replacement of scientists' value-judgments would require "not only that they fill in their own values, but that they understand what is at stake well enough in order to do so" (Brown, 2020, 72). Achieving either would require policymakers to "become technically sophisticated near the level of the scientist" (Brown, 2020, 73), which would be impractical.

radiative effects (IPCC, 2021, Ch.6), aerosols also serve as cloud nuclei, and thus influence cloud coverage and thickness, which in turn influences the Earth's radiative balance (Twomey, 1974; Albrecht, 1989). These and other effects depend on aerosol species, cloud regime, and many other circumstances and characteristics, which speaks of the complexity of climate's response to anthropogenic aerosols. Overall, ACI appear to have had a cooling effect on climate (IPCC, 2021; cf. Glassmeier et al., 2021). Since Earth System Models divide the Earth's atmosphere, land, and the upper layer of ocean into large three-dimensional grid boxes, and ACI give an example of a smaller-scale sub-grid process, the usual way to represent them is to parametrize them (Gettelman & Sherwood, 2016; Zhao et al., 2018; Gettelman et al., 2019; for a philosophical treatment, see Baldissera Pacchetti (2021) and Kawamleh (2021)).

As ACI impact Earth's radiative balance, they also play a role in estimating the climate's response to increased atmospheric carbon dioxide and other greenhouse gases. The long-term effects of atmospheric carbon dioxide can be approximately quantified with a single number, Equilibrium Climate Sensitivity (ECS), which is defined as the global-mean temperature change expected in response to a doubling of atmospheric carbon dioxide (Knutti et al., 2017). In issuing an improved sensitivity estimate, one of the major existing uncertainties concerns the effect of ACI on climate, and their representation in Earth System Models (IPCC, 2021, 1.5.3.1.; Arias et al., 2021, 1.2.2.; Meehl et al., 2020). Although the models participating in the most recent Phase of Coupled Model Intercomparison Project (CMIP6) incorporated more ACI processes than those from earlier phases, there remains more work to be done in establishing whether such amendments do in fact improve model performance in terms of the ACI's radiative forcing (Arias et al., 2021, 1.2.1). More specifically for our purposes here, in-situ observations, satellite retrievals, and large eddy simulations suggest that there are discrepancies between the effect of ACI in Earth System Models and their effect in the real world (Bender et al., 2016, 2019; Sotiropoulou et al., 2021; Tan & Storelvmo, 2019; Malavelle et al., 2017; Gryspeerdt et al., 2020.)

Focussing especially on model-based assessments of ECS, we have recently argued that value-judgments enter the estimation of ECS in multiple ways (Undorf et al., 2022). We have also noted the emerging awareness regarding the influence of values in the climate science community and encouraged its further development (Pulkkinen et al., 2022). For these reasons, it is timely to turn attention to the question of whether normative proposals regarding values could help at conducting more warranted value-judgments in research, where our focus is specifically on applying the MI ideal to the representation of ACI in an Earth System Model for the broader goal of improving climate sensitivity estimation. To explore the applicability of the MI ideal to this specific task, we can proceed by executing the four guiding activities of the MI ideal, introduced in Sect. 2. In the following section, we reconstruct the four-step deliberation of our small research group of two climate scientists developing the representation of ACI in an Earth System Model. We proceeded as follows. The MI worksheet with the four guiding activities was distributed by the philosopher overseeing this project to the two climate scientist co-authors, who individually contemplated the four questions. After this, each climate scientist returned the filled worksheet to the philosopher, who after consultation with the scientists, synthesised the answers on a document that detailed the application of the tools of the MI ideal.

The document was then assessed by the climate scientists who made some additional suggestions and improvements for clarity, where the result is presented in the next section. The climate scientists' answers to the four questions then informed the philosopher's analysis (Sects. 4 and 5) albeit the climate scientists of our team made some comments.[5]

Identify the goal or task at hand Several mutually compatible goals can be adopted in a single investigation. The following goals were identified:

(a) To improve understanding of climate sensitivity.
(b) To test the robustness of sensitivity estimates in models.
(c) To reduce discrepancy between observationally-derived data and model-derived estimates of ACI effects.

Upon closer inspection, it becomes clear that the goals range in specificity: goal (a) is the broadest whereas (c) the most specific. It should also be noted that the goals (b) and (c) are means to the end expressed in (a)—that is, (b) and (c) can be characterised as the subgoals of (a). With goal (b), robustness refers to the estimates holding up to the test of altered assumptions in models. This is in line with how robustness is understood in philosophical literature.[6]

Multiply options to carry out the task Keeping in mind the goals stated above, the next stage involves considering how to reach them. With broader goals such as (a), there are numerous avenues that could be pursued. If relying on models, one could distinguish approaches that focus on just one model (e.g., perturbed-physics-ensemble) or many (e.g., model intercomparison). Another option would not be quite so reliant on models explicitly but use them as sources of process understanding. Finally, one could use models to provide "emergent constraints," which refer to the physically explainable empirical relationships between long-term climate projections and aspects of current climate (Hall & Qu, 2006; Klein & Hall, 2015).

With goal (b), one might test robustness by examining the assumptions involved in the representation of processes that are poorly constrained by observations. In doing so, one has the option of focusing on different processes (e.g., cloud feedbacks, ocean circulation), but in this case, the focus is on ACI such as cloud lifetime and cloud brightening.

Goal (c) has already narrowed some options in virtue of focusing on ACI and by specifying the metric of improvement (reducing discrepancy between model output

[5] It should be emphasised that the reconstructed answers are not supposed to be representative of climate modellers' views more broadly, but solely provide material to inform philosophical reflection regarding the applicability of the MI ideal and its tools. Involving a small group of scientists is beneficial for our purposes here, as for Brown, the MI ideal is intended primarily to guide individual scientists and small research groups (Brown, 2020, 17–18; 229).

[6] For example, Weisberg (2006) defines robustness analysis as involving the examination of multiple models that make different assumptions about the phenomenon in question while looking for "a common prediction among these models." This view echoes the classic work by Richard Levins (Weisberg, 2006, 732; Levins, 1966, 20; Winsberg, 2018).

and observations). However, there are still some choices to be made. For example, researchers would have to specify which discrepancies between observationally-derived data and model-derived estimates to focus on.

Standards and values relevant to the situation Both different aims and the options for realising them can carry different values and standards. What values and standards are relevant to the goals and options identified above? When considering goal (a), there are important risks of error to consider. Erroneously estimating the range of climate sensitivity too low would underestimate climate change. This, in turn, could lead societies to make less mitigation and adaptation efforts than acceptable, given the anticipated impacts of a certain level of climate change. In contrast, erroneously estimating ECS too high would overestimate climate change, perhaps leading societies to spend resources in mitigation and adaptation that could have been better spent otherwise.

Aside from evaluating risks of error and their societal consequences, there are also other value-judgments relevant for goal (a). One option for reaching it is to execute a perturbed-physics-ensemble, which might be done on the grounds that they are *complete.* Another option is to combine multiple lines of evidence – as done by Sherwood and co-authors (2020) – which bears the promise of strong *evidential basis* and *robustness*.[7] If opting for a multi-model approach, there also appears to be a trade-off in values when comparing with the previous method of "model democracy," where one model would have one vote in issuing a sensitivity estimate in an ensemble of models (Knutti, 2010). Although the approach of Sherwood and co-authors (2020) constrains the range for climate sensitivity estimates significantly, it is not as *simple* and *transparent* as the previous model democratic approach, which made the weighing of evidence provided through models clearer.

For goal (b), *robustness* appears a central standard, and can be reached by examining the assumptions involved in representing processes that are poorly constrained by observations. Furthermore, when considering those parts of the models that do not have a strong physical foundation or which are not very well constrained by observations, *transparency* regarding assumptions and the relative weakness of physical foundation is important.

When considering options to carry the more specific goal (c), there are a set of values or standards relevant for both modelling and data-processing. One example is to ensure *comparability* which can be reached for example by employing a satellite simulator. Using a simulator helps to get an "apples-to-apples" comparison by rendering the results produced by a model comparable to what a satellite sees (Bodas-Salcedo et al., 2011).

Identify the legitimate stakeholders and their interests The next step of the MI framework requires considering the legitimate stakeholders of research. Note that for Brown *legitimate* stakeholders refers to those stakeholders who "either rightfully

[7] Furthermore, combining multiple lines of evidence emphasises the role of experts' subjective judgement in weighing different lines of evidence when issuing a sensitivity estimate. This suggests a more significant role for experts' own values, albeit not all subjective factors are value-judgments (Morrison, 2014).

participate in or affect the decisions in question, or who will be affected by the decision" (Brown, 2020, 21).

From the three goals, goal (a) has the broadest set of stakeholders. Since uncertainty in climate sensitivity translates into uncertainty in how large emission reductions we need for reaching a specific temperature goal, the high societal relevance of goal (a) is evident. This implies that the legitimate stakeholders of research include at least the general public, policymakers, other scientists, and future generations, albeit there have been some concerns over the "usability" (defined approximately as the actual relevance for policy-making) of ECS (Sobel, 2021).

Although both goals (b) and (c) are means to the broader goal (a), it doesn't seem plausible to state that either of the more specific goals has an equally large set of immediate stakeholders. For example, the decision to use a satellite simulator to improve the comparability of the model output and satellite data are decisions where future generations or policymakers are a more distant concern. Instead, the more immediate stakeholders are other scientists and groups that developed the instruments and tools in use since they are best positioned to evaluate the decisions made in the course of research towards goals (b) and (c).

4 Evaluating the practical applicability of the MI ideal: four obstacles

The MI ideal requires scientists to take note of the contingencies of their work; discover epistemically and morally relevant aspects of their choices; note the interests of legitimate stakeholders; imaginatively construct options and alternatives to choices; and exercise fair value-judgment in guiding decision-making (Brown, 2020, 21; 186). Now that we have followed the activities specified in the previous section, we are better positioned to assess the ideal's applicability to climate modelling. We propose that such an evaluation regarding practical applicability can be conducted by executing the following two tests:

(1) What are the obstacles for applying the tools and strategies of the ideal?
(2) Do the tools and strategies help to reach the broader aims of the ideal?

To illustrate the contrast between the two, consider that we might be testing a pair of gloves. The first test would concern whether the pair of gloves fit our hands, whereas the second test involves seeing whether the gloves can keep our fingers warm in the cold weather. In short, each test tackles a different aspect of the ideal. The first test involves evaluating whether the tools and strategies are applicable to the setting in question, where this assessment can be done by identifying obstacles for applicability. Of course, an ideal might face obstacles other than those relating to its applicability (e.g., issues with its desirability or internal coherence), but here, we predominantly discuss obstacles that make it difficult to apply an ideal to specific contexts of research. For example, the first test could reveal the irrelevance of the key concepts of the MI ideal's tools. This would be the case if there were no options to choose from, no aims, no stakeholders, and no ethical dimensions in the relevant setting. Aside from key concepts, the first test can also reveal the failure of the applicability of the assumptions behind the tools. With the presupposition of scientists' agency (understood in the sense that scientists can make choices regarding their research), we can imagine some contexts where agency is considerably restricted. Also activities recommended by an ideal could turn out to be inapplicable. If there was a requirement to engage with policy-maker stakeholders, but the research in question would not be policy-relevant, it is difficult to see how scientists could meaningfully act in the recommended way.

Aside from the above obstacles regarding key concepts, activities, assumptions, we can imagine also more concrete obstacles, such as communication of the ideal to scientists (e.g., difficulties in making it intelligible) or the ideal going under the radar for those who could potentially benefit from its guidance. It should be noted that since there was a philosopher-mediator involved in this project who introduced the ideal and its tools to scientists, we did not face such challenges. However, the fact that such issues did not arise here does not mean that they could not arise in some other investigations regarding applicability of ideal to specific research context.

In contrast to the first test, the second one concerns the relationship between the tools and the ideal. It asks whether the tools help at reaching the broader goals spelled out by the ideal. For example, if the four guiding activities would systematically fail to lead to "fair and warranted value judgment" that the MI ideal is geared towards producing, then this would call for revising the tools. Although such an assessment could be done on more theoretical grounds, here, we provide our verdict after considering the ACI case introduced above (see Sect. 5.)

In what follows, we will execute above two tests. To advertise our conclusions regarding the first test, four obstacles to applying the tools of the MI ideal are identified. The obstacles include (i) limitations regarding who gets to decide the core research priorities, (ii) limited resources, (iii) having very many options, and (iv) an extremely broad set of stakeholders to consider. (The first two obstacles concern limited opportunities to exercise one's agency. In contrast, the latter two challenges are problems of abundance; they arise when there are very many contingencies and stakeholders to consider.) Examining the former two obstacles helps to specify when the MI ideal can be applied most effectively, whereas the latter two imply that minor adjustments to the tools of the ideal are needed for determining which contingencies and stakeholders to primarily focus on.

Obstacles i & ii: limits of agency Limited agency of modellers has already been discussed in relation to the "path-dependency" of models, where models' current developers are restricted by the choices that previous modellers have made and thus "inherit" their value-judgments (Winsberg, 2012). Here, the issues concerning limited agency are somewhat different, namely that not all researchers enjoy the same opportunities to influence the core research-agenda of the project. With broader goals such as (a), the task of setting aims falls typically on the shoulders of the principal investigators who specify the goals of the investigation in funding applications. The more junior researchers who often are subsequently hired to the project (e.g., postdocs) may exercise significant agency in choosing the more specific secondary and emergent goals, but they may not be in the same position to make substantial changes to the core of the research project since they were not present when the applications were formulated.

The second factor limiting agency becomes clear when considering the second step of the MI ideal. Recall that the second step called for "identifying and imaginatively multiplying options on how to carry out the task." Executing options for carrying out tasks, however, requires resources. Options deemed most promising might not be pursued due to limited resources and computing time. For example, a perturbed-physics-ensemble might be deemed as the best option on the grounds of greater completeness, but it requires substantial project funding and computer resources. So at times, the most promising options for methodologies or tools are not pursued, and researchers might have to settle with a goal that is realistic under the circumstances.

Obstacles iii & iv: problems of abundance Aside from the factors that limit options for researchers, there is also the converse issue of having too many options to consider. With modifying the representation of ACI, scientists may be presented with multiple options on where and how to iterate the existing model. As there are many different processes to focus on, one might target the code that deals with the model's representation of macrophysics. Or, if opting to alter its representation of microphysics, one faces the choice of focussing on aerosol physics or cloud microphysics, alongside other subsequent choices involved in modifying the model's representation of ACI. The multitude of choices risks making the requirement of rendering all contingent choices into explicit ones into a very demanding request.

To answer this concern, it is important to appreciate the difference between making the reasoning behind a choice explicit and subsequently communicating it. Recognising that there are multiple options to a contingent choice can be a private reflection. The same goes with the recognition that choices might involve value-judgments and affect some stakeholders more than others. A more charitable reading of the recommendation to cultivate reflective value-judgment through considering options to choices amounts to just that: making the choices and options explicit to oneself. But this specification does not completely eradicate the question of whether researchers are expected to consider all the values, standards, and stakeholders when there are very many options to consider. Although some general principle for prioritisation would be able to address such concerns, a more plausible option is that researchers should select the choices and options that seem most relevant or significant in that context. We should not expect that scientists consider the values, standards, and stakeholders for each option for a contingent choice when there are very many options to consider, but allow them to focus on a smaller set of options they deem most significant.

Finally, there is the issue of an extremely broad set of stakeholders, which risks rendering the requirement to consider the interests of all legitimate stakeholders too demanding. Especially with the broadest goal (a) of improving understanding of climate sensitivity, there are very many stakeholders, but with the more specific goals (b) and (c), it seems more plausible that scientists with the relevant expertise form the more immediate group of stakeholders. This is because they have the ability to scrutinise the research and help to ensure that the results are strong enough to be utilised for subsequent policy-relevant decisions, and thus benefit the other stakeholder groups in the longer run.

It is worth highlighting that such a centralisation of one stakeholder group can be done without downplaying the interests of other groups. With goals (b) and (c), such a centralisation seems justified since scientists and people with relevant expertise can assess whether the proposed solutions constitute an improvement and thus help to reach the broader goal of improving sensitivity estimates. Thus, alongside distinguishing between legitimate and illegitimate stakeholders, it is worthwhile to distinguish between central and peripheral stakeholders and look for ways to align their interests.[8]

The above obstacles point towards the following specifications to the toolkit of the MI ideal: (i) in addition to the initial goals of research, the MI ideal should be applied when goals emerge during the research process; (ii) our subsequent evaluations should take into consideration that the most promising options can also be the most resource-demanding; (iii) if there are very many options, values, standards, and stakeholders to consider, scientists are encouraged to focus on the ones they deem most crucial in that setting; and finally, (iv) in settings where there are very many stakeholders, it can be fruitful to distinguish between the more central and peripheral stakeholders.

5 Evaluating the practical applicability of the MI ideal: new tools

The above obstacles to the MI ideal provide an answer to the first test: its tools are broadly applicable to the case at hand. But do they help in achieving the broader goal of the MI ideal? In Sect. 3, we saw that social value-judgments enter most obviously in the broader goal (a) in part because it has a wide set of stakeholders. In contrast, the central stakeholders in the more specific goals (b) and (c) appear to be mostly scientists. Correspondingly, the majority of the identified values and standards for goals (b) and (c) were more obviously epistemic than social.

The prevalence of scientist-stakeholders and epistemic values can of course be said to be the accurate outcome of applying the MI ideal and its tools. However, there is also the risk that applying the tools did not help to reveal the relevant social value-judgments. Here, we take that possibility seriously since the prevalence of epistemic values identified for goals (b) and (c) could suggest that more concrete guidance is needed to recognise the influence of social values. Furthermore, we argue that the nested nature of goals (a-c) implies that more attention should be given to the alignment of the specific (often technical) goals with general aims which tend to have social implications and a broader set of stakeholders. Both concerns suggest that for

[8] There are different ways in which this distinction between central-peripheral can be understood. Although providing a more substantial account is beyond the scope of this paper, it should be briefly noted that we understand the central-peripheral distinction in terms of accessibility, where those stakeholders with the technical knowledge to understand the relevant changes (i.e., who can scrutinize the results) are more central than those stakeholders who cannot do so, albeit they might be impacted by the decisions in a longer run. In this sense, the central stakeholders would act as "trustees" for other stakeholder groups.

the MI ideal to reach its central goal, some additional tools are required to accompany the existing ones.

Recognising values In Sect. 3, we saw that aside from risks of error associated with climate sensitivity estimation, the majority of the identified values are more obviously epistemic than social. Although on one hand, the dominance of epistemic values indicates their relevance for climate modelling more generally, it also can speak of the difficulty of identifying social values in the more technical decisions in research. Considering the importance of recognising the presence of social-value judgments in such aspects of research, how can we aid the process of recognising social value-judgments in settings that do not immediately strike us as involving value-judgments beyond epistemic values?

One option is to supplement the MI ideal and its toolkit with more concrete field-specific examples that illustrate how values enter even the more technical decisions in science. Recall that the four activities described in Sect. 2 came with an empty worksheet. To help to recognise social value-judgments, philosophers could accompany the four guiding activities of the MI ideal with field-specific filled worksheets that show how the four guiding questions may be answered. To make such worksheets specific to climate modelling, philosophers could draw from the existing pool of discussions concerning social values in climate modelling (Winsberg, 2012; Parker & Risbey, 2015; Parker & Winsberg, 2018; Winsberg et al., 2020). Alternatively, one could draw from more general philosophical literature on values in science. In any case, having such concrete examples of social value-judgments would give its users a model of how tools can be applied and help at drawing analogies with their own research. Tangible examples would help to demonstrate that at times, social value-judgments can influence also the more technical decisions in research where their impact might not be all that obvious.

Nested goals Apart from the difficulty of recognising how social values enter the more technical decisions involved in climate modelling, the above analysis also highlighted that goals in research can be nested.[9] That is, general goals with a broader set of stakeholders may be furthered through more specific technical goals. The nested nature of goals calls to reflect on the relationship between them. If such a reflection remains undone, there is the risk that the more technical goals are misaligned with the general goals that tend to have societal implications and have a broader set of stakeholders. To use the ACI case to illustrate, reflecting on the relationship between goals calls to consider whether reducing the discrepancy between observationally-derived and model-estimated ACI furthers the general goal of improving climate sensitivity estimation. If not, such a more specific technical goal should not be prioritised over better alternatives, or another justification for the technical aims would be needed.

[9] Goals or aims of research have been discussed especially in relation to the aims approach to values in science (Elliott, 2013; Elliott & McKaughan, 2014; Intemann, 2015). To our knowledge, the question of network of aims or nested aims has not been discussed in that literature.

What are good strategies for achieving such an alignment of technical goals and broader aims? In principle, this question can be answered by the researchers working on the more specific technical goals. However, at times, determining alignment of nested goals calls for direct engagement with users of the research output and stakeholders of the broader goal. That is, aside from theoretical evaluation of the alignment of goals, evaluating their relationship would benefit from actual engagement with potential users and stakeholders of inquiry.[10] This process could involve identifying the users' priorities (see e.g., Sobel, 2021), which can inform the revision of the broader goals and prioritisation of different technical aims.

To sum, reaching the ultimate goal of the MI ideal requires accompanying the existing tools of the ideal with some further strategies. Here, we have suggested two supplementary tools: incorporation of concrete examples of how social values enter even the more technical aspects of research, and assessment on the relationship between multiple, nested goals of research. Including these two additional strategies would aid in reaching the MI ideal's broader goal of fair and warranted value-judgment.

6 Conclusion

Despite the numerous discussions regarding how value-judgments should influence scientific research, not quite as much attention has been given to the practical applicability of proposals. The above discussion gives one example of how such an evaluation might look like. In particular, we have proposed two tests for assessing practical attainability. The first test called for evaluating the applicability of the tools and strategies of the normative ideal by identifying whether there are any obstacles for their application. The second test concerned whether the tools and strategies help to reach the broader aims of the ideal. Of course, a prerequisite for making such an assessment is that there are tools and strategies for putting ideals to practice available, which is not the case for all normative proposals.

How does the MI ideal fare? The application of the MI ideal to climate modelling revealed that there were some more minor obstacles which point towards the need for sharpening some of its existing tools and strategies. Furthermore, achieving its core aims also requires the inclusion of some additional instruments to its toolkit. In particular, it would be helpful to have more concrete examples to show how social value-judgments influence even the more technical-seeming decisions in research. Furthermore, the nested nature of goals in research calls to evaluate whether the broader aims with societal implications are aligned with the more specific technical goals.

Although the above discussion dwelled on the challenges to the practical applicability of the MI ideal, it should not be taken as a recommendation against relying on

[10] The importance of actual engagement was brought up in discussions concerning Philip Kitcher's Ideal Conversation -framework, which was criticised on the grounds that it put heavy emphasis on individual's imagination instead of actual engagement with various stakeholders (Kitcher, 2011; Douglas, 2013; Keren, 2015).

it. The ideal helps scientists to recognise how not only logic and evidence plays a role in scientific decision-making, but that contingent choices in research are guided by factors such as values and standards, which can reflect the interests of stakeholders of research. It is commendable that the MI ideal is intended to foster greater awareness of values. Without such an awareness, it is difficult to see how other normative proposals regarding value-judgments can be effectively executed (Pulkkinen et al., 2022). For this reason, ensuring the practical attainability of the MI ideal can help in the application of other normative proposals which are compatible with it.

Acknowledgements KP, SU and FB thank the editors, anonymous reviewers, and Per Wikman-Svahn for very constructive and useful comments on the manuscript. KP also thanks Björn Lundgren for insightful feedback on an early draft and the audiences in the Perspectives on Science seminar in University of Helsinki and the RIVET Expert workshop on values and deliberation in climate economics held in Berlin/online. This work was supported by FORMAS (grant no. 2018-01797).

Funding Open Access funding provided by University of Helsinki including Helsinki University Central Hospital.

Declarations

Ethical statement Research funded by FORMAS, Sweden. Grant no. 2018-01797.

Conflict of interest No conflict of interest.

References

Albrecht, B. A. (1989). Aerosols, cloud microphysics, and fractional cloudiness. *Science*, *245*(4923), 1227–1230.

Anderson, E. (2004). Uses of value judgments in science: A general argument, with lessons from a case study of feminist research on divorce. *Hypatia*, *19*(1), 1–24.

Arias, P. A., Bellouin, N., Coppola, E., Jones, R. G., Krinner, G., Marotzke, J., Naik, V., Palmer, M. D., Plattner, G. K., Rogelj, J., Rojas, M., Sillmann, J., Storelvmo, T., Thorne, P. W., Trewin, B., Achuta Rao, K., Adhikary, B., Allan, R. P., Armour, K., . . . Zickfeld, K. (2021). Technical summary. In V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge University Press (in press).

Baldissera Pacchetti, M. (2021). Structural uncertainty through the lens of model building. *Synthese*, *198*(11), 10377–10393.

Bender, F. A. M., Frey, L., McCoy, D. T., Grosvenor, D. P., & Mohrmann, J. K. (2019). Assessment of aerosol–cloud–radiation correlations in satellite observations, climate models and reanalysis. *Climate Dynamics*, *52*(7), 4371–4392.

Bender, F. A. M., Engström, A., & Karlsson, J. (2016). Factors controlling cloud albedo in marine subtropical stratocumulus regions in climate models and satellite observations. *Journal of Climate*, *29*(10), 3559–3587.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J. L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., & John, V. O. (2011). COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043.

Brown, M. J. (2020). *Science and moral imagination: A new ideal for values in science*. University of Pittsburgh Press.

Douglas, H. E. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.

Douglas, H. (2013). Science in a democratic society. *The British Journal for the Philosophy of Science*, *64*(4), 901–905.

Elliott, K. C. (2013). Douglas on values: From indirect roles to multiple goals. *Studies in History and Philosophy of Science, Part A 44*(3), 375–383.

Elliott, K. C., & McKaughan, D. J. (2014). Nonepistemic values and the multiple goals of science. *Philosophy of Science*, *81*, 1–21.

Elliott, K. C. (2017). *A tapestry of values*. Oxford University Press.

Fu, Q. (2015). Radiation transfer in the atmosphere | radiation, solar. In G. R. North, J. Pyle, & F. Zhang (Eds.), *Encyclopedia of atmospheric sciences* (2nd edition, pp. 1–4). https://doi.org/10.1016/B978-0-12-382225-3.00334-0

Gettelman, A., & Sherwood, S. C. (2016). Processes responsible for cloud feedback. *Current Climate Change Reports*, *2*(4), 179–189.

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., Lamarque, J. F., Fasullo, J. T., Bailey, D. A., Lawrence, D. M., & Mills, M. J. (2019). High climate sensitivity in the community earth system model version 2 (CESM2). *Geophysical Research Letters*, *46*(14), 8329–8337.

Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., & Feingold, G. (2021). Aerosol-cloud-climate cooling overestimated by ship-track data. *Science*, *371*(6528), 485–489.

Gryspeerdt, E., Mülmenstädt, J., Gettelman, A., Malavelle, F. F., Morrison, H., Neubauer, D., Partridge, D. G., Stier, P., Takemura, T., Wang, H., Wang, M., & Zhang, K. (2020). Surprising similarities in model and observational aerosol radiative forcing estimates. *Atmospheric Chemistry and Physics*, *20*(1), 613–623.

Hall, A., & Qu, X. (2006). Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters 33*(3). https://doi.org/10.1029/2005GL025127

Holman, B. & Wilholt, T. (2022). The new demarcation problem. *Studies in history and philosophy of science*, *91*, 211–220.

Intemann, K. (2015). Distinguishing between legitimate and illegitimate values in climate modeling. *European Journal for Philosophy of Science*, *5*(2), 217–232.

IPCC (2021). *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

Kawamleh, S. (2021). Can machines learn how clouds work? The epistemic implications of machine learning methods in climate science. *Philosophy of Science*, *88*(5), 1008–1020.

Keren, A. (2015). Science and informed, counterfactual, democratic consent. *Philosophy of Science*, *82*(5), 1284–1295.

Kitcher, P. (2011). *Science in a democractic society*. Prometheus Books.

Klein, S. A., & Hall, A. (2015). Emergent constraints for cloud feedbacks. *Current Climate Change Reports*, *1*(4), 276–287.

Knutti, R. (2010). The end of model democracy? *Climatic Change*, *102*(3), 395–404.

Knutti, R., Rugenstein, M. A. A., & Hegerl, G. C. (2017). Beyond equilibrium climate sensitivity. *Nature Geoscience*, *10*(10), 727–736.

Koskinen, I. and Kristina Rolin (2022). Distinguishing between legitimate and illegitimate roles for values in transdisciplinary research. *Studies in History and Philosophy of Science*, *91*, 191–198.

Levins, R. (1966). The strategy of model building in population biology. In E. Sober (Ed.), *Conceptual issues in evolutionary biology* (1st edition, pp. 18–27). MIT Press.

Lohmann, U. (2015). Aerosols | aerosol–cloud Interactions and their radiative forcing. In G. R. North, J. Pyle, & F. Zhang (Eds.), *Encyclopedia of atmospheric sciences* (2nd edition, pp. 17–22). https://doi.org/10.1016/B978-0-12-382225-3.00052-9

Lusk, G. (2020). Political legitimacy in the democratic view: The case of climate services. *Philosophy of Science*, *87*(5), 991–1002.

Malavelle, F. F., Haywood, J .M., Jones, A., Gettelman, A., Clarisse, L., Bauduin, S., Allan, S. P., Karset, I. H. H., Kristjánsson, J. E., Oreopoulos, L. Cho, N., Lee, D., Bellouin, N., Boucher, O., Grosvenor, D. P., Carslaw, K. S., Dhomse, S., Mann, G. W., Schmidt, A., . . . Thordarson, T. (2017). Strong constraints on aerosol–cloud interactions from volcanic eruptions. *Nature,546*(7659), 485–491.

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., & Schlund, M. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the cmip6 earth system models. *Science Advances*, *6*(26), eaba1981.

Morrison, M. (2014). Values and uncertainty in simulation models. *Erkenntnis*, *79*(5), 939–959.

Parker, W. S., & Risbey, J. S. (2015). False precision, surprise and improved uncertainty assessment. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical, and Engineering Sciences, 373*(2055), 1–13. https://doi.org/10.1098/rsta.2014.0453

Parker, W. S. , & Winsberg, E. (2018). Values and evidence: How models make a difference. *European Journal for Philosophy of Science*, *8*(1), 125–142.

Parker, W. S., & Lusk, G. (2019). Incorporating user values into climate services. *Bulletin of the American Meteorological Society*, *100*(9), 1643–1650.

Pulkkinen, K., Undorf, S., Bender, F., Wikman-Svahn, P., Doblas-Reyes, F., Flynn, C., Hegerl, G. C., Jönsson, A., Leung, G.-K., Roussos, J., Shepherd, T. G., & Thompson, E. (2022). The value of values in climate science. *Nature Climate Change.* https://doi.org/10.1038/s41558-021-01238-9

Rooney, P. (1992). On values in science: Is the epistemic/non-epistemic distinction useful. *PSA 1992: Proceedings of the 1992 Biennal Meeting of the Philosophy of Science Association, 2*, 13–22.

Rooney, P. (2017). The borderlands between epistemic and non-epistemic values. In K. C. Elliott & D. Steel (Eds.), *Current controversies in values and science*. Routledge.

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., . . . Zelinka, M. D. (2020). An assessment of earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics,* e2019RG000678.

Sobel, A. H. (2021). Usable climate science is adaptation science. *Climatic Change*, *166*(1), 8.

Sotiropoulou, G., Vignon, É., Young, G., Morrison, H., O'Shea, S. J., Lachlan-Cope, T., Berne, A., & Nenes, A. (2021). Secondary ice production in summer clouds over the antarctic coast: An underappreciated process in atmospheric models. *Atmospheric Chemistry and Physics*, *21*(2), 755–771.

Steel, D. (2017). Qualified epistemic priority: Comparing two approaches to values in science. In K. C. Elliott, & D. Steel (Eds.), *Current controversies in values and science* (pp. 49–63). Routledge.

Tan, I., & Storelvmo, T. (2019). Evidence of strong contributions from mixed-phase clouds to arctic climate change. *Geophysical Research Letters*, *46*(5), 2894–2902.

Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment (1967)*, *8*(12), 1251–1256.

Undorf, S., Pulkkinen, K., Wikman-Svahn, P., & Bender, F. A. M. (2022). How do value-judgements enter model-based assessments of climate sensitivity? *Climatic Change, 174,* 19. https://doi.org/10.1007/s10584-022-03435-7

Ward, Z. B. (2021). On value-laden science. *Studies in History and Philosophy of Science Part A*, *85*, 54–62.

Weisberg, M. (2006). Robustness analysis. *Philosophy of Science*, *73*(5), 730–742.

Winsberg, E. (2012). Values and uncertainties in the predictions of global climate models. *Kennedy Institute of Ethics Journal*, *22*(2), 111–137.

Winsberg, E. (2018). What does robustness teach us in climate science: A re-appraisal. *Synthese*, *198*(21), 5099–5122.

Winsberg, E., & Oreskes, N., & Lloyd, E. (2020). Severe weather event attribution: Why values won't go away. *Studies in History and Philosophy of Science Part A*, *84*, 142–149.

Zhao, M., Golaz, J. C., Held, I. M., Guo, H., Balaji, V., Benson, R., Chen, J. H., Chen, X., Donner, L. J., Dunne, J. P., Dunne, K., Durachta, J., Fan, S. M., Freidenreich, S. M., Garner, S. T., Ginoux, P., Harris, L. M., Horowitz, L. W., Krasting, J. P., . . . Xiang, B. (2018). The GFDL global atmosphere and land model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *Journal of Advances in Modeling Earth Systems,10*(3), 691–734.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Karoliina Pulkkinen[1,2] · Sabine Undorf[3,4] · Frida A.-M. Bender[4]

✉ Karoliina Pulkkinen
karoliina.pulkkinen@helsinki.fi

Sabine Undorf
sabine.undorf@pik-potsdam.de

Frida A.-M. Bender
frida.bender@misu.su.se

1 Aleksanteri Institute, University of Helsinki, Unioninkatu 40, 00170 Helsinki, Finland

2 Division of Philosophy and History, KTH Royal Institute of Technology, Stockholm, Sweden

3 Research Department for Climate Resilience – Climate Impacts and Adaptation, Potsdam Institute for Climate Impact Research, Potsdam, Germany

4 Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden