Title: 'Running' climate change mental models reveals poor convergence between worldviews, policy preferences and expected consequences

Authors: Claire Richert*[1], Fabio Boschetti[2,4], Iain Walker[3,6], Jennifer Price[3,6] and Nicola Grigg[5]

Affiliations:

[1]IRSTEA UMR G-EAU, France.

[2]CSIRO Marine and Atmospheric Research & Wealth from Ocean Flagship, Australia.

[3] CSIRO Social and Economic Sciences Program, Australia.

[4] School of Earth and Geographical Sciences, The University of Western Australia

[5]CSIRO Land and Water, Australia.

[6] CSIRO Climate Adaptation Flagship, Australia

*Correspondence to: Claire.richert@irstea.fr

Abstract: Understanding complex problems such as climate change is difficult for most non-scientists, with serious implications for decision making and policy support. Prominent recent research has identified the poor appreciation of the simple foundations of system dynamics at the root of this difficulty. Here we provide evidence of another causal mechanism: two different processes are at play in formulating versus evaluating mental models of the climate change debate. Ideology and worldviews define the mental model, that is, how we believe the system works today. Expectations about the future define the projected dynamical behaviour, that is, how we believe the system will look in the future. The mismatch between these cognitive tools may prevent establishing a coherent link between the mental model assumptions and consequences, thus potentially affecting decision making and behaviour changes including policy support. Simple computer models provide an avenue to 'run' mental models to assess whether a system will evolve according to our aspirations, under various policy choices.

One Sentence Summary: 'Running' climate change *mental* models reveals that *formulating* mental models of climate change and *evaluating* their likely consequences rely on different processes.

Main Text:

1 Introduction

When we are asked to provide an opinion on a complex issue (as most environmental problems are) and we take the task seriously, we are likely to follow a number of steps. First, we need to conceptualise the problem, understand the question and consider how the system underlying the issue works. We refer to this understanding as a 'mental model'. In the case of climate change, this may require a rough understanding of how the atmosphere works and its relation to human activities. Second, we need to contextualise the issue. For example, climate change might be framed as an environmental, economic, social or ethical issue, or (ideally) as a combination of the four. Finally, we will likely position the question within our overall worldviews, and relate it to our expectations, desires and fears. Mostly, mental models and their conceptualisation are likely not to be consciously worked through.

The above description focuses on the individual, but we live in a tight network of social relations and most of us will adopt, rather than develop, a mental model. In the case of climate change, few individuals have the expertise necessary to conceive a functional model of the relations between global atmospheric circulation, economy, energy production, population growth, social processes and contemporary culture. Most scientists will adopt mental models from the scientific literature. Non-scientists will do the same from the media, books, blogs, friends or family. Both will have to decide what information to trust and what to disregard. Naturally, mental models, framing and information-filtering are inter-related, since they come as a single package suited to fit predetermined worldviews and aspirations (*1-10*).

The picture we have drawn so far is complex, increasingly well understood (*11*), but incomplete. When a system has a strong dynamical component then mental models, framing and attitudes – which are static in form – are insufficient for informed decision making. The dynamics themselves need to be included and considered. In other words, the mental models need to be 'run'. It is only by doing so that we can assess whether a system will evolve to provide the outcome we aspire to.

Unfortunately, humans are poor at running mental models (*11-18*). In the specific case of climate change, Sterman (*19*) showed that misunderstanding the relation between CO_2 emissions, sequestration and accumulation often leads many people to misjudge the consequences of specific climate policies. We have replicated this experiment with members of the public, decision makers and scientists (including computer modellers) and obtained similar results (*20*). Similar conclusions were also reached by Moxnes and Saysel (*11*). It is now well established that people, even those gifted with high cognitive skills, are prone to poor decision-making because of poor understanding of basic system dynamics, unhelpful cognitive attitudes, and a host of heuristics used, often unconsciously, to cope in a complex, information-rich environment (*12, 21, 22*).

Scientists use computer models to avoid these problems caused by cognitive limitations. Furthermore, computer models can be used specifically to train our appreciation of dynamical processes, and online tools and gaming infrastructures can be used for the same purpose. However, while these tools allow users to change some parameters and explore their influence on the model outcome, the 'mental model' underlying these tools is usually predefined and fixed (e.g. via the model assumptions or parameterisation used). Here we describe a simple computer model whose parameters effectively define the underlying mental model, meaning that different parameter sets allow for fundamentally different mental models to be represented. The basis of the model is the interaction between population growth, the economy and global warming, but the relative weight of the components and the importance of the related processes can be modified via parameter setting to represent mental models as seen from a citizen supporting free-market, eco-centric or politically moderate views[1]. This effectively allows anyone using the model to explore the likely outcomes of different policy choices arising from different worldviews and different interpretations of the problem.

We examined whether the computational model's predictions from initial parameterisations provided by participants, based on their beliefs and values, would correspond with, or diverge from, the future world those same participants imagined when asked directly.

Via an online survey, we described the functioning of the APE model (see Material and Methods), and then asked 130 Australian responders to parameterise the model according to their beliefs and values by choosing a numerical value for each of 6 parameters in Table 1. Next, we asked them to predict the long-term impact of their choices on the model output variables (population size, wealth and warming), that is, to mentally execute their internal mental models. Specifically, we asked "*Assuming the policies you chose are implemented, what do you think will happen by 2100 compared to today*?". Then, we ran the APE model as parameterised by the responders. This can be seen as modelling the numerically and dynamically consistent consequences of the responders' beliefs and values, under the assumptions inherent in their mental model. In this way we could gauge the (dis)continuity between their intuitive expectations and the dynamic form of the mental model they indicated that they subscribed to.

In addition, we checked whether the responders' worldviews and attitudes i) match the values and beliefs implicit in their chosen model parameters and ii) affect the prediction of their consequences. We did so by measuring the following cognitive constructs and beliefs (*20*): attitudes toward the environment and science, beliefs in climate change, Consideration of Future Consequences, political ideology and expectations about the future evolution of Australian society (whether by 2050 Australian society will be more or less safe, skilled, wealthy, honest and friendly). We refer to the pattern of responses across these variables as a cognitive signature. The full questionnaire can be found in the Material and Methods, together with reference to the scales used and the analysis of the survey data.

The responders' chosen parameterisations were largely in line with the climate change literature and correlates well with their cognitive signature. This suggests that the responders understood the model, the role of the model parameters and their relation to their own mental model. Nevertheless, the match between

[1] Here we employ the stereotypes 'free-market supporter' for citizens who place *priority* on economic over ecological values (and vice versa for 'eco-centric' citizens), while acknowledging that concerns for environmental issues and support for market approaches are not necessarily mutually exclusive.

the projections of the mental and numerical models was poor. This is consistent with the cognitive limitations of humans identified in the above-mentioned literature. However, our results suggest a novel alternative explanation pointing to a fundamental difference between the cognitive and affective resources used in model *formulation* vs model *evaluation*.

2 Mental and numerical models

Previously (*23*) we have implemented a simple numerical model of the interaction between human Activity, Population and the Environment (APE) able to represent some mental models the public is likely to adopt in thinking about climate change. Human activity is represented by GDP per capita, which we refer to as 'wealth' in this paper. It is well known that GDP is an impoverished indicator of wealth (*24*) and we used it simply because it is still the most widely adopted and measured indicator of wealth, and would be widely understood by survey participants. In this work, we simplify the APE model further (see Material and Methods) to ensure that its overall behaviour and parameterisation can be more easily described in an online survey. Among the APE model parameters, six are particularly significant for the climate change debate (see Table 1) and hold an intuitive meaning in terms of values and beliefs. Mitigation Target and Mitigation Timeline are core targets of policy-making, since they represent the extent and urgency of mitigation initiatives. They thus represent *values*: the priority we give to addressing climate change. Climate Sensitivity and Critical Temperature represents *beliefs* about the science of climate change. The scientific community has proposed expected values and confidence ranges for these parameters (see for example (*25, 26*)) which citizens may or may not trust. Finally, Mitigation Cost and Human Carrying Capacity lie at the intersection between values and beliefs. On the one hand, they represent beliefs about how economic and social processes affect and are affected by climate change: estimates of Mitigation Cost reflect how we believe our economic system works, and Human Carrying Capacity reflects our belief of the extent to which our planet can sustain human life. Importantly, these beliefs are influenced by and can be moulded to fit specific worldviews.(*7, 10, 27, 28*)

Table 1: The model parameters which capture climate change values and beliefs: Name and symbols used in Materials and Methods (column 1), their meaning (column 2), and the numerical choices offered to the responders (column 3). In order to make the subsequent analysis clearer, these choices are ordered left-to-right according to expected eco-centric versus free-market-centric attitudes, as suggested in (*8, 29*).

Parameters	**Meaning**	**Values**
Mitigation Target (Mit_{Target})	Target percentage reduction in emissions	90, 45, 5, 0 (%)
Mitigation Timeline (Mit_{End})	Year when the target mitigation is achieved	2020, 2060, 2100
Climate Sensitivity (λ)	Increase in global average temperature per doubling of CO2 in the atmosphere	3.4 , 2.2, 1.0 (°C)
Critical Temperature (ΔT_{Crit})	The increase in temperature above which human activity will be affected	1.7, 3.3, 5.0 (°C)
Mitigation Cost (Mit_{Cost})	Cost of mitigation in % of global GDP	0, 10, 20 (%)
Human Carrying Capacity (K)	Maximum number of people the Earth could sustain	15, 27, 40 ($*10^9$)

Given its inherent simplification, the APE model explores the dynamically consistent environmental and economic consequences of different sets of beliefs and values (see Material and Methods) and thus allows us to compare the output of the *simulated* mental models with the output of the *actual* mental models as found among the general public.

3 Results

Fig 1 summarises the results by showing significant correlations ($p<0.05$). The top panel shows that the responders' choices of Mitigation Target, Mitigation Timeline, Climate Sensitivity and Critical Temperature are tightly and consistently correlated, confirming the eco-centric vs free-market-centric dimensions discussed in the literature. As a result, in the rest of the paper, we combine these parameters into a single measure reflecting the responders' values and beliefs. In addition, the distribution of these parameters matches with current scientific agreement (*26*, *30*-*32*) (see Material and Methods). The Mitigation Cost parameter is correlated only to Mitigation Timeline and appears to be overestimated (*33*-*35*).

A key outcome is in the Prediction panel. The mental models produce a positive correlation between the predictions of warming and wealth (left plot). Overall, these predictions range between a 'poor and climate-safe' future (blue arrow in the left plot) and a 'wealthy and climate dangerous' one (red arrow). However, the APE model produces a negative correlation (right plot), in which no 'poor and climate-safe' future is found. This is the main source of discrepancy between the responder's intuitive expectations and the model predictions. Note that the existence of a negative feedback between warming and wealth was specifically included in the description of the APE model as provided to the responders (see Materials and Methods).

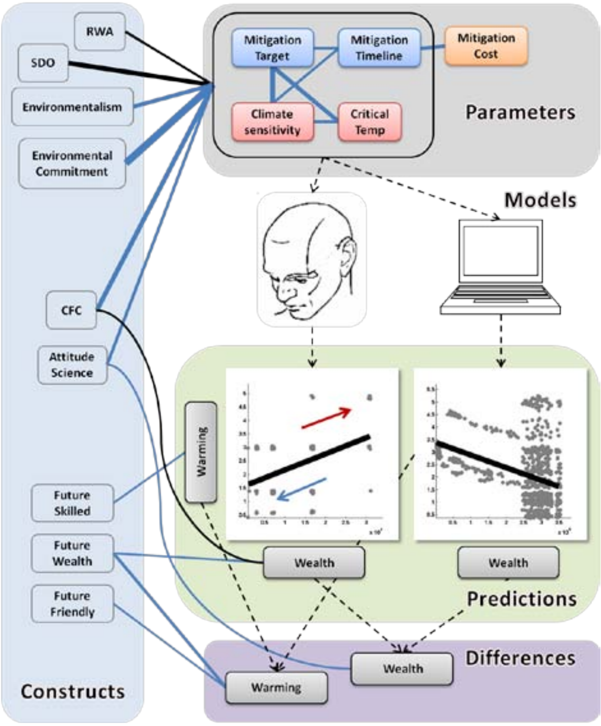


Fig 1: Visual summary of the correlations between cognitive constructs (left, blue plate) and i) model input Parameters (grey plate), ii) model Predictions (green plate) and iii) Differences between predictions from the mental and numerical model (purple plate). The differences are largely the result of the responders' predicting a positive correlation (0.431, $p<0.05$) between future warming and wealth (left plot in the Prediction panel) while the APE model run over these same parameter sets shows a negative correlation (-0.429, $p<0.05$, right plot in the Prediction panel). See 'Model Runs' in Material and Methods for a detailed discussion of the APE model output. Blue and black links represent positive and negative significant correlations ($p<0.05$), respectively, and link thickness maps correlation strength (see Material and Methods for numerical values for the correlations).

The Constructs panel on the left hand side of Fig 1 shows the cognitive constructs. Their correlations with the model parameters and results display a clear pattern. The cognitive constructs related to political ideology and attitudes towards the environment (top Constructs panel) correlate well and consistently with the APE model input parameters (see Materials and Methods). These correlations are also consistent with an eco-centric vs free-market-centric polarisation. However, political ideology and attitudes towards the environment show no correlation with the model predictions and differences.

The expectations about the future Australian society (bottom Constructs panel) show some correlations with predictions and some differences between mental models and the APE model. Among these, the strongest correlations are displayed by the expectation of future wealth. Responders who expect a wealthier future Australia display higher discrepancy between their prediction and the model prediction when the model is run with their parameter choice. They also predict, as expected, a wealthier future. Responders who expect a more skilled future Australia predict a higher level of warming, which we tentatively interpret by associating skills with technology and technology with industrial production. Unlike political ideology and attitudes towards the environment, these expectations about the future do not show any statistically significant correlation with the model input parameters.

The only constructs displaying some correlation with model input parameters as well as predictions and differences are the Consideration of Future Consequences and Attitude towards Science. Nevertheless, these correlations are considerably stronger towards the model input parameters.

4 Interpretation

The core message from the mental model projections is that our responders seem to predict alternative futures ranging between a ‘wealthy and climate-dangerous’ future and a ‘poor and climate-safe’ one. Why is it so? A few interpretations are on offer. Despite including the negative feedback between warming and wealth in the description of the APE model, most responders seem to dismiss this possibility. This suggests that our responders imagine that wealth may be impacted by initiatives to combat climate change, not by climate change itself. It could represent an associative bias arising from the oversimplified climate change debate in the media which portraits a dichotomy between ‘economic and hedonistic’ vs ‘environmental and frugal’ choices. However, in this study, this dichotomy does not correlate to ideology and attitudes, rather to expectations about the future. This points to a mismatch between the cognitive resources used to formulate the model on the one hand and to predict its dynamical evolution on the other, which leads to treating model formulation and dynamic assessment as two different processes. This is probably not surprising if we consider that, as mentioned above, mental models are often inherited, not developed, and that long-term predictions over many decades are rarely discussed, contemplated or even imagined. Our previous research indeed suggests that many responders are rarely, if ever, asked to formulate long term projections and that in doing so have to rely on cognitive resources different from the ones usually engaged in the climate change discourse.

A second interpretation points to the difficulty in assessing the dynamical relations between economic growth, population dynamics and CO2 accumulation. This would represent one more instance of the above mentioned challenges humans encounter in thinking in dynamical terms. Since we ourselves have seen this process at play in many circumstances, we do not exclude this interpretation. However, this is at odds with the observation that model projections and differences correlate with expectations about the future. If any cognitive construct were to correlate with this dynamical fallacy, the literature would suggest it would with attitudes and ideology (*27*, *28*).

Our results do not contradict the ample literature on the impact that logical fallacies (*21*, *22*), numeracy and scientific literacy (*1*, *36*) and poor appreciation of dynamical processes can have on effective decision making. Rather, they highlight a further causal path, which may interfere with and reinforce the former. Notwithstanding the finite number of options we presented during the parameterisation task, our responders seem to be drawing on a limited number of narratives available (as obtained from the media, peers, scientific literature, etc) in choosing a description of the problem (on a free-market-support vs eco-centric dimension) and a possible outcome (on an ‘economic and hedonistic’ vs ‘environmental and frugal’ dimension). Further, they appear to be using different sets of cognitive tools for the two choices, leading to description-outcomes pairing which may be dynamically inconsistent. Neither set of cognitive tools seems to focus on evaluating the degree of consistency between the two outcomes. Perhaps a ‘meta-cognitive’ process must be involved for that to happen; if so, a limitation of our work is that we have not directly assessed it.

Our analysis relies on the APE model being a reliable representation of the responders' mental model. It is possible that the responders may have misunderstood or not paid attention to the model description. Furthermore, the equations in the numerical model (see Materials and Methods), which determine the actual model dynamics, were not described in the survey. This leads us to consider whether the responders may have used a different mental model or, more generally, whether there is a single, unique numerical model which better describes the responders' mental models. This question would need to be explored in an interview or focus group setting. However, our results already elucidate some features of this hypothetical model, should it exist. First, this model would not be linear in the input (values and beliefs) parameters since no such indication is found in the data, as discussed above and captured graphically in Fig 1. The above mentioned cognitive challenges humans encounter in addressing dynamical problems lead us to question whether a more complex, non-linear model would be spontaneously developed by the responders. Second, some responders predict a 'climate-safe and poor' future which does not belong to the range of outcomes generated by the APE model as currently parameterised (Fig 1). This could be obtained via the APE model by allowing for a much larger (vastly unrealistic) mitigation cost. This may suggest that the responders misjudged the effect of mitigation cost or were implicitly using a larger mitigation cost, perhaps in the form of opportunity costs or related impacts (which are inherent in our representation of 'mitigation costs' but may not have been interpreted that way by respondents). This is unlikely though because the responders who chose high mitigation cost do not show significant differences in their predictions for either wealth or warming. While the possibility of our responders adopting different and possibly multiple mental models in their analysis cannot be dismissed, a more economical possibility at the population scale is captured by Fig 1 and proposed above: our responders' mental model is better parameterised by cognitive constructs (blue panel) than by the APE model parameters (gray panel).

Computer modelling and cognitive predispositions such as self-reflection, slow thinking, critical analysis and abstraction are recommended to address the cognitive challenges arising from dynamical processes (*12, 21, 22*) and we expect would also be beneficial (as suggested in (*37*)) in addressing the gap between model formulation (how we think the world works today) and projections (how we think the world may work in the future). In addition, our study suggests an alternative way in which computer modelling could be employed in similar training. Computer models can be run in a forward or inverse manner (*38*). Loosely speaking, forward modelling fixes the initial state and calculates future states (as in this work), while inverse modelling fixes the future state and asks what initial states may lead to it.

Our observation that expectations about the future affect the assessment of the model dynamics but not the choice of the model itself has some important implications. Debates about anthropogenic climate change and the urgency or otherwise of significant mitigation action are wracked by ideologically-based differences. Those debates rest more on analyses of the present than of the future. Inverse modelling, or back-casting, may help to bring aspirations to the fore, pushing the role of ideologies and attitudes into the background, and perhaps will lead to a better evaluation of choices currently available to us.

Materials and Methods:

1 The climate change APE numerical model

Here we describe the APE model used in this study. This model is a simplification of (*23, 39*) and describes the evolution of variables which are both abstract and aggregate, while the actual interplay of biophysical, social, economical, ecological and political processes affecting climate change is far more complex and plays out at several scales. We accept this simplification because we do not aim to model the dynamics of the *actual* system, rather the way people represent and reason about it.

The full model uses Euler integration to solve the following differential equations

$$\frac{\mathrm{d}P}{\mathrm{d}t} = P\left[k_{\mathrm{birth}}(P,A) - k_{\mathrm{death}}(P,A)\right]\left[1 - P/K\right] \tag{1}$$

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \left[k_{\mathrm{growthCap}} + k_{\mathrm{growthLab}}\frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}t} - Mit_{Cost} - k_{\mathrm{damage}}(E)\right]A \tag{2}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = c(t)k_{\mathrm{energy}}(A,P)P \tag{3}$$

$$E(\mathrm{Q}) = \frac{\lambda}{1 - \frac{5.35\lambda\gamma}{C0 + AQ}}\left[5.35\ln\left(1 + \frac{AirB \cdot Q}{C0}\right)\right] \tag{4}$$

$$c(t) = \begin{cases} C_0 - \text{Sequestration} & t < Mit_{Start} \\ C_0 e^{-\log(1-Mit_{Target})\frac{(\text{t-Mit}_{\text{Start}})}{(\text{Mit}_{\text{End}}-\text{Mit}_{\text{Start}})}} - \text{Sequestration} & t \geq Mit_{Start} \end{cases} \tag{5}$$

Where A, P, and E are the APE state variables described in the main text, k_{birth}, k_{death} and k_{energy} represent birth, death and energy use rates as a function of GDP per capita, as estimated from UN data sets (http://unstats.un.org and http://data.un.org) over the 1970-2008 period (*23, 39*); K is the Human Carrying Capacity; $k_{growthCap}$ is the economic growth due to capital and financial accumulation and $k_{growthLab}$ is the economic growth due to labor growth (*23, 39*).

Equation 4 relates cumulative CO_2 emissions since 1751 (Q) to peak temperature which is used as an approximation of warming E (*26*). Here, *C0* is the preindustrial mass of atmospheric CO_2 (= 596.4 PgC); *AirB* is the cumulative airborne fraction of CO_2 in the absence of climate feedback on the carbon cycle (=.294); λ is the Climate Sensitivity, and γ is the aggregate sensitivity of all land and ocean carbon pools to climate change (=40 PgC K^{-1}).

Equation 5 defines the carbon density of energy use c(t) (*23, 39*), which implements the climate mitigation measures. Here C_0 is the initial carbon intensity per unit energy (set to 20 g Carbon/MJ see (*40*)), *Sequestration* is the natural sequestration, which is fixed at 10% of C_0 (*40*), Mit_{Start} is the year when mitigation initiatives start to take place (2015) and Mit_{End} is the year when the planned mitigation target (Mit_{Target}) is reached.

k_{damage} represents the climate damages on the economy. We assume no damage occurs until the Critical Temperature ΔT_{Crit} is reached; past that value, we have $k_{\mathrm{damage}} = \min\left\{1, e^{r_{dam}(E-\Delta T_{Crit})} - 1\right\}$. Here we chose r_{dam} = 0.0693 K^{-1} which leads to a full economic collapse only for a warming larger than ΔT_{Crit} +10 degrees, far exceed most current predictions.

2 Model runs

In addition to the parameterisations selected by the responders, we sampled the APE model space defined in Table 1 exhaustively, in order to explore the full potential range of long-term system responses. Fig.1 shows the state of the system in the year 2100, projected over the AE plane (Warming vs Wealth).

The model runs initialised with the responders' parameterisation are shown as red dots, while the full set of exhaustive runs is shown as grey dots.

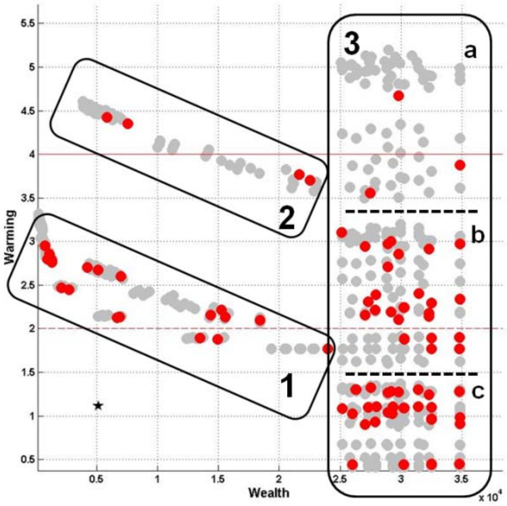


Fig.1: APE model output in year 2100. The original 3D APE output is projected over the AE (Warming vs Wealth) plane. Grey dots show all combinations of the input model parameters. Among these, the red dots show the model output for the responders' choices. The black star marks the initial conditions (year 2005). The dashed and thick red horizontal lines define the 2°C and 4°C temperature increase, which are commonly recognised as the threshold of dangerous and extreme global warming, respectively (*32*). The relation between the model parameters and the clusters (1, 2 and 3a,b,c) is described below.

We identify three regions in the output model space shown in Fig.1, which are pertinent to the responders' predictions[1]:

1) Region 1 represents the climate impact on the economy when the Critical Temperature is 1.7 °C and warming exceeds it. It is characterised by a clear negative correlation between warming and wealth. In this region, different choices of mitigation policies and different values of climate sensitivity have a considerable impact on warming and as a result, an even larger impact on wealth, which can vary between a richer and poorer future. The negative feedback between these two variables prevents reaching the 4°C extreme warming threshold, because significant damage to the economy is already imposed at relatively low temperatures.

2) Region 2 shows a similar behaviour to case 1, but for a Critical Temperature of 3.3 °C. In this case the model output lies around the 4°C extreme warming (marked by a red reference line in Fig.1) and occasionally exceeds it.

3) In Region 3 warming does not impact the economy. This occurs because warming does not exceed either a Critical Temperature of 1.7°C (Region 3c), 3.3°C (3b-c) or 5°C (3a-c). This region is characterised by a very wide variability in warming as a function of mitigation policies and climate sensitivity. These however do not impact wealth, which is affected only by mitigation cost and population size. Because mitigation cost affects GDP growth less than climate damage, this region is characterised by considerably richer future compared to today.

3 The online questionnaire

A copy of the full questionnaire can be found at: http://www.per.marine.csiro.au/staff/Fabio.Boschetti/Surveys/APE-MentalModels.pdf. It has been

[1] In order to better interpret Fig.1 it is useful to point out that sensitivity analysis (not included) shows that the model results are mostly controlled by 3 parameters: Mitigation Target, Climate Sensitivity and Critical Temperature, and to a lesser extent by Mitigation Cost. Mitigation Timeline and Human Carrying Capacity have a less significant impact.

completed by 130 individuals, chosen as unbiased representatives of the Australian population[1]. The questionnaire included a number of sections addressing different constructs and cognitive measures as well as the APE model. The questionnaire sections directly relevant to this study include:

1) Attitude toward the environment. This was explored via i) the responders' stated environmentalist attitude (Section 14, referred here as "environmentalism") and ii) the "environmental commitment and concern" module (Section 13), which we had previously tested in (*20*).

2) Attitude toward science. In order to evaluate the responders' attitude towards the specific model used in this study, we wished to assess their attitudes towards computer models in general. Unfortunately, a quarter of the responders did not know what computer modelling is. As a result, we explored their attitude toward science (*20*) (Section 10) as a proxy. The responders who claimed to know what modelling is were also asked to state how much they trusted the three stable relationships of the APE model which are not affected by the parameters. This question allows us to estimate whether the responders believe the APE model is an appropriate representation of their understanding of the problem.

3) Time preference. Environmental issues, and climate change in particular, are often described as having very slow dynamics. Hence, we wanted to analyse the relation between responders' attitude toward the future and their model parameters choices. We employed a short version of the Consideration of Future Consequences scale (Section 2, from (*41*)).

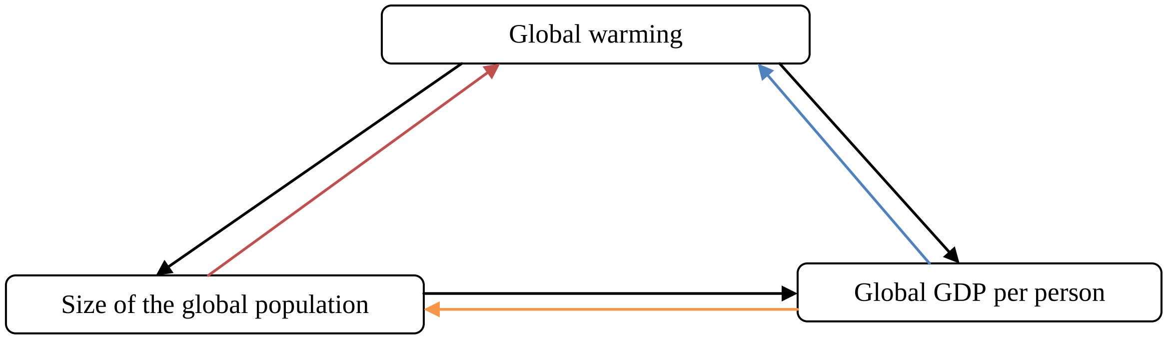
4) Ideologies. The political implications of the climate change debate have been extensively studied (*3-5*). We included the Social Dominance Orientation (SDO) scale (*42*) and the Right-Wing Authoritarianism (RWA) scale (*43*) in our questionnaire (Sections 4 and 5, respectively).

5) Expectation about the future evolution of Australia society (*44*). In particular, we asked whether the responders believe that by 2050 Australian society will be more or less safe, skilled, wealthy, honest and friendly (Section 15, questions 6-10).

We also included two questions to evaluate the extent to which the responders believe in climate change. We asked i) if they believe climate change is happening (referred as Climate Change Q1 below) and ii) what proportion of it has an anthropogenic source (Climate Change Q2). These questions were originally developed in (*45*).

In the online questionnaire, the three variables of the APE model were described as follows:

Consider a global system made up of three variables: i) global economy (represented by the global Growth Domestic Product (GDP) per person. As the global GDP per person is the value of all the goods and services produced in the world divided by the number of people, it indicates the strength of the economic activity), ii) the size of the global population (P) and iii) the state of the climate (E, for the environment, represented by the global temperature rise).



Then, the responders were presented with three stable relationships that affect the variables:

Impact of the global economy on Population:

[1] The panel used is administered by ORU, an online fieldwork company with QSOAP 'Gold Standard' and the new Global ISO 26362 standard accreditation. The ORU has a database of over 300,000 individuals from across Australia (http://www.theoru.com/).

• When GDP per person increases, the birth rate decreases. It is empirically shown (at least in developed countries) that the wealthier people become the less children they have. This in turns may lead to population decline.

<u>Impact of Activity on Environment</u>:

• When GDP per person increases, energy use per capita increases. As a result, the amount of CO_2 in the atmosphere increases.

<u>Impact of Population on Environment</u>:

• When the global population increases, the amount of CO_2 in the atmosphere increases.

Finally, the responders were presented with a description of each of the parameters listed in Table 1 and were asked to assign them one of the value chosen amongst the possibilities proposed in Table 1:

Relationships affected by beliefs

1. Maximum number of people who could live on the Earth (Earth carrying capacity):

We all know the Earth is finite in size and resources. Thus, it can't support an infinite number of human beings. In your opinion, what is the maximum number of people who could live on the Earth? (In 2012, there are approximately 7 billion people in the World).

☐ *15 billon* ☐ *27 billion* ☐ *40 billion*

2. Critical temperature:

We believe that if the global warming reaches a certain value (the "critical temperature value") the human activity will be affected and the GDP per capita will decrease. However, scientists don't agree on the value of this critical temperature.

- *The most optimistic believe that we won't see any effect on the economy until the global warming reaches 5°C.*
- *The most pessimistic believe that climate change will start to affect the economy from an increase in the global temperature of 1.7°C.*

In your opinion, what statement is the most likely?

Climate change will start to affect the economy from a global warming of:

☐ *1.7°C* ☐ *3.3°C* ☐ *5°C*

3. Climate sensitivity:

Scientists don't agree on how much the rise of CO2 in the atmosphere affects the global temperature.

- *The most optimistic believe that if the amount of CO2 in the atmosphere doubled, the rise in global temperature would be 1°C.*
- *The most pessimistic believe that if the amount of CO2 in the atmosphere doubled, the rise in global temperature would be 3.4°C.*

In your opinion, if the amount of CO2 in the atmosphere doubled, the rise on global temperature would be:

☐ *1°C (weak sensitivity)* ☐ *2.2°C (mild sensitivity)* ☐ *3.4°C (strong sensitivity)*

Relationships affected by opinions

According to your beliefs regarding the parameters above, give your opinion about the following policies:

1. *We should reduce the human global emissions of CO_2 by:*

☐ *There is no need to reduce our emissions (go to the "Predictions" subsection)*

☐ *5%* ☐ *45%* ☐ *90%*

2. *90% of the goal chosen above should be achieved by:*

☐ *2020* ☐ *2060* ☐ *2100*

3. *If we manage to reach the goal you chose, how much do you think it will cost:*

☐ *0% of the GDP* ☐ *10% of the GDP* ☐ *20% of the GDP*

4 Parameters' choice and cognitive constructs

Table 2 shows the distribution of the model parameters as chosen by the responders. Table 3 shows the correlations between the 6 APE model parameters as chosen by the responders, which are represented graphically in the top panel of Fig 1. Most of these correlations are significant and consistent with the literature.

Table 4 shows the mean and standard deviation of the cognitive constructs, their correlation with the beliefs and values model parameter choices and the adjusted R-squared from the regression over all constructs for each model parameter. For each cognitive construct, the statistically significant correlations with the model parameter choice are all of the same sign and, again, in line with the expectations as suggested by the literature. This gives us confidence that the responders had a fairly clear understanding of the parameterisation process and chose the model parameters consistent with their values and beliefs. In other words, this gives us confidence that, after parameterisation, the APE model represents a reasonable description of the responders' mental models.

A few more observations can be drawn from the analysis of Table 2, Table 3 and Table 4:

1) The distribution of the beliefs parameters matches quite well current scientific agreement (*26*, *30*-*32*).

2) The vast majority of the responders (68.5%) believe that the Earth's carrying capacity is 15 billion people. As a result, this parameter does not provide much discriminatory power and we excluded it from subsequent analysis.

3) Compared to current technical opinion (*33*-*35*) the Mitigation Cost appears to be overestimated. It is also the parameter most strongly correlated to the Mitigation Timeline. This suggests an interpretation according to which perceptions of mitigation cost may drive choices in Mitigation Timeline: the more expensive mitigation is judged, the more the mitigation initiatives are procrastinated or spread over time. Conversely, Mitigation Timeline is also significantly correlated to Mitigation Target and Climate Sensitivity. This suggests an alternative avenue, according to which Mitigation Timeline is coupled to environmental concerns. It appears that the choice of Mitigation Timeline is the parameter which best represents the environment vs economy dilemma, as stereotyped in the media.

Table 2: Percentage distribution of responders' parameters choices.

Parameters	Possible values	Percentage of responders (%)
Mitigation Target	90%	13.08
	45%	44.62
	5%	32.31
	0%	10.00
Mitigation Timeline	2020	49.23
	2060	33.85
	2100	16.92
Climate Sensitivity	3.4°C	30.77
	2.2°C	47.69
	1.0°C	21.54
Critical Temperature	1.7°C	34.62
	3.3°C	43.85
	5.0°C	21.54
Mitigation Cost	0% of the GDP	10.77
	10% of the GDP	49.23
	20% of the GDP	40.00
Human Carrying Capacity	15 billion	68.46
	27 billion	23.08
	40 billion	8.46

Table 3: Correlations among the 6 APE model parameters as chosen by the responders (*: p-value<0.05)

	Mitigation Timeline	Climate Sensitivity	Critical Temperature	Mitigation Cost	Earth carrying capacity
Mitigation Target	0.26*	0.47*	0.37*	0.14	-0.02
Mitigation Timeline	-	0.22*	0.10	0.32*	0.21*
Climate Sensitivity		-	0.27*	-0.02	0.21*
Critical Temperature			-	-0.01	0.11
Mitigation Cost				-	0.09

Table 4: Mean and standard deviation of the cognitive constructs (left) and their correlation with the beliefs and values model parameter choices (*=p-value<0.05) (right)). The bottom row (grey) shows the adjusted R-squared from the regression over all constructs for each model parameter.

Mean	Standard deviation		Mitigation Target	Mitigation Timeline	Climate Sensitivity	Critical Temperature
3.57	0.63	Consideration of Future Consequences	0.39*	0.26*	0.35*	0.20*
2.11	0.65	Social Dominance Orientation	-0.37*	-0.26*	-0.34*	-0.19*
3.62	0.83	Right-Wing Authoritarianism	-0.26*	-0.03	-0.18*	-0.11
3.72	0.61	Attitude toward science	0.31*	0.10	0.24*	0.14
2.90	1.0	Environmental commitment	0.56*	0.34*	0.49*	0.4*
3.67	0.56	Environmentalism	0.40*	0.12	0.26*	0.20*
2.37	0.93	Future safe	0.11	-0.05	0.02	0.09
2.39	0.84	Future honest	0.11	-0.04	-0.05	0.18*
2.52	0.93	Future friendly	0.14	0.01	-0.07	0.10
2.75	1.14	Future skilled	0.20*	-0.02	0.06	0.08
		Adjusted R-squared	0.36	0.10	0.24	0.14

5 Climate change beliefs and trust in the APE model

The beliefs regarding climate change and the level of trust in the APE model are not cognitive constructs but opinions about the issues addressed and tools used in this work. As a result, they are not included in the analysis of the APE model's parameterization and predictions.

As shown in Table 5, only half of the responders claim to believe that climate change is happening and is human driven. In principle, the remaining half should not agree, or agree only partially, with the APE model's fundamental assumption that economic activity drives global warming. However, when asked more specifically how much human activities contribute to climate change, less than 5% of the responders rejected the anthropogenic cause completely. The responses to the two climate change questions and the apparent lack of consistency between them match data obtained in other surveys among the Australian public (*45*) and confirm the representativeness of our sample. It suggests that both the order in which the questions are asked and the lack of more nuanced items in the proposed set of statements in Climate change Q1 may impact on the responders' choices. Finally, only one quarter of the responders do not trust the APE model (see Table 5), with the majority claiming it can be an appropriate, albeit simplified, representation of the problem.

Table 5: Responses to the two questions regarding climate change and the level of trust in the model.

Question	Possible responses	Percentage of responders (%)
Climate change Q1: "What best describes your thoughts about climate change?"	"I don't think climate change is happening."	5.38
	"I have no idea whether climate change is happening or not."	2.31
	"I think it's happening but not due to human activities."	41.54
	"I think it's happening and due to human activities."	50.77
Climate change Q2: "How much do you think humans contribute to climate change?"	0%	3.85
	15%	19.23
	30%	12.31
	45%	10.77
	60%	14.62
	75%	17.69
	90%	15.38
	100%	6.15
Trust in the model: "How much do you trust these statements?"	"Not at all"	3.08
	"Not much"	20.77
	"I am not sure"	22.31
	"Reasonably well"	43.08
	"A lot"	10.77

6 Numerical comparison between categories of responders' future projections and model numerical results

While the APE model generates a numerical estimation of future warming and wealth, the questionnaire provided the responders with choices among 5 discrete classes of relative departure from today's state ('Strong decrease', 'Decrease', 'More or less the same', 'Increase', 'Strong increase', see Questionnaire, Section 8). Such classes are purposely fuzzy, in order to simplify the responses (a numerical answer would require, among other challenges, a proper appreciation of the measurement units). As a result, there is some subjectivity in the way the five discrete classes are compared with the model numerical output. The approach we chose is as follows: i) we assigned numerical ranges to each fuzzy class, ii) each fuzzy class was assigned the median value between these ranges, iii) if numerical model and mental model prediction belong to the same fuzzy class we identified the prediction as correct, iv) if not, the difference was calculated as the distance between the numerical model outcome and the value assigned to the fuzzy class. This allows the errors to grow smoothly and continuously outside the fuzzy classes. Next, we searched for variables which could explain both the mental model predictions and their departure from the numerical models ones.

7 Responders' Comments

Immediately after the description of the core relations underlying the APE model, our survey offered the responders the opportunity to comment on the model ("Would you like to comment on these statements?"). Of the overall 130 responders, 58 provided a comment. These comments vary considerably, from agreement on the validity of the model ('*the statements seem reasonable and logical so i am inclined to believe them*'), to potential constructive solutions ('*clear that it is a behavioral change needed, people can have a high GDP but need to monitor their energy use, or energy should be renewable'*), to perplexity *('Not too sure if this is actually believable?'*), to scepticism about the model (*'too many variable not taken into account'*), to puzzling statements ('*the atmosphere needs* CO_2'), to clearly antagonistic ('*The people that design and work with this type of model are full of bullshit*'), to conspiracy-theory-like ('*This type of reasoning is typical of the distortion used by scientists, politicians and interest groups to force a 'moral argument' on us'*).

We divided these comments into 3 groups: 1) positive or constructive, 2) ambivalent, unsure or unclear and 3) negative or antagonistic. The groups differ statistically on a number of measures. As expected the largest difference is between group 1 ('positive or constructive') and 3 ('negative or antagonistic'). They differ statistically on proposed Mitigation Target (and thus on our measure of values), Climate Sensitivity, Critical Temperature and level of belief in climate change, Environmental Commitment, and, naturally, on trust on the APE model. Groups 1- 2 and groups 2-3 differ on fewer of these measures. None of the groups differ in terms of projected futures or mismatch between projections and model outcome, which confirms that ideology seems to affect model input and trust in the model, but not its dynamical evaluation or future projections.

References and Notes:

1. D. M. Kahan *et al.*, The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Clim. Change* **2**, 732 (2012).
2. A. Leiserowitz, Climate Change Risk Perception and Policy Preferences: The Role of Affect, Imagery, and Values. *Climatic Change* **77**, 45 (2006/07/01, 2006).
3. A. A. Leiserowitz, American Risk Perceptions: Is Climate Change Dangerous? *Risk Anal* **25**, 1433 (2005).
4. S. Dessai *et al.*, Defining and experiencing dangerous climate change. *Climatic Change* **64**, 11 (2004).
5. N. Adger, J. Barnett, K. Brown, N. Marshall, K. O'Brien, Cultural dimensions of climate change impacts and adaptation. *Nature Climate Change* **3**, 112 (02//print, 2013).
6. D. M. Kahan, D. Braman, P. Slovic, J. Gastil, G. L. Cohen, The Second National Risk and Culture Study: Making Sense of - and Making Progress In - The American Culture War of Fact. *GWU Legal Studies Research Paper* No. 370 (2007).
7. I. Feygina, J. T. Jost, R. E. Goldsmith, System Justification, the Denial of Global Warming, and the Possibility of "System-Sanctioned Change". *Pers Soc Psychol B* **36**, 326 (Mar, 2010).
8. Y. Heath, R. Gifford, Free-Market Ideology and Environmental Degradation. *Environment and Behavior* **38**, 48 (2006).
9. D. Kahan, Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making* **8**, 407 (2013).
10. Z. Kunda, The Case for Motivated Reasoning. *Psychological Bulletin* **108**, 480 (Nov, 1990).
11. E. Moxnes, A. K. Saysel, Misperceptions of global climate change: information policies. *Climatic Change* **93**, 15 (2009).
12. D. Dorner, *The Logic Of Failure: Recognizing And Avoiding Error In Complex Situations*. (Metropolitan Books, Ney York, 1996).
13. G. S. Halford, R. Baker, J. E. McCredden, J. D. Bain, How many variables can humans process? *Psychological Science* **16**, 70 (Jan, 2005).
14. L. B. Sweeney, J. D. Sterman, Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review* **16**, 249 (2000).
15. E. Moxnes, Overexploitation of renewable resources: the role of misperceptions. *Journal of Economic Behavior and Organization* **37**, 107 (1998).
16. E. Moxnes, Not only the tragedy of the commons: misperceptions of feedback and policies for sustainable development. *System Dynamics Review* **16**, 325 (2000).
17. M. A. Cronin, C. Gonzalez, J. D. Sterman, Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes* **108**, 116 (2009).
18. G. S. Halford, N. Cowan, G. Andrews, Separating cognitive capacity from knowledge: a new hypothesis. *Trends in Cognitive Sciences* **11**, 236 (2007).
19. J. D. Sterman, Risk Communication on Climate: Mental Models and Mass Balance. *Science* **322**, 532 (October 24, 2008, 2008).
20. F. Boschetti, C. Richert, I. Walker, J. Price, L. Dutra, Assessing attitudes and cognitive styles of stakeholders in environmental projects involving computer modelling. *Ecological Modelling* **247**, 98 (Dec, 2012).
21. D. Kahneman, *Thinking, fast and slow*. (Farrar, Straus and Giroux, New York, ed. 1st, 2011), pp. 499 p.
22. K. Stanovich, *Who is rational? studies of individual differences in reasoning* N. J. Mahwah, Ed., (Lawrence Erlbaum Associates, 1999).
23. F. Boschetti, A computational model of a mental model used to reason about climate change. *Environmental Science & Policy* **15**, 125 (Jan, 2012).
24. R. Costanza *et al.*, Time to leave GDP behind. *Nature* **505**, 283 (Jan 16, 2014).
25. M. R. Allen *et al.*, Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature* **458**, 1163 (Apr 30, 2009).
26. M. R. Raupach *et al.*, The relationship between peak warming and cumulative CO2 emissions, and its use to quantify vulnerabilities in the carbon–climate–human system. *Tellus B*, no (2011).
27. J. T. Jost, D. M. Amodio, Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion* **36**, 55 (2012).
28. J. T. Jost, J. Glaser, A. W. Kruglanski, F. J. Sulloway, Political conservatism as motivated social cognition. *Psychological bulletin* **129**, 339 (2003).
29. S. C. Gagnon Thompson, M. A. Barton, Ecocentric and anthropocentric attitudes toward the environment. *Journal of Environmental Psychology* **14**, 149 (1994).
30. J. Cohen *How Many People Can the Earth Support?* , (W. W. Norton & Company New York, 1996).
31. I. Ehrlich, F. Lui, The problem of population and growth: A review of the literature from Malthus to contemporary models of endogenous population and endogenous growth. *Journal of Economic Dynamics and Control* **21**, 205 (1997).
32. M. New, D. Liverman, H. Schroder, K. Anderson, Four degrees and beyond: the potential for a global temperature increase of four degrees and its implications. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **369**, 6 (January 13, 2011, 2011).
33. I. N. Harman, "Assessment of future global scenarios for the Garnaut Climate Change Review : an application of the GIAM framework" (CSIRO Marine and Atmospheric Research, Canberra, 2008).
34. J. Burniaux, ., et al., *The Economics of Climate Change Mitigation*. (OECD Publishing).

35. A. Bowen, G. Albertin, in *Special report on climate change: the low carbon transition*. (European Bank for Reconstruction and Development, London, 2011), pp. 20-37.
36. K. Gerardi, L. Goette, S. Meier, Numerical ability predicts mortgage default. *P Natl Acad Sci USA* **110**, 11267 (Jul 9, 2013).
37. R. Costanza, Social goals and the valuation of ecosystem services. *Ecosystems* **3**, 4 (Jan-Feb, 2000).
38. A. Tarantola, *Inverse Problem Theory*. (Elsevier, Amsterdam, 1987).
39. N. J. Grigg, F. Boschetti, M. Brede, J. J. Finnigan, A probabilistic approach to exploring low-dimensional global dynamics. *Procedia Environ Sci* **6**, 122 (2011).
40. M. R. Raupach *et al.*, Global and regional drivers of accelerating CO2 emissions. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10288 (Jun, 2007).
41. A. Strathman, F. Gleicher, D. S. Boninger, C. S. Edwards, The Consideration of Future Consequences - Weighing Immediate and Distant Outcomes of Behavior. *Journal of Personality and Social Psychology* **66**, 742 (Apr, 1994).
42. F. Pratto *et al.*, Social Dominance in Context and in Individuals: Contextual Moderation of Robust Effects of Social Dominance Orientation in 15 Languages and 20 Countries. *Social Psychological and Personality Science*, (February 19, 2013, 2013).
43. P. C. L. Heaven, D. S. Quintin, Personality factors predict racial prejudice. *Personality and Individual Differences* **34**, 625 (2003).
44. Y. Kashima *et al.*, Globalization and Folk Theory of Social Change: How Globalization Relates to Societal Perceptions about the Past and Future. *Journal of Social Issues* **67**, 696 (Dec, 2011).
45. Z. Leviston, I. A. Walker, “Baseline Survey of Australian Attitudes to Climate Change: Preliminary Report” (CSIRO Ecosystem Sciences, Perth, 2010).