

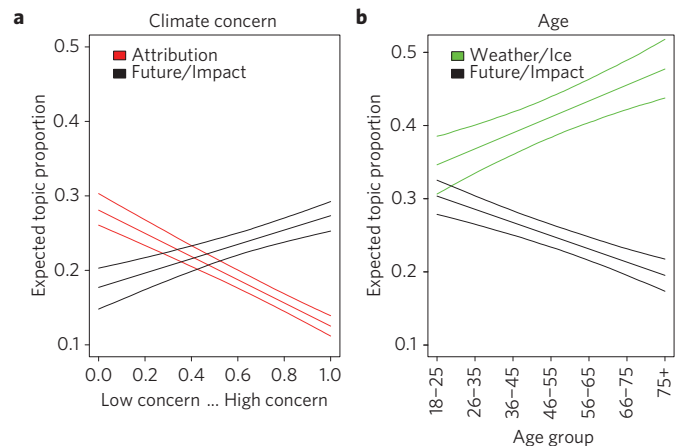
# Explaining topic prevalence in answers to open-ended survey questions about climate change

Endre Tvinnereim<sup>1\*</sup> and Kjersti Fløttum<sup>2</sup>

**Citizens' opinions are crucial for action on climate change, but are, owing to the complexity of the issue, diverse and potentially unformed<sup>1</sup>. We contribute to the understanding of public views on climate change and to knowledge needed by decision-makers by using a new approach to analyse answers to the open survey question 'what comes to mind when you hear the words 'climate change'?'. We apply automated text analysis, specifically structural topic modelling<sup>2</sup>, which induces distinct topics based on the relative frequencies of the words used in 2,115 responses. From these data, originating from the new, nationally representative Norwegian Citizen Panel, four distinct topics emerge: Weather/Ice, Future/Impact, Money/Consumption and Attribution. We find that Norwegians emphasize societal aspects of climate change more than do respondents in previous US and UK studies<sup>3-6</sup>. Furthermore, variables that explain variation in closed questions, such as gender and education, yield different and surprising results when employed to explain variation in what respondents emphasize. Finally, the sharp distinction between scepticism and acceptance of conventional climate science, often seen in previous studies, blurs in many textual responses as scepticism frequently turns into ambivalence.**

Numerous studies of public opinion about climate change show that agreement with the scientific consensus, concern about consequences and support for mitigation policies vary with age, gender, income and education<sup>7-12</sup>. However, fewer studies address differences in how climate change is interpreted and what associations are made by individuals. In this study, we make three contributions. First, we examine Norwegians' conceptions of what type of problem climate change is, and contrast this with previous studies of climate change imagery in the US<sup>3-5</sup> and UK<sup>6</sup>. Second, we test whether the structurally stable variables that have explained differences in degree on indicators such as concern or trust in science also explain differences in kind, that is, what type of association individuals choose when asked to write about climate change in their own words. Third, analysing the most representative answers of each topic, we often find emotional or affective expressions.

The overwhelming number of sub-topics that link to climate change makes it difficult to condense this issue into a few salient dimensions. The Intergovernmental Panel on Climate Change (IPCC) divides the area into three sub-fields: the physical science basis, impacts and mitigation<sup>13</sup>. One study<sup>1</sup> suggests six distinct frames: scientific uncertainty, national security, polar bears, money, catastrophe and justice/equity. Analysis of blogs shows how visions of negative impacts compete with more positive perspectives in the climate change debate<sup>14,15</sup>.

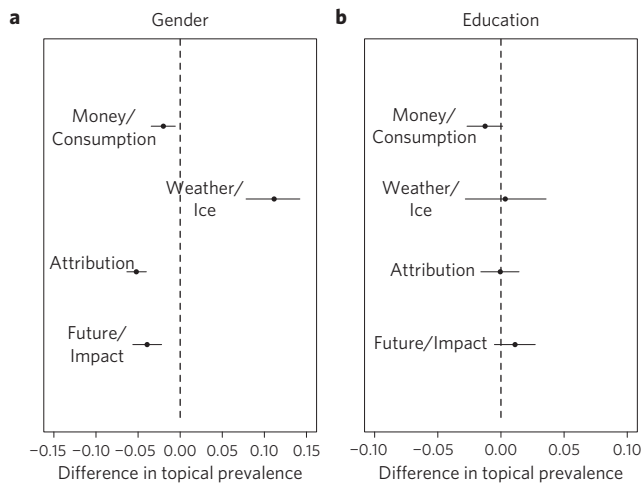


**Figure 1 | Effects of climate concern and age on topic prevalence.** Values have been generated from a regression where the outcome variable is the proportion of each document dedicated to each topic, given the selected STM model. Explanatory variables are climate concern, age, education (binary: university-level or not) and gender; those not shown on the horizontal axes are held at their sample medians. Confidence intervals indicate the 95% uncertainty range and include both regression uncertainty and measurement uncertainty from the STM model<sup>21,22</sup>.  $N=1,922$  for both models. See Supplementary Table 4 for full regression results.

Open-ended survey questions that permit respondents to use their own frame of reference, 'even if this might seem inappropriate or 'irrational' to the survey designer or analyst<sup>16</sup>, thus add great value to the study of public perceptions of climate change. US respondents emphasize ice melt, heat, 'alarmist' and 'naysayer' topics when asked to associate a word or phrase with 'global warming' in four studies from 2003 to 2010 (refs 3-5). Overall, physical images (ice melt, heat, nature, flood/sea level, weather) dominate, whereas 'naysayer' views increase over time. A study using open-ended questions to elicit reasons for supporting or opposing mitigation measures in two US states<sup>17</sup> finds four main categories of answers: economic, moral, political and technological. Men were more likely to bring up political rationales; women and young people more likely to bring up costs to self. Education, perhaps surprisingly, played no significant role predicting topic choice.

The main explanation for the relatively low number of studies of this kind has traditionally been cost, both to interviewers transcribing textual responses and to scholars analysing and categorizing the output. Recently, online survey methods and

<sup>1</sup>Uni Research Rokkan Centre, Nygårdsgaten 5, 5015 Bergen, Norway. <sup>2</sup>Department of Foreign Languages, University of Bergen, PO Box 7805, 5020 Bergen, Norway. \*e-mail: endre.tvinnereim@uni.no



**Figure 2 | Effect of gender and education on topic prevalence.** Values have been generated from a regression where the outcome variable is the proportion of each document dedicated to each topic, given the selected STM model, and with the same explanatory variables as in Fig. 1. **a,b**, Topics on the right of the zero line are more likely to be brought up by women (**a**) or those with university education (**b**). Confidence intervals (95%) include both regression uncertainty and measurement uncertainty from the STM model<sup>21,22</sup>.  $N = 1,922$  for both models. See Supplementary Table 4 for full regression results.

quantitative text analysis have brought those costs down, but to our knowledge this combination has not yet been exploited to shed light on climate change opinion. This study breaks new ground by including a greater number of responses, longer responses, a new country context (Norway) and crucially by employing automated techniques to induce a set of key topics based on mutual exclusivity and internal cohesion.

We aim to explore how diverse climate change discourses may influence and be reproduced by members of the public in their own words. Mental images of a phenomenon arguably precede cognition and thus serve as priors in decision-making, influencing how new information is processed<sup>4,5</sup>. Discourse creates, reproduces, challenges and excludes different representations of the world, thus forming the basis of decisions and actions. From this perspective, the present study permits us to uncover some of the fundamental constraints on and opportunities of climate action. Specifically, the degree to which citizens cast climate change as personal and immediate rather than distant may influence the extent to which policymakers perceive support for controversial mitigation and adaptation measures<sup>18</sup>.

The basic components of natural language—the words—typically have many meanings<sup>19</sup>. Through the textual answers, we see that ‘climate change’ is associated with many different phenomena, some related to physical reality and others to people’s subjective attitudes, their beliefs, values and interests<sup>20</sup>.

Among the textual responses, the median response length was four words and the mean length was 10.1 words (62.7 characters); the longest response had 310 words. The total corpus contained 21,470 words (110,247 characters not including spaces). Of these, the most frequent words were ‘extreme weather’ (one word in Norwegian, used 142 times), ‘weather’ (131), ‘warmer’ (94), ‘natural disaster(s)’ (78) and ‘human-made’ (77).

Through manual analysis of a range of alternative model specifications, we found that running a structural topic model with four topics yielded the most semantically coherent and distinct topics, compared with specifications with more or fewer topics (see Supplementary Methods). The selected model is shown in Table 1. We propose the following labels for the four topics.

**Topic 1: Money/Consumption.** Besides frequent references to the negative effects of consumer society, and the need to help poor countries, responses associated with this topic contain concepts related to policy, such as an international agreement and taxes, but also statements suspecting monetary motives behind climate policy. There is some overlap with Topic 3.

**Topic 2: Weather/Ice.** This topic emphasizes the physical manifestations of climate change, including predictions about the weather becoming more unstable, warmer and wetter and with more rain and landslides, melting ice and sea-level rise.

**Topic 3: Attribution.** These entries focus on what causes climate change, and notably the causal effect of human actions. Natural cycles are frequently mentioned, but often through expressions allowing for human causation at the same time.

**Topic 4: Future/Impact.** Negative consequences for humans are emphasized (danger, challenges), especially for future generations (children, grandchildren). The perspective represented here is one of the effects of climate change on humans in the future.

The most representative words for each topic—that is, words with high frequency in the given topic combined with low frequency in other topics—are given in Table 1. Four of the most representative answers of each topic, given the selected model, are shown in Table 2. Note, for example, that some of the Attribution statements express doubts about human causation of climate change.

The most prevalent topic is Weather/Ice, with close to half the document proportions in the aggregate; the Attribution and

**Table 1 | Most discriminating words by induced topic, with suggested topic labels.**

Topic	Most discriminating terms (frequency and exclusivity): Norwegian and English	Topic label	Topic proportion
1	<i>forbruk, peng*, hjelp, mest, mennesk, tror, tenk, litt, land, alt</i> consumption, money, help, most, human*, believe, think, little, land, everything	Money/Consumption	0.18
2	<i>vær, ekstremvær, smelt*, temperatur, naturkatastrof, global, uvær, flom, stig</i> weather, extreme weather, melt*, temperature, natural disaster, global, bad weather, flood, rise	Weather/Ice	0.45
3	<i>endr*, klima, syklus, jord, forhold, påvirk*, forurens*, all, menneskeskapt, endring</i> change, climate, cycle, earth, compare/condition, affect, pollut*, all/every, human-made, change	Attribution	0.15
4	<i>fremtid, alvor*, gjør, far*, verd, folk, ansv*, bil, barn, konsek</i> future, serious*, do, danger*, world, people, responsib*, car, child, consequence	Future/Impact	0.22

The terms shown are those that are the most frequent within as well as exclusive to each topic<sup>23</sup>. Labels for each topic are suggested qualitatively on the basis of the content of the terms and associated survey responses. Asterisks indicate stemming, for example, the term ‘responsib\*’ comprises ‘responsible’ and ‘responsibility’. Translations are best approximations based on readings of representative entries.

**Table 2 | Most representative responses by induced topic.****Topic 1: Money/Consumption**

[2] Bureaucrats who want to earn as much as possible from participating/being a part of expert staff who want to make money on alleged climate change. Bureaucracy. Researchers who think they know everything. Lots of people who are afraid of not making enough money if agencies are shut down. Egotistical climate bureaucrats. Money sink hole for public funds across the world.

[3] I hope and pray that the countries of the world will come to a comprehensive agreement. It is possible to do something about this problem, but rich countries have to stop being so greedy and actually help poor countries by investing in climate-friendly industry and CO<sub>2</sub>-reducing actions in poor countries as well as in their own. The oil (sovereign wealth fund) should pull out of everything dirty and environmentally unfriendly, and rather be invested in environmentally friendly technology, infrastructure, and so on, in other countries. Furthermore, the UN should set up a fund, into which all rich countries should pay an allowance of, for example, 1% of GDP, or what is mathematically needed, where the money goes to furthering environmental projects in developing countries. Subsequently, rich countries should be forced to spend another half of what they put into the fund, extra, to similar projects ... or overseas aid. And on top of that this emission reduction treaty should come, with this UN project, and additional developed-country commitments, which incentivize developing countries to sign the treaty. We, the rich countries, have created this problem, we must take most of the responsibility.

[4] That everything is changing. The climate, too. That it's a bit frightening. And a bit unfair. [I] think that because we and society are so focused on money and profit, to increase consumption and pleasure instead of sharing and thinking about future generations, we are only able to think a little about the global picture. And that catastrophes mostly fall on countries and people who have a lot to deal with already and little to protect themselves with. But who may be happier and more satisfied than us ...

[7] Rich countries have to do something to stop climate change and help poor countries with actions.

**Topic 2: Weather/Ice**

[1] More rain, storms and bad weather, higher sea level, extreme drought, extreme cold, glaciers melting.

[2] Ocean rising, temperature in ocean and air increasing, more unstable weather, floods, droughts.

[3] More bad weather such as precipitation, storms, etc. as well as melting of the poles.

[4] More extreme weather. Warmer or colder ... , floods, bad weather etc.

**Topic 3: Attribution**

[1] The climate has been changing several times in the course of the past billions of years. These changes are still going on and can barely be changed by humans, or not at all. Besides, high levels of CO<sub>2</sub> emissions are good for vegetation. The greenest periods on Earth have been those with high levels of CO<sub>2</sub> in the air.

[3] That this is something we in no way may influence to any great extent. My opinion is that this to a greater extent is cycles on Earth that are governed by the solar system.

[6] Climate change: the sum of human-made and natural changes that affect the climate on the Earth's surface and its atmosphere.

[10] The climate has always been changing and it still is. Now it's changing a bit faster than before. Part of the reason is human-made.

**Topic 4: Future/Impact**

[2] That a lot needs to be done to reduce emissions in China, India and the rest of the world; a lot needs to be done to prevent deforestation in Africa, Asia and the rest of the world; we need not only words but action against climate change.

[4] Climate change is very serious for people around the whole globe. So far we haven't seen the most serious consequences in our country, which may be one of the reasons why we don't seem to be willing to do what's needed to reduce/stop this dangerous development. I'm retired and think a lot about the problems that our children and not least grandchildren and future generations will face. The world is a bit locked into its growth economy, and the 'governing generations' aren't able to or lack the courage to take the unpopular and partly dramatic measures that are needed to reduce the 'acceleration of the wheels'. That would entail a long transition period to truly ecologically 'sustainable' production, perhaps with high levels of unemployment and necessary lowering of living standards in our rich part of the world. Therefore, we should give the young and engaged, who 'own the future', a much more important and decisive voice, since they, in any event, will be the ones who have to live with the consequences.

[5] That people no longer take it seriously because nothing is being done. It's a bit like crying wolf. Politicians are doing little because voters give very low priority to climate change, which in turn may be because little is being done. It's a vicious circle. Furthermore, Norwegians in general seem pretty egocentric on behalf of their nation. They are tired of hearing about problems that affect the rest of the world. I think this is a pity.

[7] Our grandchildren will condemn our generation for not doing what was necessary.

Selected responses from the ten most representative answers by topic, based on a qualitative assessment of representativity. The model-based representativity rank of each response is given in square brackets. The full list of top 10 responses is given in the original Norwegian in Supplementary Table 1.

Money/Consumption topics are smallest with less than one-fifth each. The categories are not directly comparable to those used in the US and UK studies, but the results are similar across the three countries in that ice melt, flooding and the weather appear prominently. However, attribution emphasis in Norway (15%) is clearly smaller than the US naysayer category (23% in 2010). Furthermore, human and societal aspects are together much more prevalent than in the US and UK studies. For example, Impacts/Future and Money/Consumption have a total topic proportion of 40% in Norway, against less than 15% for the US Alarmist category and 5% for the Disaster frame in the UK<sup>6</sup>. This suggests that Norwegians see climate change differently from their US and UK counterparts; it could also reflect a general global increase over time in attention to human and societal aspects of climate change.

Is the greater emphasis on human and societal aspects driven by any particular segments of society? We know that structurally stable variables such as age, gender and education have previously been used to explain differences in degree of concern and trust in science: do they also explain differences in the kinds of association made by citizens?

Figure 1 shows that respondents who in a closed question expressed more concern about climate change also devote significantly greater shares of their responses to the Future/Impact topic, all else equal. Conversely, those less concerned are more likely to bring up the Attribution topic. These positive relationships increase confidence in our structural topic modelling (STM) method.

We also find that older respondents are more likely to express their thoughts about climate change in terms of weather and

ice melt. Younger respondents, in contrast, emphasize human or societal effects through concern about the future, which may be explained either as a consequence of self-interest (young people are more likely to experience adverse effects from climate change) or cohort effect (young people have learnt about climate change in a period where effects on society have been emphasized). A strong cohort effect would imply increasing public attention to climate change over time, and possibly growing support for mitigation and adaptation measures. As Fig. 2 shows, women are significantly less likely than men to use the Attribution frame and more likely to respond in terms of weather and ice, all else equal. Among the 100 respondents with the highest Weather/Ice topic proportions, 67 were women; among the 100 with the highest Attribution proportions, 88 were men. This agrees with earlier findings that women express more scientifically accurate climate change knowledge and less climate scepticism than men<sup>7,9,12</sup>. Nevertheless, the result contrasts with studies using closed questions, where women tend to be more concerned than men<sup>12</sup>. In our study, this elevated concern, confirmed by the closed questions of our own poll, does not translate into a higher likelihood of using words related to future impacts.

Education has more muted effects. There are tendencies among individuals with university education to choose responses that fit the Future/Impact frame rather than the Money/Consumption frame, but only the effect on the latter topic is borderline significant. This is surprising given that education usually predicts significant variation in opinion<sup>7,8,10,12</sup>, but agrees with a study using open-ended survey questions<sup>17</sup>. Thus, although educational effects clearly exist in closed questions on climate change, education has less of an effect in determining what people think of as important in the context of climate change.

To validate our regression analyses shown in Figs 1 and 2, we undertake close readings of the 100 most representative statements of Topics 2 and 4. Specifically, among the 100 responses with the highest Attribution topic prevalence, only seven had in an earlier, closed survey question agreed with a statement that climate change is to a large extent due to human action. A reading of these responses shows that 29% nevertheless expressed ambivalent statements allowing for a human role in causing climate change. Items 6 and 10 under Attribution in Table 2 are representative of such statements. Thus, the written answers of respondents are much more nuanced than can be derived from closed survey questions. A Norwegian 'naysayer' category would need to exclude the ambivalent respondents, and thus constitute only about 10% of respondents or half the number seen in the US<sup>5</sup>.

Furthermore, an in-depth reading of the 100 most representative Topic 4 (Future/Impact) responses shows that concern is made explicit through emotional wordings such as 'afraid' and 'frightened' (19 occurrences in the top 100). The term 'serious' appears 14 times among the top 100, underlining the personal relevance of climate change to a substantial share of Norwegians.

For future closed questions, this study implies that greater insights may be garnered from allowing for expressions of ambiguity or uncertainty about the relative contributions of natural and anthropogenic drivers of climate change. Another option is to ask what type of issue respondents think climate change is, allowing respondents to choose between sub-topics identified by this study and others.

In more general terms, the present Norwegian study contributes to the relatively modest set of national studies already undertaken in countries such as the US and UK on what people associate with climate change, using their own words. Given the global character of climate change and the recurrent efforts to institute cross-national mitigation and other measures, we believe that comparative research of this kind constitutes important input to decision-making and to the understanding of public opinion formation and risk perception<sup>6</sup>.

## Methods

Methods and any associated references are available in the online version of the paper.

Received 4 December 2014; accepted 27 April 2015;

published online 1 June 2015

## References

- Hulme, M. *Why We Disagree about Climate Change: Understanding Controversy, Inaction and Opportunity* (Cambridge Univ. Press, 2009).
- Roberts, M. E. *et al.* Structural topic models for open-ended survey responses. *Am. J. Political Sci.* **58**, 1064–1082 (2014).
- Leiserowitz, A. Climate change risk perception and policy preferences: the role of affect, imagery, and values. *Climatic Change* **77**, 45–72 (2006).
- Leiserowitz, A. A. American risk perceptions: Is climate change dangerous? *Risk Anal.* **25**, 1433–1442 (2005).
- Smith, N. & Leiserowitz, A. The rise of global warming skepticism: Exploring affective image associations in the United States over time. *Risk Anal.* **32**, 1021–1032 (2012).
- Lorenzoni, I., Leiserowitz, A., de Franca Doria, M., Poortinga, W. & Pidgeon, N. F. Cross-national comparisons of image associations with "global warming" and "climate change" among laypeople in the United States of America and Great Britain. *J. Risk Res.* **9**, 265–281 (2006).
- Austgulen, M. H. & Stø, E. Norsk skepsis og usikkerhet om klimaendringer. *Tidsskr. Samfunnsforskning* **54**, 124–150 (2013).
- Dunlap, R. E. & McCright, A. M. A widening gap: Republican and Democratic views on climate change. *Environ. Sci. Policy Sustainable Dev.* **50**, 26–35 (2008).
- Engels, A., Hüther, O., Schäfer, M. & Held, H. Public climate-change skepticism, energy preferences and political participation. *Glob. Environ. Change* **23**, 1018–1027 (2013).
- Scruggs, L. & Benegal, S. Declining public concern about climate change: Can we blame the great recession? *Glob. Environ. Change* **22**, 505–515 (2012).
- Tingley, D. & Tomz, M. Conditional cooperation and climate change. *Comp. Polit. Stud.* **47**, 344–368 (2014).
- McCright, A. M. The effects of gender on climate change knowledge and concern in the American public. *Population Environ.* **32**, 66–87 (2010).
- Watson, R. T. & Albritton, D. L. *Climate Change 2001: Synthesis Report* (IPCC, Cambridge Univ. Press, 2001).
- Fløttum, K., Gjesdal, A. M., Gjerstad, Ø., Koteyko, N. & Salway, A. Representations of the future in English language blogs on climate change. *Glob. Environ. Change* **29**, 213–222 (2014).
- Morton, T. A., Rabinovich, A., Marshall, D. & Bretschneider, P. The future that may (or may not) come: How framing changes responses to uncertainty in climate change communications. *Glob. Environ. Change* **21**, 103–109 (2011).
- Stoneman, P., Sturgis, P. & Allum, N. Exploring public discourses about emerging technologies through statistical clustering of open-ended survey questions. *Public Underst. Sci.* **22**, 850–868 (2013).
- Shwom, R., Bidwell, D., Dan, A. & Dietz, T. Understanding US public support for domestic climate change policies. *Glob. Environ. Change* **20**, 472–482 (2010).
- Sunstein, C. R. On the divergent American reactions to terrorism and climate change. *Columbia Law Rev.* **107**, 503–557 (2007).
- Lyons, J. *Semantics* Vol. 1, 2 (Cambridge Univ. Press, 1977).
- Kerbrat-Orecchioni, C. *L'Énonciation: de la Subjectivité dans le Langage* (Armand Colin, 2009).
- Roberts, M. E., Stewart, B. M. & Tingley, D. *stm: R Package for Structural Topic Models* (2014); <http://structuraltopicmodel.com>
- Roberts, M. E. *et al.* Online appendix to structural topic models for open-ended survey responses. *Am. J. Political Sci.* **58**, 1064–1082 (2014).
- Bischof, J. & Airoidi, E. M. *Proc. 29th International Conference on Machine Learning (ICML-12)* 201–208 (2012).

## Acknowledgements

The Research Council of Norway provided financial support for part of the work under the Lingclim project (grant no. 220654). We are very grateful to B. M. Stewart for support on technical issues related to our structural topic models. We also thank participants at the Norwegian Citizen Panel conference, Bergen, 6–7 November 2014.

## Author contributions

E.T. conceived the study and performed the quantitative analysis. E.T. and K.F. performed the qualitative analysis and structured and wrote the text.

## Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to E.T.

## Competing financial interests

The authors declare no competing financial interests.

## Methods

STM generates mixed membership results where each textual response is assigned proportions of each topic. Choices such as model specification, number of topics and interpretation of topic cohesion and exclusivity are made qualitatively by the authors.

The topic proportions are based on word frequencies. Topical prevalence is a vector that sums up to one for each individual text or response: for example, in a three-topic model, one response may be deemed by the model to belong 70% to Topic 1 and 15% to each of Topics 2 and 3. Similarly, aggregate topic prevalence is based on overall frequencies of words associated differentially with each topic. Earlier studies using open-ended questions in the context of climate change<sup>3–5</sup> register counts of one or more topics per respondent, making aggregate topic proportions exceed 100%; these proportions are thus not directly comparable with the current study.

Responses with a high prevalence of words strongly linked to one topic are said to be representative of that topic, and close reading of representative responses is key to evaluating a model's quality. When choosing the most representative responses for close reading, we selected the 100 responses with the highest values on that topic's prevalence vector, as generated by the model.

As STM is a multimodal estimator, there is a risk that initial model conditions produce unrepresentative results. To counter this problem, we initiated over 500 potential models from randomly generated starting values, including different numbers of topics, automatically selecting the 20% with the highest expectation-maximization values for full convergence runs using the selectModel function<sup>21</sup>. The resulting 100+ different model runs were then analysed qualitatively to arrive at our preferred model shown in the paper. The qualitative analysis was based on the authors' readings of both the most discriminating terms by topic (Table 1) and most representative responses by topic (Table 2).

Our selected model contains four topics as qualitative analysis showed this number to yield the greatest semantic coherence within topics in combination with exclusivity between topics. This model reflects the general tendencies seen in the more than 100 model runs mentioned above. Specifically, most models tend to distinguish between the overall topics of impact/future/consequence, attribution/scepticism and weather/ice/sea-level rise. Supplementary Tables 1–3 show sample results of alternative model specifications containing five, six and eight topics. When greater numbers of topics were requested, a clear tendency was that ice melt and sea-level rise separated from weather events. Ice remained separate and coherent, whereas the other topics increasingly mixed at five or more topics.

Data were collected as part of the online Norwegian Citizen Panel<sup>24</sup>, based on postal recruitment of 25,000 individuals running from 6 November 2013 until 5 January 2014. Gender, age and education biases in the response rate were low<sup>25</sup>. The first wave garnered 4,905 survey subjects, amounting to a response rate of 20.1%. Of these, a randomly drawn sub-sample of 2,297 responses was used in the current study. The open-ended question had the wording: 'What comes to mind when you hear the words 'climate change'?' This produced responses from 2,115 individuals or 92%. The question appeared at the end of the study to minimize loss of respondents.

Seven other questions related to climate change were posed early in the study, in part to correlate with the open-ended answers. These questions asked about opinions on oil production in Norway, personal concern about and causes of

climate change, threat assessment, ease/difficulty of mitigation and moral obligation to reduce emissions. To minimize context effects<sup>26</sup>, the climate questions were asked early in the survey, with 41 questions (82 if counting individual battery items) about immigration, domestic terrorism, individual work situation and other demographic matters serving as a buffer. This strategy satisfies recommendations of a distance of at least six irrelevant items between related questions<sup>27</sup>. The fact that the questions are formally different (closed versus open), have dissimilar wordings and measure different constructs<sup>27,28</sup> also serves to avoid contamination effects. Close reading of the open answers finds no evidence of context effects, for example, as 'oil' (appearing early in the survey) is mentioned by only 26 respondents (1.2%) in the open answers. In contrast, 'weather' and 'ice', mentioned by 449 and 468 respondents, respectively (over 20%), were not brought up by earlier survey questions.

Concern was measured with the question 'How concerned are you about climate change?' with a five-point answer scale ranging from 'Not at all concerned' to 'Very concerned'. Views on climate science were measured with an instrument used in previous studies<sup>29</sup> asking respondents whether climate change was human-induced, natural, or not happening, with 'don't know' a fourth option. The textual responses were stemmed and stop words and punctuation removed using the SnowballC package. Subsequently, 371 terms remained. For the most frequent terms, varieties of the written standards nynorsk and bokmål were harmonized.

Figures 1 and 2 were generated on the basis of linear regression models with topic proportions for each of the four topics as dependent variables and age, gender, education and concern about climate change as explanatory variables. Detailed regression results for each model are given in Supplementary Table 4. Significance testing was performed using T-statistics based on standard errors that incorporate both estimation uncertainty from the topic induction process<sup>21,22</sup> and estimation uncertainty from the regression models. The figures thus show the most conservative uncertainty estimates available.

**Data deposition.** A replication data set with R code has been deposited at the Harvard Dataverse Network: E.T.; K.F., 2015, 'Replication data for: Explaining topic prevalence in answers to open-ended survey questions about climate change', <http://dx.doi.org/10.7910/DVN/28689> Harvard Dataverse Network [Distributor] V2 [Version].

The full data set is available from the Norwegian Social Science Data Services (NSD): [http://www.nsd.uib.no/nsddata/serier/norsk\\_medborgerpanel.html](http://www.nsd.uib.no/nsddata/serier/norsk_medborgerpanel.html)

## References

- Ivarsflaten, E. *et al.* *Endre Tvinneheim and Jacob Aars. Norsk Medborgerpanel [Norwegian Citizen Panel] Round 1: Data file* (Univ. Bergen, 2014).
- Hogestøl, A. & Skjervheim, Ø. *Norwegian Citizen Panel 2013: First Wave Methodology Report* (Univ. Bergen, 2014).
- Tourangeau, R. & Bradburn, N. *Handbook of Survey Research* (Emerald, 2000).
- Weijters, B., Geuens, M. & Schillewaert, N. The proximity effect: The role of inter-item distance on reverse-item bias. *Int. J. Res. Mark.* **26**, 2–12 (2009).
- Tourangeau, R., Rasinski, K. A., Bradburn, N. & D'Andrade, R. Belief accessibility and context effects in attitude measurement. *J. Exp. Soc. Psychol.* **25**, 401–421 (1989).
- Leviston, Z., Walker, I. & Morwinski, S. Your opinion on climate change might not be as common as you think. *Nature Clim. Change* **3**, 334–337 (2012).