Imprecise probability assessment of tipping points in the climate system

Elmar Kriegler[a,b,1], Jim W. Hall[c,d], Hermann Held[a], Richard Dawson[c,d], and Hans Joachim Schellnhuber[a,e,f]

[a]Potsdam Institute for Climate Impact Research, PO Box 60 12 03, 14412 Potsdam, Germany; [b]Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213-3890; [c]School of Civil Engineering and Geosciences, Newcastle University, Newcastle-upon-Tyne NE1 7RU, United Kingdom; [d]Tyndall Centre for Climate Change Research, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE1 7RU, United Kingdom; [e]Environmental Change Institute, Oxford University, Oxford OX1 3QY, United Kingdom; and [f]Tyndall Centre for Climate Change Research, University of East Anglia, Norwich NR4 7TJ, United Kingdom

Edited by William C. Clark, Harvard University, Cambridge, MA, and approved February 2, 2009 (received for review September 16, 2008)

Major restructuring of the Atlantic meridional overturning circulation, the Greenland and West Antarctic ice sheets, the Amazon rainforest and ENSO, are a source of concern for climate policy. We have elicited subjective probability intervals for the occurrence of such major changes under global warming from 43 scientists. Although the expert estimates highlight large uncertainty, they allocate significant probability to some of the events listed above. We deduce conservative lower bounds for the probability of triggering at least 1 of those events of 0.16 for medium (2–4 °C), and 0.56 for high global mean temperature change (above 4 °C) relative to year 2000 levels.

climate change | expert elicitation

The potential of large-scale changes in the earth system as a result of anthropogenic climate change has received increasing attention (*cf.* refs. 1 and 2), fuelled by observations of, among other things, accelerated ice discharge from Greenland and West Antarctica (see ref. 3 for an overview). Nonetheless, the assessment of the likelihood of such changes under global warming has largely defied quantification due to insufficient data, and a limited ability to model the underlying processes (2). As a consequence, it is difficult to account properly for the possibility of major changes in the earth system in climate policy assessments, although the potentially large socioeconomic impacts of such events are a source of concern (4, 5).

To overcome this unsatisfactory situation, we have elicited beliefs about major changes in the Atlantic meridional overturning circulation (AMOC), the Greenland ice sheet (GIS), the West Antarctic ice sheet (WAIS), the Amazon rainforest and the El Niño/Southern Oscillation (ENSO) from experts in the field. These systems have been cited as candidates for harbouring large-scale discontinuities, or "tipping points," where a small change in a driver, such as global mean temperature (GMT) can result in a disproportionate response of the system (2, 6). Building upon Lenton et al.'s (6) broad review of a range of potential tipping points, here we seek to quantify beliefs about critical transitions in the 5 components of the climate system listed above. Each event of "crossing a potential tipping point" was precisely defined in terms of the final state of the transition process (Table S1). Our aim was to produce policy-relevant information in terms of a set of subjective probabilities for "triggering" those transition processes under different scenarios of future GMT increase. In this context, "triggering (the crossing of) a tipping point" denotes the event of initiating the transition, or making its future initiation inevitable.

In the Bayesian paradigm, subjective probabilities constitute a measure of degree of belief as reflected in an individual's disposition to act (as opposed to the frequentist paradigm in which probabilities are thought of as limiting frequencies) (7, 8). The use of subjective probabilities is closely linked to decision analysis, which tries to identify best courses of action based on quantified preferences and beliefs together with a set of normative criteria for rational decision making (*cf.* ref. 9). To better inform decision making processes, formal elicitation protocols have been developed for assessing subjective probabilities of experts in the field (see ref. 10 for an overview). Such protocols established procedures to avoid common biases in the assessment of probabilities, drawing on a large body of literature on heuristics and framing effects in decision making under uncertainty (11). A common criticism is that expert elicitations do not add to the body of scientific knowledge unless verified by data or theory. In the context of risk analysis and decision making, however, expert elicitations have proved to be a unique tool for systematically gathering and projecting scientific information in complex policy problems (12, 13). It is increasingly recognized that they can play a valuable role for informing climate policy decisions (14). Formal elicitations have already been conducted in various areas of climate science (*cf.* refs. 15–17).

Subjective probabilities used in the context of normative decision theories can be interpreted as betting rates in a risk-neutral (linear utility) environment (18). For eliciting such probabilities, proper scoring rules have been proposed that reward the specification of the probability value that reflects the expert's "true" belief (*cf.* ref. 19). In practice, however, it is more common to assess probability values directly by measuring strength of belief in reference to well-defined frequencies. Owing to the specific challenges of judging the prospect of tipping points in the climate system, we have admitted imprecise probability assessments in this study.

The current knowledge base about tipping points is poor, with very limited data and process understanding that would allow experts to update their beliefs (2). Imprecise probability theory (20) offers a rigorous framework to capture potentially ambiguous beliefs. Such beliefs are described by an interval of subjective probabilities whose bounds can be interpreted as lower and upper betting rates in the context of generalized normative decision theories (*cf.* ref. 21) (see Methods and *SI Appendix* 1, Section 2). From a practical point of view, probability assessments aim to elicit some probability $\in [0,1]$ that best characterizes the expert's belief, whereas imprecise probability assessments seek to exclude those probabilities $\in [0,1]$ that would be incommensurate with the expert's belief. It is the general philosophy of this study to present a conservative assessment of the information available from experts.

The expert elicitation was conducted between October 2005 and April 2006 with a computer-based interactive questionnaire completed individually by participants. A total of 52 experts participated in the elicitation (see Table S2 for names and affiliations). The questionnaire included 7 events of crossing a tipping point. Five of them are discussed here whereas the remaining 2, concerned with

Author contributions: E.K., J.W.H., and H.-J.S. designed research; E.K., J.W.H., and R.D. performed research; E.K., J.W.H., and H.H. contributed new reagents/analytic tools; E.K., J.W.H., and R.D. analyzed data; and E.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: kriegler@pik-potsdam.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0809117106/DCSupplemental.

SUSTAINABILITY SCIENCE

Table 1. Excerpts from definition of the events of crossing a tipping point, and distribution of responses from experts

Reorganisation of the Atlantic Meridional Overturning Circulation (CMOC) "that involves a permanent shutdown of convection in the Labrador Sea and a drastic reduction in deep water overflow across the Greenland-Scotland ridge by at least 80%."
(A) 16 (B) 4 [Reasons: remote; model results inconclusive] (C) 2 [local cooling overwhelmed by overall warming trend]

Melt of the Greenland ice sheet (MGIS) ". . . an alternative state that is largely ice-free."
(A) 13 (B) 1 [Reason: too far into the future] (C) 1 [fastest melt is 600 years, too slow to be dangerous]

Disintegration of the West Antarctic ice sheet (DAIS) ". . . in which West Antarctica becomes an archipelago . . . "
(A) 14 (B) 1 [Reason: uncertainty about time scale of disintegration; possibility of collapse due to glacial readjustments only]

Dieback of the Amazon rainforest (AMAZ) ". . . in which at least half of its current area is converted to raingreen forest, savannah or grassland. Besides climate change, a second driver . . . is land use change from human activity . . . factor out this driver by assuming . . . that not more than 20 percent of the current rainforest will be deforested by human activity in the long run."
(A) 10 (B) 1 [Reason: vegetation change inconclusive for assessing feedbacks on climate] (C) 3 [Global effects limited; CO_2 sink to source conversion not a dangerous feedback on the climate]

Shift to a more persistent El Niño regime (NINO) ". . . a shift of the ENSO mean state towards El Niño like conditions."[†]
(A) 10 (A*)[‡] 1 (B) 3 [Reasons: original definition remote; Model results, paleo record inconclusive] (C) 1 [impacts of El Niño superimposed on a warmer world uncertain]

Option A: event will lead to potentially dangerous climate change, and willing to answer questions about its probability; option B: elicitation of probabilities not appropriate; option C: event will not lead to potentially dangerous climate change. See Table S1 for more details of the expert response.
[†]Definition was changed during the final phase of the elicitation during which participants were allowed to revise their statements (see Table S1).
[‡]Expert specified probabilities only for the original definition of the event. The response is not included here.

a dieback of boreal forests and a decline of the ocean carbon sink, were judged by experts to be of more speculative nature and are discussed in *SI Appendix* 2 and Figs. S1 and S2. The questionnaire proceeded in 4 parts: (*i*) selection of tipping points, (*ii*) ranking of tipping points in terms of sensitivity to global warming and uncertainty about underlying physical mechanisms, (*iii*) elicitation of lower and upper probabilities for the event of triggering the (crossing of) selected tipping points, and (*iv*) assessment of interactions between tipping points. The results from the ranking exercise (part *ii*) are reported in ref. 6, whereas this article focuses on the elicited probability statements (parts *iii* and *iv*).

Participation in our study was voluntary, which may have introduced a self-selection bias toward experts with higher concerns about tipping points. The possibility of such a bias will have to be judged on the basis of the list of participating experts (Table S2). In addition, nonzero probabilities of triggering major changes in the climate system may have emerged simply because we confronted experts with those particular events. We believe that such an "availability bias" is mitigated by the use of imprecise probabilities, which allows experts to register concern in terms of a nonzero upper probability, while at the same time expressing doubt with a zero lower probability. This article will therefore focus on the lower probability estimates provided by the experts.

Results and Discussion

Overview of Expert Response. Table 1 shows a break down of the expert response (see Table S1 for more details). Participants were allowed to choose the subset of tipping points they wished to comment upon, but were encouraged to restrict themselves to their area of expertise. Participants were asked for a self-assessment of their expertise on the selected tipping points, ranging on a scale from "1: Active researcher" to "4: Leading expert." This identified subsets of "core experts," defined as those who gave their highest self-assessment for the particular tipping point in question. An exception was made for the 2 pairs of cryosphere (Greenland and West Antarctic ice sheet) and biosphere tipping points (Amazon rainforest and boreal forests). Core experts on tipping point A of a pair (A, B) were also assigned "core" status on B if self-assessments matched the combinations (4, 3), (4, 2), (3, 2), or (2, 1) for (A, B). Participants were free to refuse to specify probabilities of triggering a tipping point (options B and C, Table 1) and, in 21% of all responses, experts exercised this option. Nine of 52 experts declined to estimate probabilities at all. The elicitation process is described in *SI Appendix* 1.

Fig. 1 shows the elicited probability intervals for "triggering" the events in Table 1 conditional on 3 scenarios representing low, medium and high warming (numerical values listed in Table S3). These scenarios are specified in terms of corridors of GMT increase relative to the year 2000 (Fig. 1 upper row). The corridors run out to 2200 to allow for a long-time perspective that is particularly important for the assessment of major changes in the ocean and the cryosphere. The long time horizon distinguishes our study from other assessments of abrupt climate change focusing on the 21st century (e.g., ref. 17). In particular, it extends beyond the time frame of the IPCC Fourth Assessment Report's (AR4) conclusion that "abrupt climate changes .. are not considered likely to occur in the 21st century" (ref. 2, p 818). Nonetheless, the slow transition time scales of particularly the cryosphere can extend far beyond a policy relevant time horizon of at most 200 years, which requires us to focus on the triggering of the transition process rather than the reaching of its final state. Because the former event can be difficult to observe, the additional cognitive demand on the experts may add to ambiguity in beliefs.

The probability of an event *B* conditional on some environmental variable is typically assessed across a range of environmental conditions, in our case described by the temperature corridors C1, C2, C3. Fig. 1 shows the change in the probability of triggering a tipping point (CMOC, row 2; MGIS, row 3; DAIS, row 4; AMAZ, row 5; NINO, row 6) from low (*Left*) to high warming (*Right*). Each panel summarizes the probability intervals of the respondents, thus providing information about the spread of assessments across experts. The dependence of individual estimates on the amount of warming can be traced by focusing on the bins with identical expert label across panels. A trend toward higher probabilities with increasing warming is clearly visible for all tipping points. With the exception of expert A1, all individual lower and upper probability values are monotonically increasing with temperature corridor. Because a corridors specifies a range of GMT trajectories, the spread between lower and upper probability of tipping may incorporate not only the ambiguity in expert beliefs, but also the range of trajectories within the corridor.

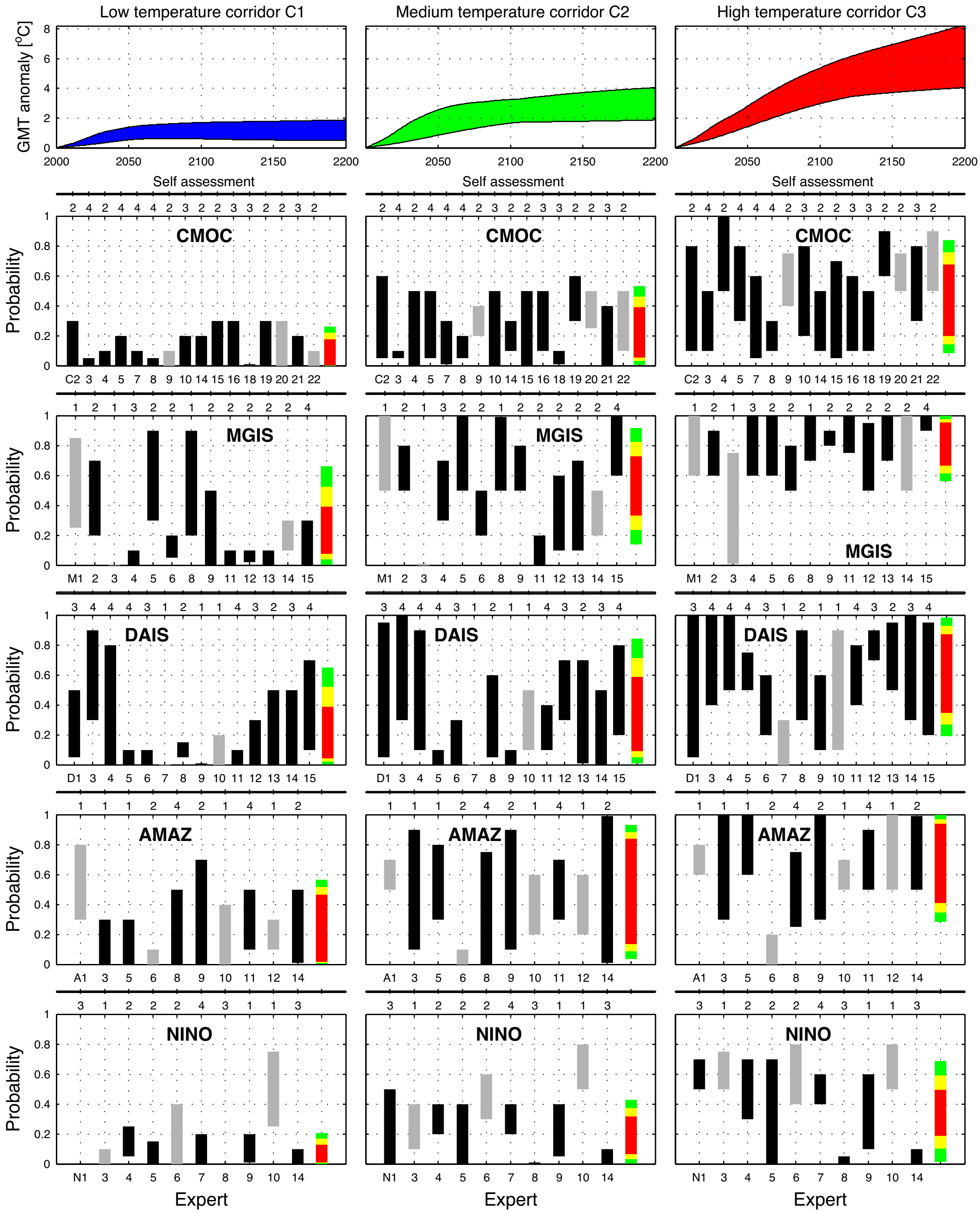

Fig. 1. Probability intervals from experts for the events CMOC, MGIS, DAIS, AMAZ, and NINO (see Table 1) conditional on 3 different corridors for future global mean temperature (GMT) increase to 2200 (relative to year 2000 temperatures, see top row). The presentation of expert opinions has been anonymized by numbering a random permutation of experts (shown below each panel). Labels are tipping point specific as indicated by the preceding letters C, M, D, A, and N. The self-assessment of experts is shown above each panel. Probability estimates of core experts (see text for an explanation) are depicted in black, and the remaining estimates are shown in gray. The rightmost bar in each panel shows the aggregation of probability intervals from core experts based on increasingly restrictive assumptions about expert weights: (*i*) weights are allowed to vary by ±100% (green) or ±50% (yellow) around uniform weights, and (*ii*) unweighted average of lower and upper bounds (red). The increasing strength of assumptions leads to nested probability intervals (Red < Yellow < Green). If bounds fall onto each other, the color of the outer interval is not seen.

Fig. 1 reveals that the experts' ambiguity about the probability of triggering a tipping point (as measured by the distance between their lower and upper probability assignments) is large. One-third of all estimates (38% of estimates from core experts) cover at least half the range of the unit probability interval, and several of them express near ignorance. In addition, expert intervals scatter widely. Nonetheless, there is a considerable amount of information contained in the expert assessments. The prospect of triggering a tipping point may be considered "remote" if the upper probability $P^*(B) < 0.1$. It may be labeled "significant" if the lower probability

$P_*(B) \geq 0.1$, and "large" if $P_*(B) \geq 0.5$. It can be seen that the majority of experts assessed the prospect of triggering a tipping point as "not remote" for all cases except for "CMOC, corridor C1" (41% of experts) and "NINO, corridor C1" (33% of experts). For all tipping points, for high climate change, the majority of experts regard the probability of triggering as "significant," and in the case of MGIS and AMAZ as "large." For medium climate change (corridor C2), the majority sees a "significant" probability of MGIS and AMAZ, and 50% of core experts judge the probability of MGIS to be "large." These results can be compared with a discussion of reasons for concern in the IPCC Third Assessment Report, which included large-scale climate system discontinuities (4). The qualitative assessment of increasing risk from such discontinuities for small to medium warming and high risk above 4 °C warming is broadly supported by our collection of expert estimates.

Some aggregation of the probability information in Fig. 1 is required for making it accessible to decision analysis. We have conducted a sensitivity analysis of pooling rules (Methods and *SI Appendix* 1, Section 3), and found weighted averages for lower and upper expert probabilities, respectively, a so-called linear opinion pool (22), the most satisfactory choice. The assessment of expert weights is typically attempted by cross- or self-evaluation of experts, or scoring past expert performance if available (13). Beyond the use of self-evaluation to identify core experts for each tipping point, none of this was practical in our application. Like others (ref. 20, chapter 5), we are skeptical of using a uniform weighting function to capture ignorance about the quality of expert statements. Therefore, we compare uniform weighting with 2 cases where the expert weights can vary by ±50% and ±100%, respectively, around uniform weights. Fig. 1 shows the probability intervals from linear pooling under those 3 assumptions about expert weights for all 5 tipping points and GMT corridors (rightmost bar in each panel). We restricted the aggregation to core experts. Obviously, the pooled probability intervals will widen with increasing imprecision in expert weights. We note that ambiguity in individual expert beliefs, and ambiguity in the weighting of expert estimates, represent 2 different layers of imprecision that contribute to the overall imprecision of the pooled probability estimates.

Reorganization of AMOC (CMOC). The probability intervals for CMOC are marked by comparatively low values for corridor C1, with a large increase in predominantly the upper probability bounds toward corridor C3. We associate the response pattern with the inconclusive nature of evidence from model intercomparison studies of a collapse of AMOC in the long run (2, 23, 24). A previous study (17) conducted detailed interviews with 12 scientists in the field (5 of whom participated in our survey) on the response of the AMOC to climate change. In qualitative agreement to our results, Zickfeld et al. (17) report probability estimates for the shutdown of AMOC (until 2100) in the range of 0–0.2 for low (<2 °C), 0–0.6 for medium (2–4 °C), and 0.05–0.95 for high climate change (4–8 °C) from those experts who considered a shutdown possible.

Shift of ENSO (NINO) and Dieback of the Amazon Rainforest (AMAZ). The expert response for NINO shows a similar pattern to CMOC. However, as GMT increases, experts tend to fall into 2 groups. The response of experts N8 and N14 were motivated by model intercomparison studies that showed no consistent trend in El Niño amplitude and frequency under climate change (25, 26). In contrast, other experts assume an increase in the probability of more persistent El Niño conditions in a warmer world, as found in a subset of models that best simulate the tropical Pacific climatology (27). The prospect of a dieback of the Amazon rainforest is closely linked to future changes in ENSO. Vegetation models driven with a strong drying of the Amazon basin have shown a dieback (28), but the magnitude of potential precipitation decrease over the Amazon remains controversial. Expert responses for AMAZ cluster above a probability of 0.5 for corridor C3 with the exception of expert A6, who believes that more persistent El Niño conditions in a warmer world are remote.

Melt of the GIS (MGIS). The expert response for MGIS differs markedly from NINO and CMOC. The upper bounds of some probability intervals already reach 0.9 for corridor C1. The lower bounds increase strongly toward larger warming. The exception is expert M3, who judged MGIS likely to be avoidable in the year 2200 regardless of the magnitude of warming (a similar response was given by expert D7 for DAIS). This points to an important controversy to what extent positive feedback such as (*i*) increased ablation due to changes of ice sheet topography and surface properties and (*ii*) rapid ice discharge due to lubrication of the ice sheet base (29) affect the stability of GIS. Dynamic deglaciation processes were discussed, but not incorporated in model-based inferences about GIS stability presented in IPCC AR4 (2). The expert response in our study might indicate a larger concern about such feedback, reinforced by data about rapidly increasing ice discharge from Greenland (30). However, this conclusion cannot be drawn unanimously from our elicitation, because the assessment of the likelihood of ice sheet decay will also depend on the extrapolation of the GMT corridors beyond 2200, which was left to the discretion of the experts. IPCC AR4 gives a threshold of 1.9–4.6 °C warming above preindustrial (≈1.3–4.0 °C above 2000, which intersects corridor C1 and C2 here) at which the surface mass balance of GIS becomes negative (ref. 2, p. 829). After the elicitation, more information about the extent of GIS during the last interglacial has become available (31, 32). This information might affect expert estimates for low climate change (corridor C1), but less so for the other 2 corridors describing temperature changes above the last interglacial.

Disintegration of WAIS (DAIS). The response pattern for DAIS is marked by large uncertainty among experts. This reflects the fact stated in IPCC AR4 that "no quantitative information is available from the current generation of ice sheet models as to the likelihood" of a disintegration of WAIS (ref. 2, p. 819). In the absence of model studies, a survey of expert opinions was conducted by ref. 16 revealing disagreement on the likely mechanism and time scale of a WAIS disintegration. In our study, experts mentioned specifically (*i*) the uncertain role of ongoing glacial readjustments and (*ii*) the lack of data on the size of WAIS and buttressing ice shelves in previous interglacials as contributing factors to the uncertainty. Point *ii* is closely related to a finding in IPCC AR4 identifying the role of surrounding ice shelves for the stability of WAIS as a major uncertainty (ref. 2, p. 817). At the time of the study of Vaughan and Spouge (16), recent evidence of ice shelf disintegration and subsequent acceleration of ice flows (3) was not available. This evidence may now be reflected in the high upper probability bounds for medium and high climate change. We note that experts express such concern despite the fact that quantitative model studies are lacking. This points to the strength of expert elicitations in providing a holistic picture of beliefs incorporating not only model results, but also insights from empirical data and theoretical considerations.

Interactions Between Tipping Points. The probability of triggering a tipping point may be increased or reduced depending upon whether or not a tipping point in another subsystem has already been crossed (6). For the tipping points selected under option A or B (Table 1), we asked participants whether knowing that another tipping point on the list had been triggered would (increase/decrease/increase or decrease/have no effect on) their estimate of the probability of triggering the particular tipping point in question at a later point in time. As depicted in Fig. 2, a majority of respondents identified an effect of some kind in 12 of 20 possible combinations of preceding and succeeding tipping point, highlighting the intricate web of interactions between these sensitive components of the earth system. As a matter of concern, the majority of experts anticipated

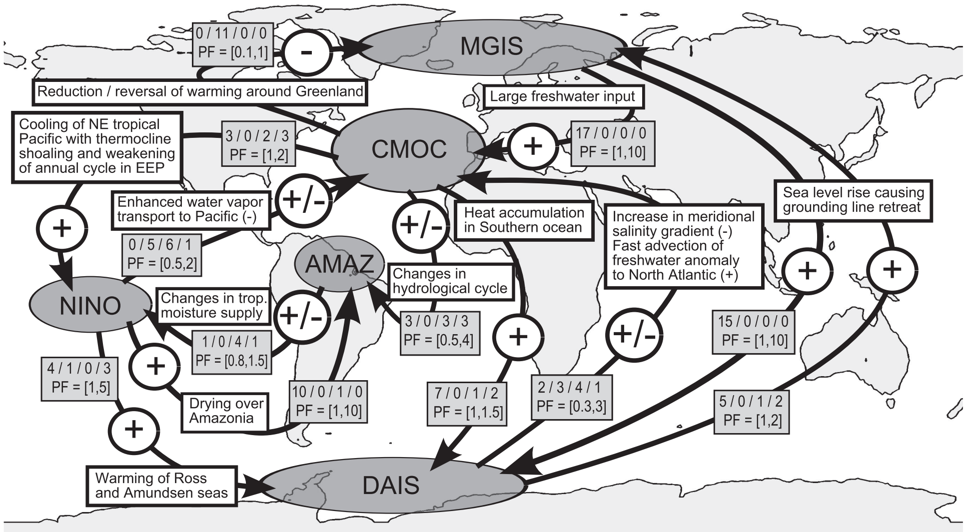

Fig. 2. Sketch of the main interactions between tipping points as described by participating experts. Pairs of tipping events *A*, *B* are connected with a directed arrow $A \rightarrow B$ if (*i*) at least 4 experts judged that some effect of triggering *A* on the probability of triggering *B* exists and (*ii*) they outnumbered the experts who saw no effect of *A* on *B*. Each arrow is accompanied by information about (*i*) whether the majority of experts identified an increase (+), decrease (−), or uncertainty in the direction (±) of change in the probability of triggering *B* after *A* occurred (white circles), (*ii*) the number of experts supporting an increase/decrease/uncertain direction/no effect (top line, gray boxes), (*iii*) the range of probability ratios *PF* by which the probability of triggering *B* is believed to change after *A* occurred (union of *PF* intervals from only those experts that provided estimates in accordance with the type of effect attributed to each arrow; bottom line, gray boxes), and (*iv*) the dominant physical mechanism(s) described by some of the experts (white rectangles). Verbal descriptions of mechanisms for NINO → DAIS and CMOC → AMAZ could be obtained only from 1–2 experts, and therefore remain particularly speculative. For CMOC → NINO, the described mechanism was identified only recently (35), and was hinted at, but not fully disclosed at the time of elicitation.

an aggravating effect in 7 of 12 cases. The depicted interaction between the ice sheets via sea level rise and associated grounding line retreat might have to be reevaluated due to a recent finding of grounding line stability (33) that was not available at the time of elicitation. However, the coupling between the 2 ice sheets via bipolar see-saw [mediated via AMOC (34)] is unaffected by this finding.

To quantify the interactions for those tipping points *B* for which the respondents had provided probability intervals, we asked them to specify lower and upper bounds on the factor *PF* by which their probability of triggering *B* would [increase/decrease/increase (upper bound) or decrease (lower bound)] if they learned that tipping point *A* had already been triggered (see Methods and *SI Appendix* 1, Section 4). Fig. 2 records the union of *PF* intervals from those experts who agreed with the majority on the particular direction of the effect. These intervals are conservative in as much as they always include the possibility that tipping point *A* has no effect on *B*, i.e., $PF = 1$.

Because all tipping points listed in Table 1 may inflict large damages on socioeconomic systems, and may also aggravate the probability of triggering yet another tipping point (Fig. 2), we deduced bounds on the overall probability of triggering at least 1 of the 5 tipping points (henceforth called event ONE) conditional on the 3 GMT corridors. Including the information about the probability ratios above (Fig. 2), we have calculated the probability intervals for ONE, for (*i*) a sample of possible combinations of probability estimates from core experts for different tipping points and (*ii*) the pooled probability intervals for triggering the individual tipping points (see *Methods* and *SI Appendix* 1, Section 4). In the latter case, assuming linear pooling with (*ii.a*) uniform weighting of expert estimates (red bars, Fig. 1) and (*ii.b*) ±100% imprecision in expert weights (green bars), we calculate lower bounds on the probability of triggering ONE of 0.10 (*ii.a*) and 0 (*ii.b*) for low (<2 °C, corridor C1), 0.36 (*ii.a*) and 0.16 (*ii.b*) for medium (2–4 °C, corridor C2), and 0.69 (*ii.a*) and 0.56 (*ii.b*) for high GMT change (>4 °C, corridor C3). The upper probability bounds are uninformative (=1 for all cases and corridors, except of 0.86 for case *ii.a* and corridor C1). As shown in Fig. S3, we find that the probability intervals for ONE derived in case (*ii.b*) encompass the probability intervals for at least 96% of expert combinations investigated in case (*i*) and therefore can be considered a very conservative estimate. These results are reported in the abstract.

Conclusion

Our results indicate that the large uncertainty among experts about the prospect of triggering major changes in the climate system (Fig. 1) does not necessarily imply that such events are considered to be remote. We have presented subjective probability estimates, elicited from 43 experts in the field, that do not support the notion of high consequence, low probability events. Even when allowing experts to express large ambiguity in their beliefs, and, using conservative assumptions for aggregating their probability estimates, we find significant lower probability bounds for triggering major changes in the climate system. In many cases, they are considerably higher than the probability allocated to catastrophic events in current climate damage assessments, e.g., in ref. 36, Table 4.9 (25% loss of gross world product with probability 0.012 for 2.5 °C and 0.068 for 6 °C warming) and ref. 37 (5–20% loss with probability 0.1 ($T - 5$ °C) for a warming $T \geq 5$ °C), although this comparison is limited by the fact that the economic losses from the major climate changes discussed here are uncertain.

It should be noted that expert beliefs are tied to the time of elicitation, and may be updated in light of new information. With regard to a possible self-selection bias mentioned in the introduction, we note that spread and imprecision in expert responses do not indicate an overly coherent cohort of experts, as reflected by the list of participants (Table S2). With regard to our hypothesis that the "availability bias" is mitigated by the use of imprecise probabilities we note that experts registered a zero lower probability for crossing a tipping point in 71% of cases for low, 33% for medium, and 10% for high climate change. Face to face interviews with carefully selected experts as used in ref. 17 remain the gold standard for expert elicitations (12), but become increasingly infeasible if large numbers of experts are to be included (38). It is reassuring that ref. 17 found a pattern of expert beliefs about a shutdown of AMOC that is in qualitative agreement to our results. Our questionnaire design with interactive consistency checks and a subsequent revision phase involving extensive E-mail communication mitigated differences in comprehension and commitment among participating experts.

Methods

Interpretation and Elicitation of Imprecise Probabilities. In the Bayesian tradition, the probability of some event B is identified with the certainty equivalent of a bet on B assuming linear utility of payoffs (18). Consider a bet that pays \$1 if B occurs. An expected utility maximizer holding probability $P(B)$ would buy the bet for a price $\$p < \$P(B)$, and sell the bet for a price $\$q > \$P(B)$. Thus, $P(B)$ constitutes the certainty equivalent or fair betting rate of the bet on B. In the presence of ambiguity about the probability $P_*(B) \leq P(B) \leq P^*(B)$, the individual will become more conservative and accept the bet only if the price is reduced below $\$p_* < \$P_*(B)$ (supremum buying price). Likewise, the individual may issue the bet only for a price above $\$q^* > \$P^*(B)$ (infimum selling price) (20). For $p_* = q^*$, the Bayesian case of a fair betting rate is recovered. For $p_* = 0$ and $q^* = 1$, the individual would not indulge in any type of betting on the event B, signaling a state of complete ignorance. It is important to note that the concept of imprecise probability is very different from the assumption of a second-order (meta)probability distribution $F(p_B)$: $[0,1] \rightarrow [0,1]$ on the probability $P(B)$. From an expected utility point of view, second order probabilities are an ill-defined concept. Because the fair betting rate on the event B would then be described by the expectation $\langle P\rangle(B) = \int_0^1 p_B dF(p_B)$, only the mean value of the distribution $F(p_B)$ is relevant for the betting decision. Therefore, a decision maker holding a second-order probability $F(p_B)$ on $P(B)$ is indistinguishable from a decision maker holding a first-order probability $\langle P\rangle(B)$. The presence of imprecision indicates ambiguity about the probability $P(B)$ that cannot be resolved by some second-order (meta)probability model.

The elicitation of lower and upper probabilities was designed to reduce typical biases of overconfidence and anchoring (11). Participants were asked (*i*) which of the 2 complementary events "Triggering" and "Not Triggering" they judge to be more probable, and (*ii*) which in a set of linguistic probability labels based on the IPCC categorization of uncertainty (39) they find incommensurate with their belief. They were then asked to (*iii*) provide conservative lower and upper bounds on the probability of tipping taking into account their previous judgments in (*i*) and (*ii*). The full information was subjected to interactive cross-checks for basic consistency of expert statements. Compiled results were presented to participants, who were then given the opportunity to revise their individual submissions. Further details on this method are provided in *SI Appendix* 1, Sections 1 and 2.

Aggregation of Expert Probabilities. Notwithstanding their theoretical limitations, we have explored 2 prominent axiom-based aggregation rules, the linear and logarithmic opinion pool (22), together with other pooling rules, in particular a proposal by Nau (40) based on the betting interpretation of lower and upper probabilities (Fig. S4). A detailed discussion of our implementation and comparison of various pooling rules is given in *SI Appendix* 1, Section 3. We found the linear opinion pool to be most robust against outliers in expert estimates (Fig. S5). For any rule, there remains the problem of specifying meaningful expert weights (36). We compared a standard (but not particularly defensible) assumption of uniform expert weights with weaker assumptions incorporating ambiguity about the expert weights. In the latter case, expert weights were adjusted within $\pm 50\%$ or $\pm 100\%$ of uniform weights so as to minimize (maximize) the pooled lower (upper) probability (*SI Appendix* 1, Section 3).

Derivation of the Joint Probability of Triggering ONE. The probability of triggering ONE is constrained by the marginal lower and upper probabilities of triggering the individual tipping points, and the probability ratio $PF = P(B|A$ before $B)/P(B)$ capturing the effect of preceding tipping point A on succeeding tipping point B. PF will in general depend on the magnitude of the marginal probability $P(B)$, and hence on the conditioning GMT corridor. Nevertheless, we only asked for a single generic interval capturing the range of PF to reduce the complexity of the question. The consequences of this assumption are explored in (Fig. S6). Because we asked for cause-effect relationships and not correlations to elicit the probability ratios, their inclusion in the calculation required a careful consideration of the sample space (*SI Appendix* 1, Section 4).

ACKNOWLEDGMENTS. We thank the experts who participated in this study for their cooperation and Till Kuhlbrodt, Tim Lenton, Wolfgang Lucht, Michael Oppenheimer, Carol Turley, Martin Visbeck, Robert Nau, Teddy Seidenfeld, and Granger Morgan for advice in designing and analyzing the questionnaire. E.K. was supported by Marie Curie International Fellowship MOIF-CT-2005-008758 (to E.K.) within the 6th European Community Framework Program, and his research infrastructure was partly provided by Climate Decision Making Center at Carnegie Mellon University through National Science Foundation Grant SES-0345798; H.H. was supported by Volkswagen Foundation Grant II/78470; and Implementation of the computer-based questionnaire by R.D. was funded by the Tyndall Centre for Climate Change Research.

1. Alley RB, et al. (2003) Abrupt climate change. *Science* 299:2005–2010.
2. Meehl GA, Stocker TF (2001) in *Climate Change 2007—The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the IPCC, eds Solomon S, et al. (Cambridge Univ Press, Cambridge, UK), pp 747–846.
3. Alley RB, Clark PU, Huybrechts P, Joughin I (2005) Ice-sheet and sea-level changes. *Science* 310:456–460.
4. Smith JP, Schellnhuber H-J, Mirza MQ (2001) in *Climate Change 2001: Impacts, Adaptation, and Vulnerability*, Third Assessment Report of the IPCC, eds McCarthy JJ, Canziani OF, Leary NA, Dokken DJ, White KS (Cambridge Univ Press, Cambridge), pp 913–967.
5. Schellnhuber HJ, Cramer W, Nakicenovic N, Wigley T, Yohe G, eds (2006) *Avoiding Dangerous Climate Change* (Cambridge Univ Press, Cambridge, UK).
6. Lenton TM, et al. (2008) Tipping elements in the earth's climate system. *Proc Natl Acad Sci USA* 105:1786–1793.
7. de Finetti B (1974) *Theory of Probability* (John Wiley & Sons, Chichester, UK).
8. Savage LJ (1954) *The Foundation of Statistics* (John Wiley & Sons, New York).
9. Raiffa H (1968) *Decision Analysis: Introductory Lectures on Choices Under Uncertainty* (Addison-Wesley, Reading, MA).
10. Garthwaite PH, Kadane JB, O'Hagan A (2005) Statistical methods for eliciting probability distributions. *J Amer Stat Ass* 100:680–701.
11. Kahneman D, Slovic P, Tversky A (1982) *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge Univ Press, Cambridge, UK).
12. Morgan MG, Henrion M (1990) *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge Univ Press, New York).
13. Cooke RM (1991) *Experts in Uncertainty* (Oxford Univ Press, Oxford).
14. Oppenheimer M, O'Neill BC, Webster M, Agrawala S (2007) The limits of consensus. *Science* 317:1505–1506.
15. Morgan MG, Keith DW (1995) Subjective judgments by climate experts. *Environ Sci Technol* 29:468–476.
16. Vaughan DG, Spouge JR (2002) Risk estimation of collapse of the West Antarctic ice sheet. *Clim Change* 53:65–91.
17. Zickfeld K, et al. (2007) Expert judgements on the response of the Atlantic meridional overturning circulation to climate change. *Clim Change* 82:235–265.
18. Ramsey FP (1931) In *The Foundations of Mathematics and Other Logical Essays*, ed Braithwaite RB (Routledge and Kegan Paul, London), pp 156.
19. Winkler RL (1967) The quantification of judgment: Some methodological suggestions. *J Amer Stat Ass* 62:1105–1120.
20. Walley P (1991) *Statistical Reasoning with Imprecise Probabilities* (Chapman and Hall, London).
21. Troffaes MCM (2007) Decision making under uncertainty using imprecise probabilities. *Int J Approx Reas* 45:17–29.
22. Clemen RT, Winkler RL (1999) Combining probability distributions from experts in risk analysis. *Risk Analysis* 19:187–203.
23. Gregory JM, et al. (2005) A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing CO_2 concentration. *Geophys Res Lett* 32:L12703.
24. Stouffer RJ, et al. (2006) Investigating the causes of the response of the thermohaline circulation to past and future climate changes. *J Clim* 19:1365–1387.
25. Collins M, CMIP Modelling Groups (2005) El Niño- or La Niña-like climate change? *Climate Dyn* 24:89–104.
26. van Oldenborgh GJ, Philip SY, Collins M (2005) El Niño in a changing climate: A multi-model study. *Ocean Science* 1:81–95.
27. Guilyardi E (2006) El Niño-mean state-seasonal cycle interactions in a multi-model ensemble, *Climate Dyn* 26:329–348.
28. Cox PM, Betts RA, Collins M, Harris PP, Huntingford C, Jones CD (2004) Amazonian forest dieback under climate-carbon cycle projections for the 21st century. *Theor Appl Clim* 78:137–156.
29. Zwally HJ, et al. (2006) Surface melt-induced acceleration of Greenland ice-sheet flow. *Science* 297:218–222.
30. Rignot E, Kanagaratnam P (2006) Changes in the velocity structure of the Greenland ice sheet. *Science* 311:986–990.
31. Otto-Bliesner BL, et al. (2006) Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* 311:1751–1753.
32. Willerslev E, et al. (2007) Ancient biomolecules from deep ice cores reveal a forested Southern Greenland. *Science* 317:111–114.
33. Alley RB, Anandakrishnan S, Dupont TK, Parizek RB, Pollard D (2007) Effect of sedimentation on ice-sheet grounding-line stability. *Science* 315:1838–1841.
34. Stocker TF (1998) The seesaw effect. *Science* 282:61–62.
35. Timmermann A, et al. (2007) The influence of a weakening of the Atlantic Meridional Overturning Circulation on ENSO. *J Climate* 20:4899–4919.
36. Nordhaus WD, Boyer J (2000) *Warming the World* (MIT Press, Cambridge, MA).
37. Stern N (2006) *The Economics of Climate Change* (Cambridge Univ Press, Cambridge, UK).
38. Keith DW (1996) When is it appropriate to combine expert judgments? *Clim Change* 33:139–143.
39. IPCC (2001) *Climate Change 2001: The Scientific Basis*, SPM Footnote 7 (Cambridge Univ Press, Cambridge, UK).
40. Nau RF (2002) The aggregation of imprecise probabilities. *J Stat Plan Inf* 105:265–282.