This the the authors' preprint (before reviewer comments, subsequent revisions, and typesetting) of an article with the same title and author line accepted for publication at *Climatic Change*.

Attention to values helps shape convergence research

Casey Helgeson[a], Robert E. Nicholas[a,b], Klaus Keller[c,a], Chris E. Forest[b,a], and Nancy Tuana[d]

[a]Earth and Environmental Systems Institute, Penn State, State College PA 16801
[b]Department of Meteorology and Atmospheric Science, Penn State, State College PA 16801
[c]Department of Geosciences, Penn State, State College PA 16801
[d]Department of Philosophy, Penn State, State College PA 16801

Abstract:
Convergence research is driven by specific and compelling problems and requires deep integration across disciplines. The potential of convergence research is widely recognized, but questions remain about how to design, facilitate, and assess such research. Here we analyze a seven-year, twelve-million-dollar convergence project on sustainable climate risk management to answer two questions. First, what is the impact of a project-level emphasis on the values that motivate and tie convergence research to the compelling problems? Second, how does participation in convergence projects shape the research of postdoctoral scholars who are still in the process of establishing themselves professionally? We use an interview-based approach to characterize what the project specifically enabled in each participant's research. We find that (a) the project pushed participants' research into better alignment with the motivating concept of convergence research and that this effect was stronger for postdoctoral scholars than for more senior faculty. (b) Postdocs' self-assessed understanding of key project themes, however, appears unconnected to metrics of project participation, raising questions about training and integration. Regarding values, (c) the project enabled heightened attention to values in the research of a large minority of participants. (d) Participants strongly believe in the importance of explicitly reflecting on values that motivate and pervade scientific research, but they question their own understanding of how to put value-focused science into practice. This mismatch of perceived importance with poor understanding highlights an unmet need in the practice of convergence science.

1 Introduction

Many pressing societal problems—such as pandemics, antibiotic resistance, global climate change, and sustainable development—span established academic disciplines. For problems like these, improving understanding of underlying trade-offs and providing effective decision support requires integrating expertise and insights across disciplines and stakeholders (National Research Council 2014; Institute of Medicine 2005; National Academies of Sciences, Engineering, and Medicine 2019). A number of partly-overlapping labels are used to describe such integrative research, including *interdisciplinary*, *transdisciplinary*, and *convergence* research (we will refer collectively to ITC research) (Huutoniemi et al. 2010; Institute of Medicine

2005; National Academies of Sciences, Engineering, and Medicine 2019; National Academies of Sciences, Engineering and Medicine 2021; National Research Council 2014; National Science Foundation n.d.). In light of its standing among the “10 Big Ideas” of the US National Science Foundation (NSF) (National Science Foundation n.d.), here we prioritize the concept of *convergence research*, characterized as research “driven by a specific and compelling problem” and involving “deep integration across disciplines” (National Science Foundation n.d.). Nevertheless, we situate our study with respect to the broader umbrella of ITC research, given the overlap between concepts. A series of National Academies reports synthesizes current understanding and best practices for facilitating ITC research (National Research Council 2014; Institute of Medicine 2005; National Academies of Sciences, Engineering, and Medicine 2019). One theme in these reports is a continuing need for more *meta-research* or *science of science* (Fortunato et al. 2018; Ioannidis 2018) to better understand the social and intellectual processes conducive to ITC research (National Research Council 2014; Institute of Medicine 2005; National Academies of Sciences, Engineering, and Medicine 2019).

A growing literature is responding to this need, using a variety of methods including bibliometric analyses (Porter et al. 2006; Anzai et al. 2012; Abramo, D’Angelo, and Zhang 2018; Porter et al. 2010; Kodama, Watatani, and Sengoku 2013), surveys of research participants (Tress, Tress, and Fry 2005; Cummings and Kiesler 2005; van Rijnsoever and Hessels 2011; Teirlinck and Spithoven 2015), interviews (Corley, Boardman, and Bozeman 2006; Lundershausen 2018; Polk 2014; Siedlok, Hibbert, and Sillince 2015; Tress, Tress, and Fry 2005; Wall, Meadow, and Horganic 2017), ethnographic observation (MacLeod and Nersessian 2014; Polk 2014; Siedlok, Hibbert, and Sillince 2015), document analysis of proposals, meeting minutes, or other sources (Gaziulusoy et al. 2016; Corley, Boardman, and Bozeman 2006; Cummings and Kiesler 2007; Siedlok, Hibbert, and Sillince 2015), analysis of data from funding agencies (Bromham, Dinnage, and Hua 2016; Cummings and Kiesler 2007), and reflections on personal experience (Freeth and Caniglia 2020; Gaziulusoy et al. 2016; König et al. 2013; Lang et al. 2012; McLeish and Strang 2016; West, van Kerkhoff, and Wagenaar 2019).

Such studies pursue a variety of aims, including measuring collaboration and interdisciplinarity (Porter et al. 2006, 2007; Porter, Roessner, and Heberger 2008; Anzai et al. 2012; Abramo, D’Angelo, and Di Costa 2012; Abramo, D’Angelo, and Zhang 2018; Sylvan Katz and Martin 1997), identifying barriers (Brister 2016; Gaziulusoy et al. 2016; Lang et al. 2012; MacLeod 2018), evaluating coordination mechanisms (Cummings and Kiesler 2005, 2007), measuring funding success (Bromham, Dinnage, and Hua 2016), linking outcomes to individual- or project-level characteristics (Polk 2014; Teirlinck and Spithoven 2015; Tress, Tress, and Fry 2005; van Rijnsoever and Hessels 2011), and developing frameworks for guiding ITC research (Corley, Boardman, and Bozeman 2006; Freeth and Caniglia 2020; König et al. 2013; Lang et al. 2012; Siedlok and Hibbert 2014; West, van Kerkhoff, and Wagenaar 2019) or for studying and evaluating ITC research (Bark, Kragt, and Robson 2016; Huutoniemi et al. 2010; Klein 2008; Kodama, Watatani, and Sengoku 2013; MacLeod and Nagatsu 2018; McLeish and Strang 2016; Wall, Meadow, and Horganic 2017). A majority of such studies also leverage their findings to posit generalized insights and recommendations, targeting all levels of organization from participants and project leaders through institutions and funding agencies.

This growing body of research has provided important insights, but many questions remain. Our study addresses two open questions. The first concerns the importance of values in ITC research. Disciplines and institutions have their own cultures, defined in part by shared values (Institute of Medicine 2005; Laursen, Gonnerman, and Crowley 2021). These values include both *ethical values* regarding the societal importance of research (Bessette et al. 2017; Diekmann and Peterson 2013; Vezér et al. 2018) and *epistemic values*—also called *knowledge* values (Rhoten 2003)—regarding the scientific importance of research questions and the features that make a piece of research respectable and rigorous (Vezér et al. 2018; Mayer et al. 2017; MacLeod 2018). Non-academic partners and other stakeholders bring their own sets of values (O'Brien and Wolf 2010; Polk 2014; Tschakert et al. 2017). Openness to other perspectives and values may be a common characteristic of successful ITC researchers (Institute of Medicine 2005; National Academies of Sciences, Engineering, and Medicine 2019), and divergence in values (both ethical and epistemic) can be a powerful obstacle to ITC research (Brister 2016; MacLeod 2018; MacLeod and Nagatsu 2018). While some ITC integration tools include a role for values (Robinson et al. 2016; Laursen, Gonnerman, and Crowley 2021), there is an absence of project-level studies investigating efforts to improve ITC research through strategies that foreground values.

A second question concerns what junior scholars, who are still in the process of establishing themselves professionally, learn through participation in ITC projects. Barriers and best practices for ITC research vary by career stage (Institute of Medicine 2005; National Research Council 2014). Postdoctoral scholars in particular can play an important role in collaborations and connections within ITC projects (National Research Council 2014; Rhoten 2003), and postdoctoral experiences may provide "the best opportunity for researchers to train deeply in a new discipline" (Institute of Medicine 2005). Despite considerable attention to the training of postdoctoral scholars (henceforth postdocs) in science generally, the literature offers little targeted insight into what postdocs learn from participating in ITC projects or how these experiences shape their research.

To address these questions, we combine participant interviews (n=29), bibliometric analysis (155 publications), and individual-level project data in a mixed-methods case study of a large convergence-science project, the Network for Sustainable Climate Risk Management (SCRiM). SCRiM was a $11.9-million research network funded from 2012 to 2019 through the NSF's Sustainability Research Networks (SRN) program (National Science Foundation n.d.), a precursor to the current Growing Convergence Research (National Science Foundation n.d.) and Sustainable Regional Systems Research Networks (National Science Foundation n.d.) programs. SCRiM spanned a number of disciplines, including Earth sciences, statistics, engineering, economics, decision- and risk-analysis, philosophy, and the social sciences. The project included twenty-six funded senior personnel across eight institutions as well as postdocs, graduate and undergraduate students, and administrative staff. SCRiM's mission was to identify *“sustainable, scientifically sound, technologically feasible, economically efficient, and ethically defensible climate risk management strategies”* (“The Network for Sustainable Climate Risk Management (SCRiM)” n.d.).

A distinctive feature of SCRiM allows us to address our question concerning the role of values in convergence research. SCRiM included an element of heightened attention to the values shaping and embedded within research. Project leaders brought this focus on values into SCRiM through the concept of *coupled ethical-epistemic analysis*, which refers to a practice of deliberating over method choices and research design with explicit consideration for both ethical and epistemic values, including trade-offs among values (Tuana 2013; Tuana et al. 2012; Valles, Piso, and O’Rourke 2019). For example, design choices that benefit real-world relevance and applicability of potential findings (ethical values) may sometimes trade off against the feasibility and trustworthiness of the required analysis (epistemic values), or against the prospect of fundamental scientific insights (more epistemic values). The practice of coupled ethical-epistemic analysis was discussed at SCRiM’s annual all-hands meetings, taught in the project’s annual summer school, and further developed and applied through a number of collaborative SCRiM co-supported publications (Bessette et al. 2017; Mayer et al. 2017; Vezér et al. 2018; Garner, Reed, and Keller 2016; Helgeson et al. 2021; Wong et al. 2017).

We examine the impact and reception of this focus on values in SCRiM by two routes. First, we ask SCRiM participants whether and how the project shaped their own research. Through this general, open-ended question, we survey *all changes* in the character or content of individuals’ research that participants themselves attribute to their participation in SCRiM. We identify common elements within participant responses (including, but not limited to, the values focus) and quantify their frequency across participants. Second, we ask a series of targeted questions about specific cross-cutting themes identified by the authors as representing SCRIM’s overall research agenda. One of these themes is *coupled ethical-epistemic analysis*. Using a quantitative rating scale, we asked participants to indicate how well they *understood* each of these themes and also how *important* they considered it for achieving the project’s research goals. This allows us to compare perceptions and attitudes across themes and participants.

Due to its size and duration, SCRiM provides a relatively large pool of postdoc participants whose research and training may have been shaped by their participation in a convergence research project. We investigate the participation of postdocs in convergence research by interviewing both faculty and postdoc participants in SCRiM and partitioning the results of our analyses to reveal contrasts between these two groups. We explore individual variation in quantitative responses through further steps of analysis examining potential explanatory variables drawn from individual-level project data (e.g., funding level) and from a bibliographic analysis of coauthorship patterns within SCRiM-supported publications.

2 Participant interviews

We invited all SCRiM-funded faculty and postdocs for interviews. Twenty-nine took part (66% response rate), with roughly equal numbers of faculty (n=15) and postdocs (n=14). (SCRiM also funded thirty-four graduate and undergraduate students, whom we omitted from the study for two reasons: to minimize risk of improper influence in the recruitment of institutionally and deferentially vulnerable research participants, and because students have typically done too little prior research to answer a key question regarding SCRiM's influence on their work.)

All interviews were conducted by a single interviewer in February and March of 2018 via teleconference (Zoom). Interviews were recorded and professionally transcribed. The transcripts were reviewed and corrected by the interviewer with reference to the recordings. (Due to technical difficulties, one interview was conducted by phone and not recorded; in this case, the interviewer took notes by hand.) Table S2 displays the full list of interview questions and the subset on which we report here.

To prompt responses on how SCRiM shaped participants' research, we asked participants to describe any aspects of their SCRiM-supported research that were "uniquely enabled" by the project, in contrast to what they might have done outside the project, *but with the same level of funding* (question 2, Table S2). While the question involves a hypothetical comparison, it is one that each individual is well-placed to make, since they are speaking only for themselves and their own research trajectory. By framing the question in this way, we focus on the "value added" of the project understood in causal terms.

We subsequently raised three cross-cutting themes identified in advance as characteristic of SCRiM's overall research agenda. The first theme was ***multi-objective robust decision analysis***, an approach to decision analysis that evaluates strategies based on their performance across a wide range of possibilities while eschewing premature assumptions about the relative importance of diverse objectives (Kasprzyk et al. 2013; Hadka et al. 2015). The second theme, ***identifying and characterizing deep uncertainties,*** refers to a commitment to questioning modeling assumptions and incorporating into analyses even those uncertainties that cannot be uncontroversially quantified with a single probability distribution (Kwakkel and Pruyt 2013; Lempert et al. 2006). The third theme was ***coupled ethical-epistemic analysis*** (see above). We asked participants to indicate how well they *understood* each theme and also how *important* they considered it for achieving the project's research goals—in both cases using a rating scale from one to five (questions 3–5, Table S2).

3 Transcript analysis and results

Two coders (also called raters) analyzed participant responses to the "uniquely enabled" question by first developing and testing a set of concepts corresponding to common themes within participant responses, then coding the responses (flagging instances of the concepts as they occur in the transcripts) to enable quantification of those concepts' frequency and distribution (Saldana 2015). One quarter of the transcripts were independently double-coded as a check on coder subjectivity. We coded transcripts using the Brat rapid annotation tool (Stenetorp et al. 2012) followed by processing and plotting in R (R Core Team 2013). See SI for details of code development, the final codebook, and inter-rater reliability calculations.

Despite variation in the topics and questions addressed by participants' research, we identified four general attributes shared across participant responses to the "uniquely enabled" question. *Bridging disciplines* refers to the crossing of disciplinary boundaries in one's research. *Decision relevance* refers to a greater focus on actionable insights or the relevance of research to real-world decisions. *Treatment of uncertainties* refers to expanded or improved treatment of uncertainties in methods and findings. *Attention to values* refers to greater attention to ethical values motivating the research and ethical assumptions embedded within the research. Table 1 displays interview excerpts illustrating each attribute. Figure 1 shows their frequency across participants, including the breakdown by faculty versus postdocs.

Table 1: Example interview excerpts illustrating four general attributes of the research enabled by the convergence research project SCRiM.

Attribute	Example quotations from participant interviews
bridging disciplines	Well, I think none of [my research] would have happened [without the network] because I am a hydrologist by training.
	There were feedbacks as I interacted with folks across the disciplines—the questions that I had been asking changed.
	So it was a project that immediately spanned geophysics, statistics, and a little economics, decision analysis, those sorts of things.
decision-relevance	One of the first projects that I was working on when I first started was making sea-level rise projections and then linking it to coastal defense and decision-making.
	Okay so maybe we can or can't make reliable projections of future sea-level change, but quite beyond whatever numbers we come up with here, what does this actually mean for decisions that might be made?
	I've just been starting to get involved in a project that involves more collaboration with people on the policy end and the state government end.
treatment of uncertainties	I just got into this whole area that I wouldn't have otherwise of quite advanced, for me, statistics and interpretations of results.
	There's a lot of work in where the statistics side meets the, kind of, geophysics side of things like quantifying uncertainties in a more rigorous or robust way.
	So that was the main perspective change, focussing on extreme events and uncertainty rather than other things that are, like, more focussed-on in ecology.
attention to values	Thinking about values and ethics in a different way and really thinking about, well, you know, if we're measuring or modeling these things either inappropriately or at such a high level that it doesn't matter anyway, then what are we doing?
	I remember all those discussions on values and how values affect framing of objectives and how then that affects what you get out of a decision problem in environmental sciences.
	. . . a lot more attentiveness to, you know, the ethical component.

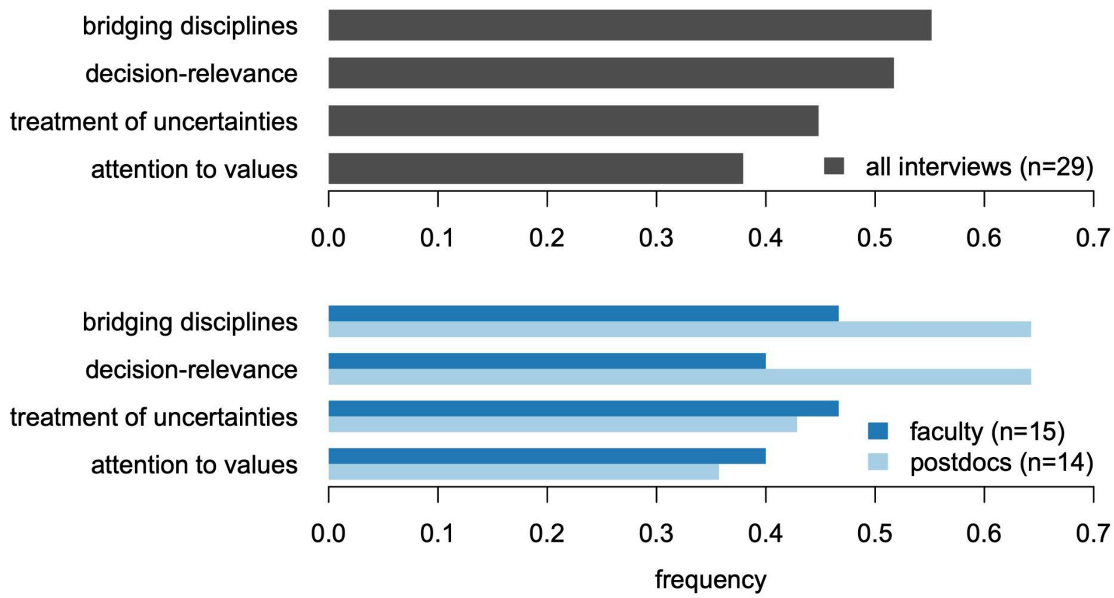

Figure 1: Four attributes of the changes to individuals' research that resulted from participation in the convergence research project SCRiM: bridging disciplines, decision relevance, treatment of uncertainties, and attention to values. These attributes summarize ways that research done through the project diverged from what each participant considered their own business-as-usual trajectory in the absence of participating in SCRiM. Bars show the fraction of participants whose description of what SCRiM enabled in their own work includes the attribute (inter-rater reliability = 0.92). The top panel shows all participants in aggregate; the bottom panel separates faculty from postdocs.

Regarding the cross-cutting themes specified in advance, all three themes were judged highly important, with *coupled ethical-epistemic analysis* rated (narrowly) the highest (Figure 2). Understanding varied more across themes, with *coupled ethical-epistemic analysis* notably lower than the other two (Figure 2). The responses on *coupled ethical-epistemic analysis* thus stand out for the—perhaps surprising—combination of highest perceived importance together with lowest perceived understanding.

All participants answered the understanding questions, but one-third did not answer the importance questions as they felt unable to make the required judgments. Figure 2 includes all responses, with each histogram normalized to display response frequencies. An alternative approach would be to discard understanding responses from participants who did not also answer the importance questions; this shifts each understanding distribution slightly to the right (i.e., participants who felt unable to judge importance had relatively low understanding). Most participants supplied integer answers to these rating-scale questions, but some (about 20%)

gave intervals such as “three to four.” In order to include all data in the histograms (Figure 2), intervals were treated as a weighted combination of integers.

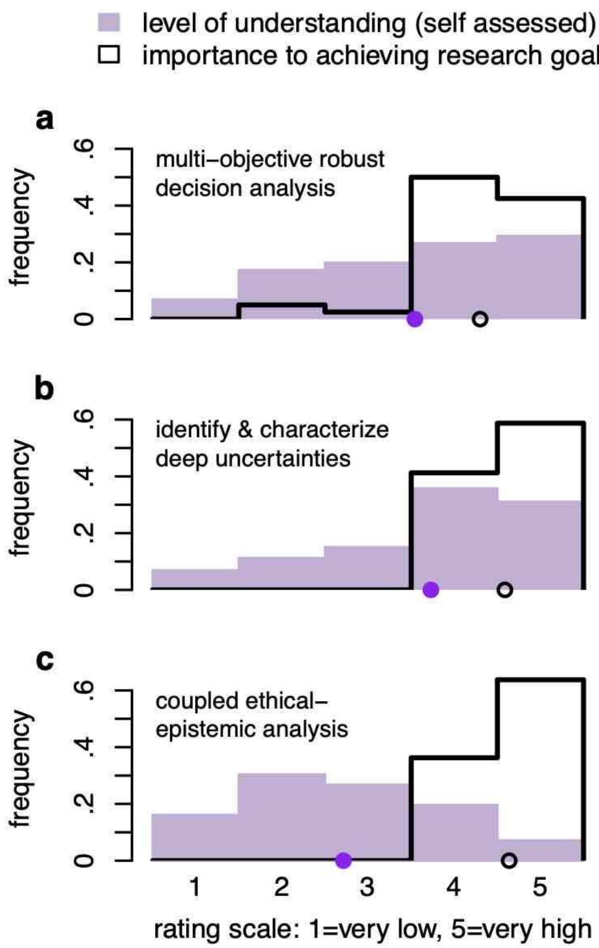

Figure 2: Histograms showing the distribution of participant responses on perceived importance and self-assessed understanding of three cross-cutting themes representing the research agenda of the project SCRiM: a) multi-objective robust decision analysis, b) identifying and characterizing deep uncertainties, and c) coupled ethical-epistemic analysis. See the main text for explanations of the three themes. Points on the x-axis indicate means.

4 Further analysis

To further contextualize and interpret results reported above, we bring to bear two additional sources of data: project management records and a database of SCRiM-supported research publications. The management records include individual-level details of project personnel, funding levels, and attendance at SCRiM’s annual all-hands meetings and summer schools.

4.1 Attention to values in uniquely enabled research

A large minority of participants (roughly forty percent) described “uniquely enabled” departures from their previous research trajectory that we subsequently coded as “attention to values” (Table 1; Figure 1). As a check on our coders’ interpretation of researcher statements, and to provide more detail on how this “attention to values” was expressed in research outputs, we examined the SCRiM-supported publications of this subset of participants. Among the publications co-authored by each participant, we identified those publications whose content best corresponded to the participant’s statements about “uniquely enabled” research that were subsequently coded as “attention to values.” Table 2 lists the publications identified through this process and summarizes aspects in which each exhibits attention to values. We emphasize that Table 2 is not an exhaustive accounting of values components across SCRiM-supported research, but rather a summary of those instances where at least one co-author (among the researchers interviewed) volunteered the attention-to-values aspect of the research as something that was uniquely enabled by their participation in SCRiM.

Of the eleven interview participants whose “uniquely enabled” research descriptions were coded as “attention to values,” nine are coauthors of papers listed in Table 2. As of the cutoff date used in our publications analysis (see below), the remaining two participants had not (yet) published research corresponding to their research descriptions coded as “attention to values.”

Table 2. The table lists SCRiM-supported publications containing research that interview participants’ considered “uniquely enabled” by SCRiM and that we subsequently categorized as including “attention to values” (Figure 1). Entries in the right column briefly summarize how each publication exhibits attention to values.

Publication	Summary of how the publication exhibits “attention to values.”
(Adler et al. 2017)	Examines ethical assumptions within calculations of the social cost of carbon (SCC). Gives new estimates of SCC by varying ethical assumptions about the distribution of impacts with special attention to a prioritarian philosophy of social welfare.
(Bakker et al. 2017)	Produces novel sea-level rise projections by questioning common implicit motivating assumptions in study design. Aims to enable identification of robust coastal adaptation strategies by erring on the side of under- rather than over-confident projections.
(Bakker, Applegate, and Keller 2016)	Develops a new computational model of Greenland ice sheet melting with design choices shaped by concern for low-probability, high-impact futures and decision makers’ interest in the reliability or robustness of adaptation strategies in the face of uncertainty.
(Bessette et al. 2017)	Investigates the values and goals of New Orleans residents in connection with coastal flood risk management through an interview-based study eliciting both knowledge about climate risks and concerns for climate impacts and management outcomes.
(Garner and Keller 2018)	Demonstrates evaluation of coastal flood risk management strategies using a more inclusive approach to values by decomposing a standard measure of outcome desirability into multiple objectives and revealing trade-offs among these objectives.

(Garner, Reed, and Keller 2016)	Demonstrates global mitigation policy analysis with a standard economic measure of policy success decomposed into four component objectives. Reveals trade-offs among these objectives and reflects on policy options rejected by standard analyses.
(Lempert, Groves, and Fischbach 2013)	Argues that in contexts of deep uncertainty, predict-then-act approaches to decision support can lead to gridlock or overconfidence while context-first approaches like robust decision making can better facilitate ethical deliberations needed to evaluate actions.
(Mayer et al. 2017)	Investigates the ethical and epistemic values driving model development and study design in climate risk research by interviewing researchers and presenting results in terms of where specific values come into play within a scientific modeling workflow.
(Quinn et al. 2018)	Demonstrates evaluation of reservoir management strategies while incorporating large-ensemble exploratory modeling of extreme precipitation changes motivated by decision makers' concerns for high reliability in meeting management objectives.
(Simpson et al. 2016)	Develops an immersive, 3D data visualization platform for displaying trade-offs among multiple policy objectives in global climate mitigation policy analysis to support multi-objective analyses and help realize their potential to facilitate ethical deliberation.
(Singh, Reed, and Keller 2015)	Demonstrates decision analysis for ecosystem management that departs from standard problem formulations with a single value function to use multiple objectives with no pre-specified weights to reflect multiple stakeholder perspectives on contested values.
(Tuana 2013)	Argues for benefits of including philosophers in interdisciplinary STEM research, such as identifying ethical aspects of research design and methods choices, highlighting implicit priorities in disciplinary practices, and aligning research with real-world needs.
(Vezér et al. 2018)	Calls attention to diverse consequences of model choice in risk analysis, including for computational tractability, transparency, flexibility of modeling frameworks, and capacity to address issues that matter to stakeholders in light of their values and goals.
(Ward et al. 2015)	Proposes a stylized environmental management problem as a new benchmark for evaluating optimization algorithms. Key problem features include many control variables, uncertain threshold response, and competing objectives with value disagreement.
(Wong et al. 2017)	Develops a new computational model of regional sea-level rise with explicit discussion of values motivating design choices, including transparency, accessibility, and flexibility, to promote sharing of computer code and reproducibility of results in the geosciences.

4.2 Variation in participant understanding of project themes

Returning to participants' rating-scale responses on SCRiM's cross-cutting themes (Figure 2), self-assessed understanding showed far more variation than the importance judgments. To investigate this variation, we use the mean of each participant's three understanding scores in a further step of analysis. Since few participants entered the network with expertise in more than one theme, we use mean understanding across the three themes as an indicator of individual intellectual integration into the project. Drawing from the data sources noted above, we construct three potential explanatory variables for this understanding indicator: (a) months of project funding received by the participant, (b) number of project-wide events attended by the participant, and (c) an index of interdisciplinary coauthorship within SCRiM publications.

To calculate the coauthorship index, we assign each SCRiM participant a primary *departmental affiliation* by reviewing participants' institutional websites (Table S8). For each publication, we then count the number of unique affiliations among SCRiM coauthors and subtract one. Summing across the publications coauthored by a participant yields that individual's coauthorship index. In other words, the index sums the number of *instances* of coauthorship with SCRiM coauthors outside the participant's home department. As a supplement to this department-based index, we repeat the same calculation using a purpose-built *disciplinary* classification (Table S9) that groups participants somewhat differently than departmental affiliations.

For calculating the coauthorship indices (and also for the document analysis above) we consider all SCRiM-supported research publications published as of March 2019 (155 publications). The rationale for this date (one year after completing the interviews) was to allow for research already underway at the time of the interviews to be included in the analysis. We used the Python package pubStats (developed for this project, and available via GitHub repository) to calculate coauthorship indices and other analytical outputs used qualitatively to guide the coauthorship analysis.

Figure 3 plots the understanding indicator against the potential explanatory variables, with faculty (top row) separated from postdocs (bottom row). These variables indeed appear relevant to faculty understanding (Figures 3a–c), with 90% confidence intervals for the slope of linear regression lines excluding zero for the coauthorship index (both variants) and bounded by zero (roughly) for both funding and attendance (see Figure 3 caption for confidence intervals). In contrast, postdoc understanding shows no statistically significant linear relationship with the potential explanatory variables (Figures 3d–f).

5 Discussion

The motivation for our study is the great and pressing need—from both scientific and societal perspectives—for improving the practice of convergence research. Specifically, we have targeted two research gaps within existing literature on the design, conduct, and evaluation of projects falling under the somewhat broader umbrella of ITC research. These gaps concern the role of values and the training of postdocs within such projects. We begin with postdocs.

5.1 Training and participation of postdocs

We identified four general attributes within participants' descriptions of what SCRiM "uniquely enabled" in their own research (Table 1; Figure 1). Among these, the two that occured most frequently align with the defining elements of convergence research. Convergence research is (1) driven by a specific and compelling problem and (2) involves deep integration across disciplines (National Science Foundation n.d.). The most common attribute, *bridging disciplines*, echoes "deep integration across disciplines," while the second most common, *decision* *relevance*, can be seen as an indicator of research "driven by a specific and compelling problem." The frequencies of these two attributes were notably higher among postdocs than among faculty, each occurring in two thirds of postdoc participants' research (Figure 1). Indeed, it is only because of postdocs that these attributes are, overall, the most common ones; among faculty, the more SCRiM-specific attributes (*treatment of uncertainties* and *attention to values*) were no less common than those characteristic of convergence research in general.

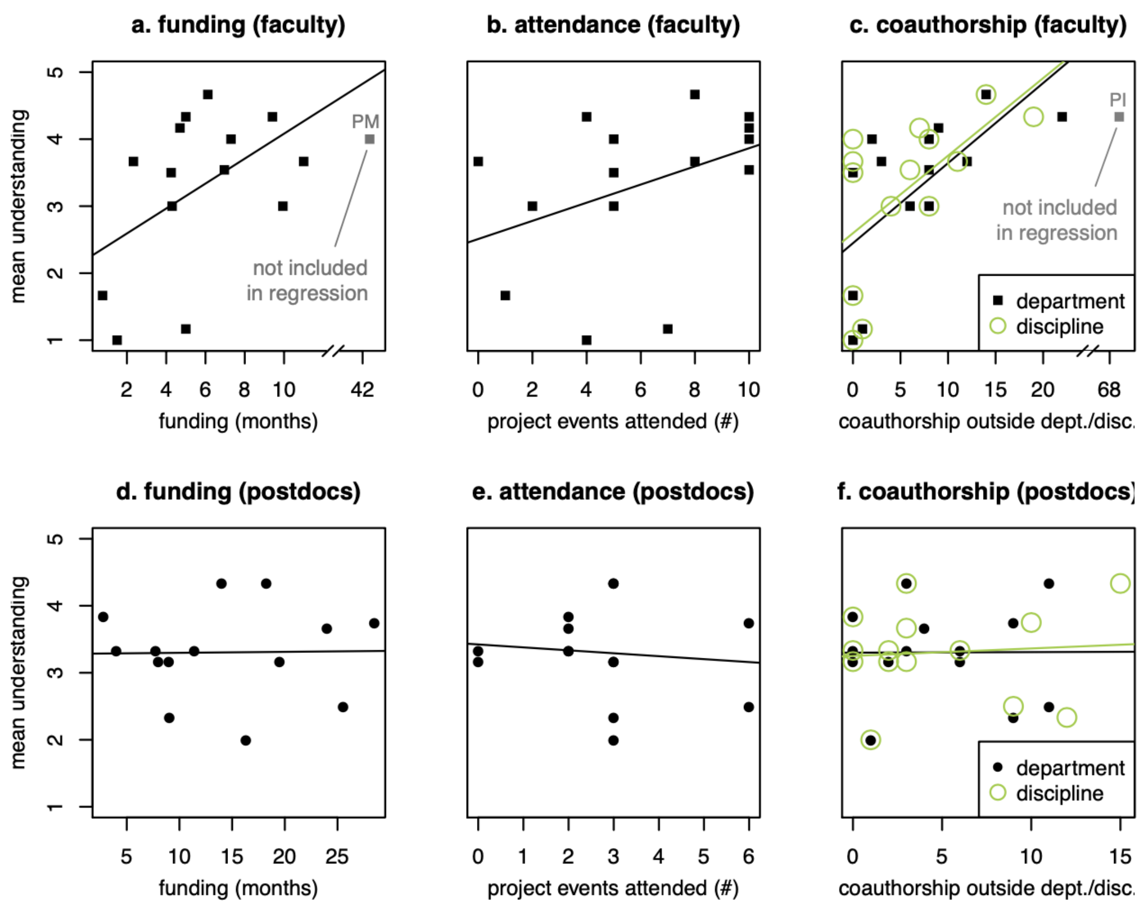

Figure 3: Self-assessed understanding of SCRiM research themes plotted against potential explanatory variables. Each point is one participant. Mean understanding refers to the average of an individual's three self-assessed understanding scores. Figures 3a–3c (top row) show faculty. Figures 3d–3f (bottom row) show postdocs. The project manager was removed from (a) and the lead PI was removed from (c) prior to calculating regression lines and confidence intervals; see supplemental information for discussion. 90% confidence intervals for the slope of linear regression lines: a. (.01, .36), b. (-.02, .29), c. dept. (.05, .19), disc. (.04, .20), d. (-.04, .04), e. (-.25, .16), f. dept. (-0.09, 0.09), disc. (-.06, -.08).

In this way, SCRiM appears to have pushed the research of postdocs in a convergence-science direction more reliably than it did the research of faculty. One possible explanation is that faculty often have broader project portfolios than postdocs, and these portfolios may be, overall, less focused on convergence research. Another possibility is that faculty research agendas are more rigid and less easily influenced. Our study suggests that insofar as postdoc research agendas are more flexible, this flexibility was exploited to bend those agendas towards the concept of convergence research. These results reinforce prior findings that postdocs can play an important role in collaborations within ITC projects (National Research Council 2014; Rhoten 2003) and that postdoc experiences can provide a good opportunity for interdisciplinary training (Institute of Medicine 2005).

While participation in SCRiM clearly made postdocs' research more interdisciplinary and more problem-driven, our analysis of individual variation in understanding of key project themes leaves us with a puzzle about postdoctoral training. In contrast to faculty, we find no statistically significant relationships between postdocs' understanding of key project themes and our potential explanatory variables (Figure 3). Yet it seems fair to assume that the amount of training that postdocs received through SCRiM is likely linked to those variables. So why does this training not improve perceived understanding of the project?

One possible explanation is that the week-long SCRiM summer school (attended by most postdocs as part of an onboarding process) raised project understanding by an amount that dwarfed the contribution of further experience and training within the project. While we suspect project understanding among postdocs would be somewhat better correlated with our explanatory variables in the absence of such a summer school, we do not believe this provides a full explanation. Average understanding of the key project themes among postdocs was just above three on a scale of one to five—slightly lower than faculty and with evident room for improvement. Further training shows no sign of providing this improvement despite large variation across postdocs in the proxies (proposed explanatory variables) we have used to quantify that training.

Another possible explanation is that postdocs are typically advised to focus on specific and tractable research questions and that this narrower focus detracts from developing an understanding of the project as a whole. On this hypothesis, variation in project understanding would be explained by the variation in specific research questions addressed by each postdoc: postdocs developed an understanding of key project themes only where this was directly relevant to their specific research questions. We consider this question an important topic for future investigation.

5.2 Values in convergence research

SCRiM's focus on values was motivated by two hypotheses: first, that explicit deliberation about ethical and epistemic values can support the design of research that is relevant to real-world problems, and second, that such deliberation can also support communication across academic cultures and between analysts and stakeholders. Our finding of "uniquely enabled" *decision relevance* and *bridging disciplines* in a majority of participants' research (Figure 1) is consistent with these motivating hypotheses. Our study does not, however, specifically assess what contribution SCRiM's focus on values made to those outcomes.

We consider the finding of "uniquely enabled" *attention to values* in the research of nearly forty percent of participants' (Figure 1) an encouraging level of uptake for an aspect of research typically not emphasized in scientific projects. Moreover, examination of corresponding SCRiM-supported publications shows a variety of substantive engagements with the values components of convergence research in general and climate risk management in particular (Table 2). Most of these papers can be placed into one or more of four (overlapping) categories of attention to values: interview-based studies of the values of stakeholders and researchers (e.g., Bessette et al. 2017; Mayer et al. 2017); discussion and demonstration of how risk analysis can better facilitate ethical deliberation (e.g., Garner, Reed, and Keller 2016; Lempert, Groves, and Fischbach 2013; Simpson et al. 2016; Tuana 2013); values-inclusive policy analysis that illuminates trade-offs among contested values (e.g., Adler et al. 2017; Garner and Keller 2018; Garner, Reed, and Keller 2016; Quinn et al. 2018; Singh, Reed, and Keller 2015); and examination of motivating values behind design choice in geoscience model building and hazard characterization (e.g., Bakker et al. 2017; Bakker, Applegate, and Keller 2016; Quinn et al. 2018; Wong et al. 2017).

When asked directly about the theme of c*oupled ethical-epistemic analysis*, participants rated its importance to achieving project research goals (see mission statement above) very highly in absolute terms and highest among the three cross-cutting themes (Figure 2). We see this broad agreement on the importance of the theme as validation of SCRiM's focus on values and encouragement to further pursue the potential for enabling convergence research through foregrounding values.

At the same time, participants' self-assessed understanding of the *coupled ethical-epistemic analysis* theme was relatively low (Figure 2). This was presumably due, at least in part, to the theme's relative novelty for most participants. But other comments suggest this is not the only reason. One participant said of *coupled ethical-epistemic analysis* "I'm a lot more aware of it," but "I'm not sure we've got a process down for, you know, incorporating it—it's not, sort of, part of the crank-turning." Another said "I like to quantify things, and I thought it was just descriptive and trying to label things, but I never saw it as a means of doing quantitative analysis." A third participant said "This was always something that was, to me, a bit of a catchphrase more than an actual application." These comments suggest that realizing the benefits of *coupled ethical-epistemic analysis* may require more codified approaches that better operationalize the practice, as well as improved integration into quantitative methods, tools, and workflows.

The values theme was led by participants from the humanities—specifically, philosophy (cf. Nagatsu et al. 2020; Robinson et al. 2016; Tuana 2020)—and the tensions in the quotes above regarding quantification and operational processes may derive in part from the disciplinary “distance” between humanists and scientists. As one participant noted: “the ethical-epistemic stuff is different—*different sort of people*—and so it’s not as well-oiled a machine as some of the other stuff.” More broadly, communication can also be more difficult between more “distant” fields. Discussing challenges within SCRiM, one participant highlighted the “difficulty of incorporating the philosophy world-view.” Another noted that some SCRiM researchers “come from quite different disciplines, especially [philosopher], and so we never sort of worked, talking the same language that much.”

Another factor that may have compounded barriers to deeper understanding of *coupled ethical-epistemic analysis* was the relatively few project participants focused primarily on the values theme. In terms of funding duration, humanists received just four percent of SCRiM’s total person-months of funding, (compared with, e.g., twenty-eight percent for meteorology/climatology, thirteen for geosciences, and ten for economics). In our experience of interdisciplinary projects, the number of opportunities for learning and bridge-building between one’s own field and another specific field is roughly proportional to the number of project participants from that given field. As a result, the composition of personnel determines a disciplinary “center of mass” towards which members’ cross-disciplinary learning is oriented. Poorly-represented fields exert less influence on other participants’ direction of travel. (Moreover, participants from poorly-represented *regions* of disciplinary space disproportionately shoulder the burdens of “long distance” cross-disciplinary learning and communication.) Allocating the representation of fields within project personnel may be an important lever for establishing a project’s disciplinary center of mass, and through this, orienting participants’ limited cross-disciplinary bandwidth toward bridges that need further development.

5.3 Uniquely enabled research

As a part of our study, we asked SCRiM participants to tell us what, if anything, the project enabled them to do in their research that they would not have done, or not been able to do, outside the project but with the same funding (Section 2). We used qualitative data analysis, with quantified inter-rater reliability (see SI), to summarize common themes that we saw in participant responses (Table 1; Figure 1). We then dug deeper into one of those themes (*attention to values*) by surveying project-supported research corresponding to participant statements about uniquely enabled research (Table 2). Here this analysis formed one part of our investigation of values and postdoc participation in convergence research, but thinking more broadly, we see this component of our study as a promising approach for characterizing and assessing the *content* of ITC research projects. As noted above, the “uniquely enabled” question and subsequent steps of analysis provide one lens on the “value added” of a project at the level of changes to individual participants’ research trajectories. This approach leverages each project participant’s intimate knowledge of their own research and project collaborations.

Addressing the simple question of what effect (if any) a funded project had on participants’ research is vital for evaluating and learning from ITC projects. Yet few studies characterizing ITC project outcomes explore the content of the research outputs, with most focusing instead on process and collaboration (Freeth and Caniglia 2020; Gaziulusoy et al. 2016; Hessels, De Jong, and Brouwer 2018; König et al. 2013; Polk 2014), researcher satisfaction (Katoh et al. 2018; Tress, Tress, and Fry 2005), or categories of outputs and outcomes (Corley, Boardman, and Bozeman 2006; Cummings and Kiesler 2005, 2007; Steger et al. 2021; Teirlinck and Spithoven 2015; Tress, Tress, and Fry 2005). A number of bibliometric analyses posit and measure indicators of interdisciplinarity in the publications of a project or researcher (Abramo, D'Angelo, and Di Costa 2012; Abramo, D'Angelo, and Zhang 2018; Bark, Kragt, and Robson 2016; Anzai et al. 2012; Porter, Roessner, and Heberger 2008), but these provide only a partial view through the imperfect lenses of author affiliations and journal classifications. Our study demonstrates a complementary approach to illuminating the content of research carried out within a project.

6 Conclusions

Here we apply our own backgrounds, perspectives, and experiences to interpret and translate the findings of this study into actionable ideas. The following recommendations are in no way comprehensive, nor do they exhaust what we have learned through SCRiM. The scope of these recommendations is limited to issues directly informed by this study. The scope is further restricted to ideas that are not already widely appreciated in published advice (see introduction). Our goal is an incremental expansion of existing understanding of best practices for the design and facilitation of convergence research projects. Other take-aways from the study are possible; these are ours:

1. **Follow the values.** Scrutinize why the research matters, to whom, and in what ways. Does the research design serve these motivations well? Acknowledge potential tensions between real-world relevance, scientific feasibility, disciplinary norms, and professional rewards. Navigate these tensions transparently. Further work is needed on how best to structure and guide these activities and improve their interface with the workflows and quantitative mindset of STEM researchers.
2. **Exploit new training opportunities.** Large convergence projects offer new kinds of training opportunities. These include chances to expand one's vision by appreciating the breadth of a project and how all of the pieces fit together. Otherwise successful postdoc training may not reliably advance this big-picture project understanding (and associated capacity to envision and lead convergence research). Where such capacity-building is a priority, new training mechanisms may be needed.
3. **Support disciplinary islands.** Participants whose disciplinary culture is far from a project's "center of mass" can face greater integration challenges. To support the integration of concepts and methods from relatively isolated disciplines, consider devising dedicated coordinating mechanisms or rethinking the project design to adjust the disciplinary balance of the project team.
4. **Leverage participant expertise in assessment.** Each project participant knows their own research far better than anyone else. This intimate knowledge can and should be exploited to better understand the outcomes of convergence funding within the research

of project participants. “Bottom up” approaches that begin from project participants’ perspective and experience can supplement “top down” bibliographic approaches to quantifying interdisciplinarity and other aspects of research content.

Declarations

Funding: This work was supported by the National Science Foundation through the Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507 and by the Penn State Center for Climate Risk Management (CLIMA).

Conflicts of interest: The authors are not aware of any affiliations, funding sources, or relationships that could be perceived as potential sources of bias.

Code availability: All code used in this study is available via GitHub repositories:

- https://github.com/scrim-network/pubStats
- https://github.com/scrim-network/convergenceStudyFigures

Data availability: The data used in this study are supplied to the greatest extent consistent with protecting our study participants. See SI Section 8 and Figure S3 for details on the data provided and withheld. Supplied data are available via GitHub repository:

- https://github.com/scrim-network/convergenceStudyFigures

Ethics approval and consent to participate: All research was conducted with approval from the Institutional Review Board of Pennsylvania State University (study ID: STUDY00008957). All participants provided informed consent.

References

Abramo, Giovanni, Ciriaco Andrea D’Angelo, and Flavia Di Costa. 2012. “Identifying Interdisciplinarity through the Disciplinary Classification of Coauthors of Scientific Publications.” *Journal of the American Society for Information Science and Technology* 63 (11): 2206–22.

Abramo, Giovanni, Ciriaco Andrea D’Angelo, and Lin Zhang. 2018. “A Comparison of Two Approaches for Measuring Interdisciplinary Research Output: The Disciplinary Diversity of Authors vs the Disciplinary Diversity of the Reference List.” *Journal of Informetrics* 12 (4): 1182–93.

Adler, Matthew, David Anthoff, Valentina Bosetti, Greg Garner, Klaus Keller, and Nicolas Treich. 2017. “Priority for the Worse-off and the Social Cost of Carbon.” *Nature Climate Change* 7 (6): 443–49.

Anzai, Tomohiro, Ryoichi Kusama, Hiroyuki Kodama, and Shintaro Sengoku. 2012. “Holistic Observation and Monitoring of the Impact of Interdisciplinary Academic Research Projects: An Empirical Assessment in Japan.” *Technovation* 32 (6): 345–57.

Bakker, Alexander M. R., Patrick J. Applegate, and Klaus Keller. 2016. “A Simple, Physically Motivated Model of Sea-Level Contributions from the Greenland Ice Sheet in Response to

Temperature Changes.” *Environmental Modelling & Software* 83 (September): 27–35.
Bakker, Alexander M. R., Tony E. Wong, Kelsey L. Ruckert, and Klaus Keller. 2017. “Sea-Level Projections Representing the Deeply Uncertain Contribution of the West Antarctic Ice Sheet.” *Scientific Reports* 7 (1): 3880.
Bark, Rosalind H., Marit E. Kragt, and Barbara J. Robson. 2016. “Evaluating an Interdisciplinary Research Project: Lessons Learned for Organisations, Researchers and Funders.” *International Journal of Project Management* 34 (8): 1449–59.
Bessette, D. L., L. A. Mayer, B. Cwik, M. Vezér, K. Keller, R. J. Lempert, and N. Tuana. 2017. “Building a Values-informed Mental Model for New Orleans Climate Risk Management.” *Risk Analysis* 37 (10): 1993–2004.
Brister, Evelyn. 2016. “Disciplinary Capture and Epistemological Obstacles to Interdisciplinary Research: Lessons from Central African Conservation Disputes.” *Studies in History and Philosophy of Biological and Biomedical Sciences* 56 (April): 82–91.
Bromham, Lindell, Russell Dinnage, and Xia Hua. 2016. “Interdisciplinary Research Has Consistently Lower Funding Success.” *Nature* 534 (7609): 684–87.
Corley, Elizabeth A., P. Craig Boardman, and Barry Bozeman. 2006. “Design and the Management of Multi-Institutional Research Collaborations: Theoretical Implications from Two Case Studies.” *Research Policy* 35 (7): 975–93.
Cummings, Jonathon N., and Sara Kiesler. 2005. “Collaborative Research Across Disciplinary and Organizational Boundaries.” *Social Studies of Science* 35 (5): 703–22.
———. 2007. “Coordination Costs and Project Outcomes in Multi-University Collaborations.” *Research Policy* 36 (10): 1620–34.
Diekmann, Sven, and Martin Peterson. 2013. “The Role of Non-Epistemic Values in Engineering Models.” *Science and Engineering Ethics* 19 (1): 207–18.
Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, et al. 2018. “Science of Science.” *Science* 359 (6379). https://doi.org/10.1126/science.aao0185.
Freeth, Rebecca, and Guido Caniglia. 2020. “Learning to Collaborate While Collaborating: Advancing Interdisciplinary Sustainability Research.” *Sustainability Science* 15 (1): 247–61.
Garner, Gregory, and Klaus Keller. 2018. “Using Direct Policy Search to Identify Robust Strategies in Adapting to Uncertain Sea-Level Rise and Storm Surge.” *Environmental Modelling & Software*. https://doi.org/10.1016/j.envsoft.2018.05.006.
Garner, Gregory, Patrick Reed, and Klaus Keller. 2016. “Climate Risk Management Requires Explicit Representation of Societal Trade-Offs.” *Climatic Change* 134 (4): 713–23.
Gaziulusoy, A. Idil, Chris Ryan, Stephen McGrail, Philippa Chandler, and Paul Twomey. 2016. “Identifying and Addressing Challenges Faced by Transdisciplinary Research Teams in Climate Change Research.” *Journal of Cleaner Production* 123 (June): 55–64.
Hadka, David, Jonathan Herman, Patrick Reed, and Klaus Keller. 2015. “An Open Source Framework for Many-Objective Robust Decision Making.” *Environmental Modelling & Software* 74 (December): 114–29.
Helgeson, Casey, Vivek Srikrishnan, Klaus Keller, and Nancy Tuana. 2021. “Why Simpler Computer Simulation Models Can Be Epistemically Better for Informing Decisions.” *Philosophy of Science* 88. https://philpapers.org/rec/HELWSC-3.
Hessels, Laurens K., Stefan P. L. De Jong, and Stijn Brouwer. 2018. “Collaboration between Heterogeneous Practitioners in Sustainability Research: A Comparative Analysis of Three Transdisciplinary Programmes.” *Sustainability: Science Practice and Policy* 10 (12): 4760.
Huutoniemi, Katri, Julie Thompson Klein, Henrik Bruun, and Janne Hukkinen. 2010. “Analyzing Interdisciplinarity: Typology and Indicators.” *Research Policy* 39 (1): 79–88.

Institute of Medicine. 2005. *Facilitating Interdisciplinary Research*. Washington DC: National Academies Press.

Ioannidis, John P. A. 2018. “Meta-Research: Why Research on Research Matters.” *PLoS Biology* 16 (3): e2005468.

Kasprzyk, Joseph R., Shanthi Nataraj, Patrick M. Reed, and Robert J. Lempert. 2013. “Many Objective Robust Decision Making for Complex Environmental Systems Undergoing Change.” *Environmental Modelling & Software* 42 (April): 55–71.

Katoh, Shogo, Giancarlo Lauto, Tomohiro Anzai, and Shintaro Sengoku. 2018. “Identification of Factors to Promote Interdisciplinary Research: A Trial at COINS.” *2018 Portland International Conference on Management of Engineering and Technology (PICMET)*. https://doi.org/10.23919/picmet.2018.8481881.

Klein, Julie T. 2008. “Evaluation of Interdisciplinary and Transdisciplinary Research: A Literature Review.” *American Journal of Preventive Medicine* 35 (2 Suppl): S116–23.

Kodama, Hiroyuki, Kenji Watatani, and Shintaro Sengoku. 2013. “Competency-Based Assessment of Academic Interdisciplinary Research and Implication to University Management.” *Research Evaluation* 22 (2): 93–104.

König, Bettina, Katharina Diehl, Karen Tscherning, and Katharina Helming. 2013. “A Framework for Structuring Interdisciplinary Research Management.” *Research Policy* 42 (1): 261–72.

Kwakkel, Jan H., and Erik Pruyt. 2013. “Exploratory Modeling and Analysis, an Approach for Model-Based Foresight under Deep Uncertainty.” *Technological Forecasting and Social Change* 80 (3): 419–31.

Lang, Daniel J., Arnim Wiek, Matthias Bergmann, Michael Stauffacher, Pim Martens, Peter Moll, Mark Swilling, and Christopher J. Thomas. 2012. “Transdisciplinary Research in Sustainability Science: Practice, Principles, and Challenges.” *Sustainability Science* 7 (1): 25–43.

Laursen, Bethany K., Chad Gonnerman, and Stephen J. Crowley. 2021. “Improving Philosophical Dialogue Interventions to Better Resolve Problematic Value Pluralism in Collaborative Environmental Science.” *Studies in History and Philosophy of Science. Part B. Studies in History and Philosophy of Modern Physics* 87 (June): 54–71.

Lempert, Robert J., David G. Groves, and Jordan R. Fischbach. 2013. “Is It Ethical to Use a Single Probability Density Function.” *Santa Monica, CA: RAND Corporation*. https://www.rand.org/content/dam/rand/pubs/working_papers/WR900/WR992/RAND_WR992.pdf.

Lempert, Robert J., David G. Groves, Steven W. Popper, and Steve C. Bankes. 2006. “A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios.” *Management Science* 52 (4): 514–28.

Lundershausen, Johannes. 2018. “The Anthropocene Working Group and Its (inter-)disciplinarity.” *Sustainability: Science Practice and Policy* 14 (1): 31–45.

MacLeod, Miles. 2018. “What Makes Interdisciplinarity Difficult? Some Consequences of Domain Specificity in Interdisciplinary Practice.” *Synthese* 195 (2): 697–720.

MacLeod, Miles, and Michiru Nagatsu. 2018. “What Does Interdisciplinarity Look like in Practice: Mapping Interdisciplinarity and Its Limits in the Environmental Sciences.” *Studies in History and Philosophy of Science* 67 (February): 74–84.

MacLeod, Miles, and Nancy J. Nersessian. 2014. “Strategies for Coordinating Experimentation and Modeling in Integrative Systems Biology.” *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution* 322 (4): 230–39.

Mayer, Lauren A., Kathleen Loa, Bryan Cwik, Nancy Tuana, Klaus Keller, Chad Gonnerman, Andrew M. Parker, and Robert J. Lempert. 2017. “Understanding Scientists’ Computational

Modeling Decisions about Climate Risk Management Strategies Using Values-Informed Mental Models.” *Global Environmental Change: Human and Policy Dimensions* 42: 107–16.
McLeish, Tom, and Veronica Strang. 2016. “Evaluating Interdisciplinary Research: The Elephant in the Peer-Reviewers’ Room.” *Palgrave Communications* 2 (1): 1.
Nagatsu, Michiru, Taylor Davis, C. Tyler DesRoches, Inkeri Koskinen, Miles MacLeod, Milutin Stojanovic, and Henrik Thorén. 2020. “Philosophy of Science for Sustainability Science.” *Sustainability Science* 15 (6): 1807–17.
National Academies of Sciences, Engineering, and Medicine. 2019. *Fostering the Culture of Convergence in Research: Proceedings of a Workshop*. Washington DC: National Academies Press.
National Academies of Sciences, Engineering and Medicine. 2021. *Measuring Convergence in Science and Engineering: Proceedings of a Workshop*. The National Academies Press.
National Research Council. 2014. *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond*. Washington DC: The National Academies Press.
National Science Foundation. n.d. “‘Growing Convergence Research Program Solicitation.’” Nsf.gov. Accessed May 15, 2020a. https://www.nsf.gov/pubs/2019/nsf19551/nsf19551.htm.
———. n.d. “NSF’s Ten Big Ideas.” Accessed July 6, 2020b. https://www.nsf.gov/news/special_reports/big_ideas/index.jsp.
———. n.d. “‘Sustainability Research Networks Program Solicitation.’” Nsf.gov. Accessed December 1, 2020c. https://www.nsf.gov/pubs/2011/nsf11574/nsf11574.htm.
———. n.d. “‘Sustainable Regional Systems Research Networks Program Solicitation.’” Nsf.gov. Accessed November 19, 2021d. https://www.nsf.gov/pubs/2020/nsf20611/nsf20611.htm.
O’Brien, Karen L., and Johanna Wolf. 2010. “A Values-Based Approach to Vulnerability and Adaptation to Climate Change: A Values-Based Approach.” *Wiley Interdisciplinary Reviews. Climate Change* 1 (2): 232–42.
Polk, Merritt. 2014. “Achieving the Promise of Transdisciplinarity: A Critical Exploration of the Relationship between Transdisciplinary Research and Societal Problem Solving.” *Sustainability Science* 9 (4): 439–51.
Porter, Alan L., Alex S. Cohen, J. David Roessner, and Marty Perreault. 2007. “Measuring Researcher Interdisciplinarity.” *Scientometrics* 72 (1): 117–47.
Porter, Alan L., David J. Roessner, and Anne E. Heberger. 2008. “How Interdisciplinary Is a given Body of Research?” *Research Evaluation* 17 (4): 273–82.
Porter, Alan L., J. David Roessner, Alex S. Cohen, and Marty Perreault. 2006. “Interdisciplinary Research: Meaning, Metrics and Nurture.” *Research Evaluation* 15 (3): 187–95.
Porter, Alan L., David J. Schoeneck, David Roessner, and Jon Garner. 2010. “Practical Research Proposal and Publication Profiling.” *Research Evaluation* 19 (1): 29–44.
Quinn, Julianne D., Patrick M. Reed, Matteo Giuliani, Andrea Castelletti, Jared W. Oyler, and Robert E. Nicholas. 2018. “Exploring How Changing Monsoonal Dynamics and Human Pressures Challenge Multireservoir Management for Flood Protection, Hydropower Production, and Agricultural Water Supply.” *Water Resources Research* 54 (7): 4638–62.
R Core Team. 2013. “R: A Language and Environment for Statistical Computing.” Vienna, Austria. http://cran.univ-paris1.fr/web/packages/dplR/vignettes/intro-dplR.pdf.
Rhoten, Diana. 2003. “A Multi-Method Analysis of the Social and Technical Conditions for Interdisciplinary Collaboration.” *Final Report, National Science Foundation BCS-0129573*. http://ssrc-cdn1.s3.amazonaws.com/crmuploads/new_publication_3/a-multi-method-analysis-of-the-social-and-technical-conditions-for-interdisciplinary-collaboration.pdf.
Rijnsoever, Frank J. van, and Laurens K. Hessels. 2011. “Factors Associated with Disciplinary

and Interdisciplinary Research Collaboration." *Research Policy* 40 (3): 463–72.
Robinson, Brian, Stephanie E. Vasko, Chad Gonnerman, Markus Christen, Michael O'Rourke, and Daniel Steel. 2016. "Human Values and the Value of Humanities in Interdisciplinary Research." *Cogent Arts & Humanities* 3 (1): 1123080.
Saldana, Johnny. 2015. *The Coding Manual for Qualitative Researchers*. SAGE.
Siedlok, Frank, and Paul Hibbert. 2014. "The Organization of Interdisciplinary Research: Modes, Drivers and Barriers." *International Journal of Management Reviews* 16 (2): 194–210.
Siedlok, Frank, Paul Hibbert, and John Sillince. 2015. "From Practice to Collaborative Community in Interdisciplinary Research Contexts." *Research Policy* 44 (1): 96–107.
Simpson, Mark, Jan Oliver Wallgrün, Alexander Klippel, Liping Yang, Gregory Garner, Klaus Keller, Danielle Oprean, and Saurabh Bansal. 2016. "Immersive Analytics for Multi-Objective Dynamic Integrated Climate-Economy (DICE) Models." In *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, 99–105. ISS '16 Companion. New York, NY, USA: Association for Computing Machinery.
Singh, Riddhi, Patrick M. Reed, and Klaus Keller. 2015. "Many-Objective Robust Decision Making for Managing an Ecosystem with a Deeply Uncertain Threshold Response." *Ecology and Society*. https://doi.org/10.5751/es-07687-200312.
Steger, Cara, Julia A. Klein, Robin S. Reid, Sandra Lavorel, Catherine Tucker, Kelly A. Hopping, Rob Marchant, et al. 2021. "Science with Society: Evidence-Based Guidance for Best Practices in Environmental Transdisciplinary Work." *Global Environmental Change: Human and Policy Dimensions* 68 (May): 102240.
Stenetorp, Pontus, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun 'ichi Tsujii. 2012. "Brat: A Web-Based Tool for NLP-Assisted Text Annotation." In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 102–7. Avignon, France: Association for Computational Linguistics.
Sylvan Katz, J., and Ben R. Martin. 1997. "What Is Research Collaboration?" *Research Policy* 26: 1–18.
Teirlinck, Peter, and André Spithoven. 2015. "How the Nature of Networks Determines the Outcome of Publicly Funded University Research Projects." *Research Evaluation* 24 (2): 158–70.
"The Network for Sustainable Climate Risk Management (SCRiM)." n.d. Scrim.psu.edu. Accessed May 26, 2020. http://www.scrimhub.org/about/overview/.
Tress, Bärbel, Gunther Tress, and Gary Fry. 2005. "Researchers' Experiences, Positive and Negative, in Integrative Landscape Projects." *Environmental Management* 36 (6): 792–807.
Tschakert, Petra, Jon Barnett, Neville Ellis, Carmen Lawrence, Nancy Tuana, Mark New, Carmen Elrick-Barr, Ram Pandit, and David Pannell. 2017. "Climate Change and Loss, as If People Mattered: Values, Places, and Experiences: Climate Change and Loss, as If People Mattered." *Wiley Interdisciplinary Reviews. Climate Change* 8 (5): e476.
Tuana, Nancy. 2013. "Embedding Philosophers in the Practices of Science: Bringing Humanities to the Sciences." *Synthese* 190 (11): 1955–73.
———. 2020. "Values-Informed Decision Support: The Place of Philosophy." In *Philosophy for the Real World: An Introduction to Field Philosophy with Case Studies and Practical Strategies*, edited by Evelyn Brister And, 143–59. Taylor & Francis/Routledge.
Tuana, Nancy, Ryan L. Sriver, Toby Svoboda, Roman Olson, Peter J. Irvine, Jacob Haqq-Misra, and Klaus Keller. 2012. "Towards Integrated Ethical and Scientific Analysis of Geoengineering: A Research Agenda." *Ethics, Policy & Environment* 15 (2): 136–57.
Valles, Sean A., Zachary Piso, and Michael O'Rourke. 2019. "Coupled Ethical-Epistemic

Analysis as a Tool for Environmental Science." *Ethics, Policy & Environment* 22 (3): 267–86.
Vezér, Martin, Alexander Bakker, Klaus Keller, and Nancy Tuana. 2018. "Epistemic and Ethical Trade-Offs in Decision Analytical Modelling." *Climatic Change* 147 (1): 1–10.
Wall, Tamara U., Alison M. Meadow, and Alexandra Horganic. 2017. "Developing Evaluation Indicators to Improve the Process of Coproducing Usable Climate Science." *Weather, Climate, and Society* 9 (1): 95–107.
Ward, Victoria L., Riddhi Singh, Patrick M. Reed, and Klaus Keller. 2015. "Confronting Tipping Points: Can Multi-Objective Evolutionary Algorithms Discover Pollution Control Tradeoffs given Environmental Thresholds?" *Environmental Modelling & Software* 73 (November): 27–43.
West, Simon, Lorrae van Kerkhoff, and Hendrik Wagenaar. 2019. "Beyond 'linking Knowledge and Action': Towards a Practice-Based Approach to Transdisciplinary Sustainability Interventions." *Policy Studies* 40 (5): 534–55.
Wong, Tony E., Alexander M. R. Bakker, Kelsey Ruckert, Patrick Applegate, Aimée B. A. Slangen, and Klaus Keller. 2017. "BRICK v0. 2, a Simple, Accessible, and Transparent Model Framework for Climate and Regional Sea-Level Projections." *Geoscientific Model Development* 10 (7): 2741.

Supplemental Information

For the manuscript:
Attention to values helps shape convergence science

Contents

---

1 Breakdown of personnel and participation by institution

Sections 1 and 2 of the article report numbers of SCRiM personnel and their interview participation rates. Here we provide a breakdown of those numbers by institution (Table S1). The institutional organization of SCRiM followed a "hub and spokes" design, with a larger number of personnel (including the PI) at a central "hub" institution and smaller numbers of personnel at each of seven "spokes" institutions.

Table S1: Numbers of funded faculty (PI, co-PIs, site-leads, and senior personnel) and postdocs at each SCRiM institution. Numbers following the slash report how many participated in the interviews. E.g., institution a (the hub) hosted ten faculty among whom eight were interviewed. (Letters assigned to institutions are the same as those used in Table S10 and Figure S2.)

institution	a	b	c	d	e	f	g	h
# funded faculty	10/8	1/1	1/1	2/1	3/1	1/0	5/1	1/1
# postdocs	15/13	1/0	2/2	1/0	0/0	1/0	0/0	0/0

2 Interview questions

The main text reports our analysis of a subset of the interview questions. Table S2 provides the full list of questions.

Table S2: The full list of questions used in the participant interviews. In this study, we report our analysis of responses to the questions highlighted in yellow.

1	What do you find to be the unique features of SCRiM's approach to research?
2	Can you describe any aspects of your own research that were uniquely enabled by SCRiM? In other words, anything you were able to do thanks to participating in SCRiM, that you wouldn't have accomplished on your own with the same funding but without the network.
3	One element of SCRiM research is many-objective robust decision analysis.
a	What is your level of understanding of many-objective robust decision analysis, on a scale from one to five, where one is very low and five is very high?
b	In your view, how important is many-objective robust decision analysis for achieving SCRiM research goals, on a scale of one to five, where one is not important and five is very important?
c	Does your work include many-objective robust decision analysis?
d	What about your work before SCRiM?
e	(Various follow-ups, depending on answers to previous questions.)
4	Another element of SCRiM research is identifying and characterizing decision-relevant deep uncertainties.
a	What is your level of understanding of identifying and characterizing decision- relevant deep uncertainties, on a scale from one to five, where one is very low and five is very high?
b	In your opinion, how important is identifying and characterizing decision-relevant deep uncertainties for achieving SCRiM research goals, on a scale of one to five, where one is not important and five is very important?
c	Does your work involve identifying and characterizing decision-relevant deep uncertainties?
d	What about your work before SCRiM?
e	(Various follow-ups, depending on answers to previous questions.)
5	Another element of SCRiM research is coupled ethical-epistemic analysis.
a	What is your level of understanding of coupled ethical-epistemic analysis, on a scale from one to five, where one is very low and five is very high?
b	In your opinion, how important is coupled ethical-epistemic analysis for achieving SCRiM research goals, on a scale of one to five, where one is not important and five is very important?
c	Does your work include coupled ethical-epistemic analysis?
d	What about your work before SCRiM?
e	(Various follow-ups, depending on answers to previous questions.)
6	Has participation in SCRIM changed your career trajectory?
7	Has participation in SCRIM changed your approach to communicating research?
8	In your opinion, in what ways has the SCRIM network been successful?
9	In your view, what could have been done better?
10	Is there anything else that you would like to add?

3 Details of transcript analysis

Here we provide further details on the qualitative data analysis for Figure 1. To analyze responses to question 2 (see Table S2), we used a common approach to qualitative data in which concepts of interest are defined and the locations in the data where those concepts occur are flagged, allowing for quantitative summary of the frequency and distribution of the concepts. The list of concepts, together with descriptions of what constitutes an occurrence of each concept, is called the codebook. The activity of flagging parts of the data with these concepts (or codes) is called coding. In this case, two coders worked collaboratively through two phases of analysis: developing the codebook and coding the transcripts.

3.1 Initial codebook development

Coder 1 (also the interviewer) studied the transcripts and then devised and iteratively refined an initial set of codes. The criteria applied during this process were to generate: (a) a relatively small set of codes, (b) each conceptually clear and distinct from the others, that (c) together encompass the salient elements that occur multiple times across participant responses. These criteria led to an initial coding system consisting of two groups of codes with eight codes in each group (Table S3). The codes in group A label phases of research (motivation, design, execution), types of activity (collaboration, exposure, mentoring), and things that support research (infrastructure, tools), with no regard for the topics or themes addressed in or through the research. Codes in group B label aspects of the intellectual content of the research activities or the focus of the research. Only group-B results are reported in the main text. Group-A coding was ultimately used only as a preliminary step in filtering and preparing transcripts.

Table S3: The initial codebooks (abbreviated) for group-A and group-B codes, developed by coder 1 prior to clarification and testing involving coder 2.

group A (phases, activities, and supporting elements of research)	
motivation	Motivations or purpose behind research activities
design	Formulation or design of research questions or analyses
research	Research activities and results (general catch-all)
collaboration	Collaborating or consulting on specific research projects
exposure	Communication or interaction outside of research projects
mentoring	Mentoring or guidance (both formal and informal)
infrastructure	Research infrastructure such as computing resources
tools	Reusable, shareable research tools (typically software)
group B (intellectual content, focus, and other features of research)	
uncertainty	Emphasizes or expands treatment of uncertainties
values	Involves attention to values
decision	Emphasizes decision-relevance or actionable insight
skill sets	Skill sets combined
disciplines	Disciplines or disciplinary knowledge integrated
new topic	Topic or application is new to the researcher
stakeholders	Stakeholder participation
sustained	Long-term or repeated interactions

3.2 Codebook testing and revision (group-A codes)

Coder 2 read the draft codebook for group A, and coders 1 and 2 iteratively refined the codebook across multiple rounds of discussion and revision focussed on clarifying the meaning of the codes and establishing a shared understanding of those meanings. These revisions produced a revised codebook. (These steps were taken without reference to any transcripts. Coder 2 saw the interview transcripts only later, at the stage of applying the final codebook.)

At this point, three researchers who did not participate in the study were interviewed (using slightly modified question wording) to generate additional transcripts for coding practice. These researchers were two postdoctoral scholars working with SCRiM's PI on similar topics (though not funded through SCRiM), and one SCRiM-funded doctoral student. These additional interviews were recorded and transcribed by the interviewer. We refer to the resulting transcripts as the practice transcripts. (Otherwise, transcripts refers to the transcribed responses of the twenty-nine study participants.)

Coders 1 and 2 independently coded the practice transcripts using the revised group-A codebook. To quantify inter-rater reliability of this practice coding, the practice transcripts were blocked into paragraph-sized passages, and each coder's results were summarized in terms of which codes were present (one or more instances) within each block of text. In this way, each coder can be seen as having made x times y binary judgments (present/absent), where x is the number of text blocks and y the number of codes. With the results so characterized, Cohen's kappa (Cohen, 1960) can be applied to quantify inter-rater reliability. The "practice" column of Table S4 reports kappa scores for each code and for the set of codes.

Table S4: Inter-rater reliability of coding during development and final application of code group A. f refers to frequency of agreement and κ refers to Cohen's Kappa (Cohen, 1960).

	practice	sample		lumped		reduced	
code	κ	f	κ	f	κ	f	κ
motivation	.00	.63	-.20				
design	-.66	.63	.14	1.0	1.0	1.0	1.0
research (unspecified)	.00	.88	.75				
collaboration	.40	.88	.71	.88	.71		
exposure	1.0	.88	.71	.88	.71		
mentoring	.57	1.0	–	1.0	–		
infrastructure	–	1.0	1.0	1.0	1.0		
tools	.00	1.0	1.0	1.0	1.0		
overall	.32	.86	.60	.96	.89	1.0	1.0

The coders then examined each other's coding of the practice transcripts and every instance of disagreement was discussed to discover the source of the disagreement, converge on a consensus coding, further refine the meaning and application conditions for the codes, and revise the codebook. This produced the final codebook for group A (Table S5).

3.3 Coding transcripts (group-A codes)

To provide a check on coder subjectivity in the analysis of participant responses, a portion of the transcripts were double coded. (Coder 1 was also the interviewer and had previously studied

Table S5: Final codebook for group-A codes. Coding with group-A was ultimately used only as a preliminary step to filter and prepare transcripts for analysis with group-B codes (Table S6).

code	short description / notes
motivation	Motivations or purpose behind research activities
	Talk of the consequences of doing research, such as "I learned a lot about . . . " should not be coded *motivation* without indication that this consequence was anticipated and considered as a reason for pursuing the research.
design	Formulation or design of research questions or analyses
	Talk of new "methods" or "approaches" may indicate *design* (though not if mainly professional development, e.g., "I learned new methods"). Shifting topics is *design* only if the shift is driven by change in research questions.
collaboration	Collaborating or consulting on specific research projects
	The word "we" (rather than "I") may suggest collaboration but should be coded as such only if there is further emphasis on the collaborative nature of the research.
research	Catch-all for research activities and results discussed in general terms
	When possible, use the more specific codes *motivation*, *design*, or *collaboration* instead.
mentoring	Mentoring or guidance (formal or informal)
	If within a joint research project, remember to also apply *collaboration*.
exposure	Communication or interaction outside of research projects
	Applies to within-network interaction not directed towards a specific research project. Includes passive exposure to network members' research or expertise, for example, at summer schools and all-hands meetings.
infrastructure	Research infrastructure such as computing resources and staff
tools	Reusable, shareable research code/software

the transcripts in detail during the development of the initial draft codebook; coder 2 was seeing the transcripts for the first time during this coding exercise.) The two coders independently coded eight participant responses (28%) using the final group-A codebook developed above. The eight participants were selected randomly (uniform distribution) after removing the five longest responses (for efficiency in the double-coding exercise).

To quantify inter-rater reliability, each participant's full response was treated as a single block of text and coding results were summarized in terms of presence (one or more instances) or absence of each code within each block. (Note that this is the same resolution at which we report coding results in the main text of the paper.) In this way, each coder can be seen as having made 8*8=64 binary (present/absent) judgments. Characterizing the coding results in this way, the "sample" column of Table S4 reports both the frequency of agreement and the kappa score for each code separately and the set of codes together.

Low inter-rater reliability for the codes representing phases of research (motivation, design, and execution) indicated a high degree of subjectivity in the application of those codes, despite efforts to reduce this during codebook development. We therefore lumped those three codes under the general concept of undifferentiated research activity, defining the new code research as the logical disjunction of motivation, design, and execution. Agreement on this broad (and less informative) research code was very high (see column "lumped" in Table S4).

Though inter-rater reliability on the codes mentoring, infrastructure, and tools was high, these codes occurred very rarely in the eight double-coded responses and we set them aside at that point. We decided at this point to use the group-A coding only as a filter to determine which responses—and within a response, which parts—would be coded using the group-B codes. We would apply the group-B codes only to statements that did in fact refer to the participant's own SCRiM-supported research, as indicated by the group A codes research and collaboration (see column "reduced" in Table S4). While question 2 explicitly asks participants to talk about their own research, interview subjects often talk about other things as well, and we used the group-A codes—in the very reduced form just two codes—to separate out the statements that in fact refer to the participant's SCRiM-supported research.

3.4 Codebook testing and revision (group-B codes)

Testing and revision of the group-B codes followed the same steps as for group A. First, coders 1 and 2 discussed and revised the draft codebook (Table S3) to clarify codes and establish shared understandings of their meaning. Second, coders 1 and 2 applied the revised codebook independently to the practice transcripts. Inter-rater reliability was quantified as per the group A method, and every instance of disagreement was discussed in order to find the source of disagreement, converge on a consensus coding, and further refine the codebook. During this process, we trimmed and consolidated the original eight codes (Table S3) down to the four codes in the final group-B codebook (Table S6).

Table S6: Final codebook for group-B codes, including short code descriptions, additional notes, indicator words, and guiding examples. See Table S7 for inter-rater reliability.

code	short description / notes and indicator words / example changes to research
uncertainty	Discussion of uncertainty
	Includes typical use of the words: extremes, extreme outcomes, extreme values, GEV functions/distributions, tails, probability, distributions, ensembles, scenarios, statistics, calibration, risk, reliability, or certainty.
	• expanded or improved treatment of uncertainty • increased attention to method choices in addressing uncertainty • better integration of treatment of uncertainty with other aspects of research (including aspects under the values and decision-relevance codes)
relevance	Discussion of research on decisions or of relevance of one's research to decisions
	Includes normal use of the terms mitigation, adaptation, integrated assessment, decision analysis, decision making, RDM, or actionable insight, as well as discussion of the information needs of stakeholders or decision makers.
	• shifting to more decision-relevant research topics • more attention to actionable insights • improving the decision-relevant aspects of the work • better integration of decision relevance with other aspects of research
values	Discussion of ethical values
	Can be anyone's values. Includes talk of objectives, reasons, or why people care about risks and climate/weather impacts.
	• greater attention to values shaping or woven into research • explicit consideration of values not previously appreciated • shifting to research topics where values play a larger role • better integration of attention to values with other aspects of research (including interaction with uncertainties or decision-relevance)
disciplines	Discussion of combining disciplinary knowledge or crossing disciplinary lines
	Applies to both research content and team makeup. We don't judge depth of integration. Talk of fields, areas, or skill sets may indicate disciplines.
	• Incorporating new other-discipline knowledge or perspectives • collaborating with people from disciplines not already routinely collaborated with

Our refinement of one code in particular deserves some additional discussion. Initially, the "values" code was meant to track attention to values understood broadly to include *both* ethical and epistemic values. But we found that some attention to epistemic values is nearly always apparent in participants' discussion of their "uniquely enabled" research. Since coding for omnipresent concepts is largely uninformative, we narrowed the code to address only the rarer instances of attention to ethical values. Because attention to epistemic values is so common in our interview transcripts, many passages that we coded as (ethical) "values" also include consideration of epistemic values and may indicate coupled ethical-epistemic analysis. Still, the "values" code is narrower than coupled ethical-epistemic analysis as the code does not require concurrent consideration of epistemic values or of trade-offs between values across categories.

3.5 Coding transcripts (group-B codes)

Coders 1 and 2 independently applied the final group-B codebook to eight transcripts (coding only those statements previously coded as "research" or "collaboration" during group-A coding). For this double-coding, we began with the same eight transcripts used for the group-B double-coding exercise, though two of those did not contain any statements coded as "research" or "collaboration" and therefore contained no text in need of group-B codes. We replaced these with two more transcripts chosen randomly (uniform distribution) from the remaining transcripts (again excluding the five longest). Inter-rater reliability was quantified as per the group A procedure and is displayed per code and overall in Table S7. Coder 1's coding of these and the remaining transcripts is the basis for quantitative results reported in the main text.

Table S7: Frequency of agreement (f) and inter-rater reliability (κ) for the final group-B codes (Table S6) on the random sample of eight transcripts (28%) that were double coded.

code	f	κ
uncertainty	1.0	1.0
relevance	1.0	1.0
values	.86	.70
disciplines	1.0	1.0
overall	.96	.92

4 Project events considered in Figure 3

Here we describe the events considered by the "project events attended" variable used in Figures 3b and 3e. Each year, SCRiM held two events that brought project members together in-person from across the network: (1) an all-hands project meeting, and (2) an interdisciplinary summer school in climate risk management. The potential explanatory variable "project events attended" counts the number of these events attended by each participant prior to their interview. The interviews took place during the sixth year of SCRiM. At that time, five events of each type had been held, so individuals' attendance counts vary from zero to ten.

All-hands meetings were generally two days long. These meetings were centered around report-outs from the project teams working on twelve sub-projects within SCRiM. There was also time allocated for sub-project team meetings and break-out sessions on cross-cutting topics. All-hands meetings provided an important opportunity for SCRiM project members to learn about sub-projects and research themes that they were not otherwise directly involved with and to deepen connections between their own research and the big-picture vision of SCRiM. For this reason, we consider attendance at all-hands meetings as an *a priori* plausible contributor to participants' overall intellectual integration within the project as assessed by their understanding of the three cross-cutting themes discussed in the main text.

Summer schools were five days long and covered a wide range of interdisciplinary topics in climate risk management research including a significant focus on each of SCRiM's three cross-cutting themes. Summer schools were taught primarily by SCRiM-funded faculty and attended mainly by graduate students and postdocs from outside the project. The summer school was also used as an onboarding activity for new SCRiM postdocs, most of whom attended the summer school during their first year working for the project. Some postdocs returned later as instructors, and many faculty participated as instructors multiple years. Our attendance counts include both participation as a student/attendee or as an instructor.

5 Departmental and disciplinary classifications in Figure 3

See Tables S8 and S9 for the departmental and disciplinary classification schemes used in the coauthorship analysis underlying Figures 3c and 3f. Departmental affiliations were obtained from participants' institutional websites. The disciplinary scheme was created "by hand," and as a result includes some subjectivity both in the choices of where to draw lines between disciplines and in where to place individuals within the categories. The point of the disciplinary classification was not to achieve a perfect or objectively correct classification (which is not possible), but rather to provide a second perspective on the analysis based on departmental affiliation.

Table S8: Departmental classification scheme used in the coauthorship analysis for Figures 3c and 3f. Person counts include only SCRiM participants who coauthored at least one SCRiM-supported publication before the cut-off date used in the publications analysis for Figure 3.

	name of department/unit (across multiple institutions)	# people
1	earth and environmental systems institute	16
2	department of meteorology and atmospheric science	6
3	department of geosciences	5
4	department of statistics	4
5	department of civil and environmental engineering	3
6	department of environmental sciences	3
7	energy and resources group	3
8	(institution x)	3
9	dept. of agricultural economics, sociology, and education	2
10	department of agricultural and resource economics	2
11	department of earth and environment	2
12	department of economics	2
13	applied research laboratory	1
14	department of energy and mineral engineering	1
15	department of environmental and resource economics	1
16	department of philosophy	1
17	sustainable development institute	1

Table S9: Disciplinary classification scheme used in the coauthorship analysis for Figures 3c and 3f. Person counts include only SCRiM participants who coauthored at least one SCRiM-supported publication before the cut-off date used in the publications analysis for Figure 3.

	disciplinary category	# people
1	meteorology/climatology	11
2	economics	10
3	geoscience	8
4	civil/environmental engineering	6
5	policy analysis/decision support	6
6	philosophy	5
7	statistics	4
8	applied math/computer science	3
9	operations research	2
10	sustainable development	1

6 Outlier participants in Figure 3

We removed one participant each from the linear regressions in Figures 3a and 3c and from the associated confidence interval calculations. Here we explain our rationale for those choices and present the results of the same analyses if those participants are instead included.

Figure 3a plots participants' mean understanding (of three SCRiM research themes) against the number of months of funding received by the participant prior to their interview. One participant, the project manager, had a dual role consisting primarily (in terms of funding duration) of management but also some research. Because of this management role, the project manager received nearly four times as many funding months as the next highest faculty participant. And because these months of funding represent a different type of participation compared to other faculty participants, we believe it is better to leave the project manager out of the regression relating project understanding to months of funding. Figure S1a* shows the results of both approaches, and the figure caption includes the associated confidence intervals on the slope of the linear regression lines.

Figure 3c plots participants' mean understanding of project themes against indices of within-project coauthorship across departmental and disciplinary lines. One participant, the principal investigator, scored more than three times as high as the next highest participant on these indices. Because the principal investigator has a unique position with respect to project publications (as the "anchor author" on a very large number of publications) we believe that including the PI in this regression involves a misleading apples-to-oranges comparison. Figure S1c* shows the linear regression lines with and without the PI, and the figure caption includes the associated confidence intervals on the slope of those lines.

7 Coauthorship across institutions in SCRiM

Our analysis of SCRiM-supported publications included an additional component not reported in the main text. We report the findings of this additional component here.

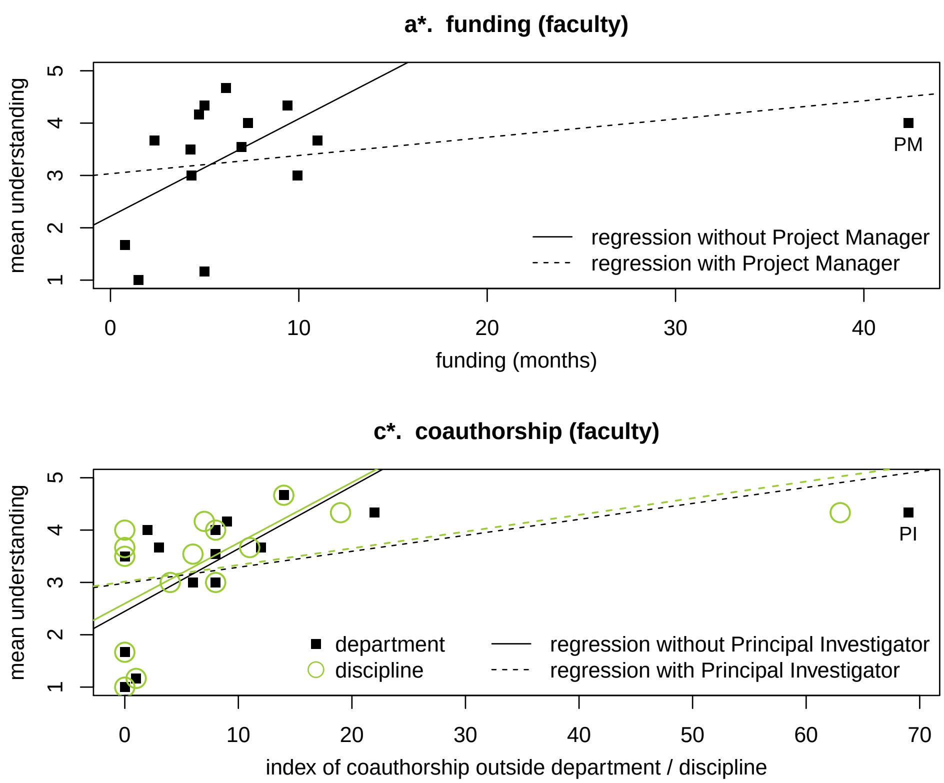

Figure S1: Figures 3a and 3c redrawn with added regression lines and no axis breaks. 90% confidence intervals for slopes of regression lines: (a*) without PM: (.01, .36); with PM: (-.02, .09); (c*) without PI: dept. (.05, .19), disc. (.04, .20); with PI: dept. (0, .06), disc. (0, .06).

The institutional organization of SCRiM followed a "hub and spokes" design, with a proportionally larger number of personnel (as well as the PI) at a central "hub" institution and smaller numbers of personnel at each of seven "spokes" institutions.

Using the database of SCRiM-supported publications discussed in the main text, we examined patterns of coauthorship across the project's eight institutions. Table S10 shows the numbers of publications coauthored across institutions, and Figure S2 summarizes the same data visually. All but two spoke institutions coauthored publications with the hub, but spokes did not (with one exception) coauthor with other spokes. Looking across the spoke institutions, the proportion of publications that include a SCRiM coauthor from another institution (i.e., the fraction of the circle that is filled in) varies widely (compare, e.g., institutions b and c).

Table S10: Publication counts by institution and for cross-institution coauthorship. Letters a–h represent institutions that participated in SCRiM. Numbers on the diagonal report the total number of SCRiM-supported publications coauthored by participants at each institution. Numbers off the diagonal report the number of publications with coauthors from both the row and column institutions. These results include publications through March 2020 (one additional year beyond the cut-off time used in the publications analysis for Figure 3).

	a	b	c	d	e	f	g	h
a	91	1	12	1	1	0	4	0
b	–	24	0	0	0	0	0	0
c	–	–	18	0	0	0	0	0
d	–	–	–	14	0	2	0	0
e	–	–	–	–	10	0	0	0
f	–	–	–	–	–	10	0	0
g	–	–	–	–	–	–	8	0
h	–	–	–	–	–	–	–	1

8 Data availability and protection of study participants

We supply all code used in our analysis and also sufficient data to reproduce our figures. The data files that we provide are "downstream" from antecedent data sources that cannot be fully shared because they contain personal information about our study participants or could be used to identify our participants. Figure S3 shows our full data-analysis process and clarifies what is provided and what is withheld to protect study participants.

References

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement 20*(1), 37–46.

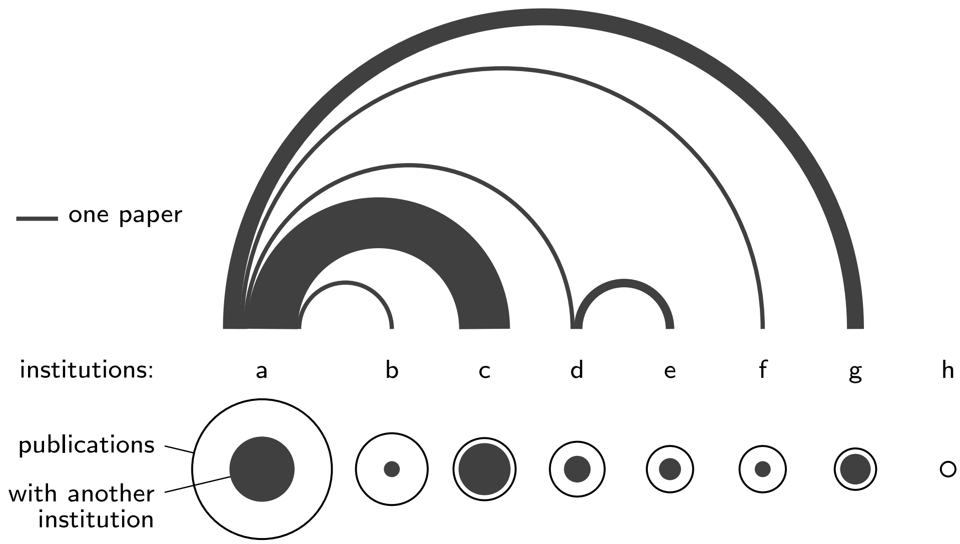

Figure S2: An infographic visualizing the data in Table S10. Letters a–h represent institutions that participated in SCRiM. Circles below the institution labels represent the number of publications with at least one coauthor from that institution (number of pubs. proportional to area). Outer circles include all publications; inner circles (filled) include only those with a coauthor from another SCRiM institution. Arcs show which institutions coauthored with which (number proportional to thickness). No publication had coauthors from more than two SCRiM institutions.

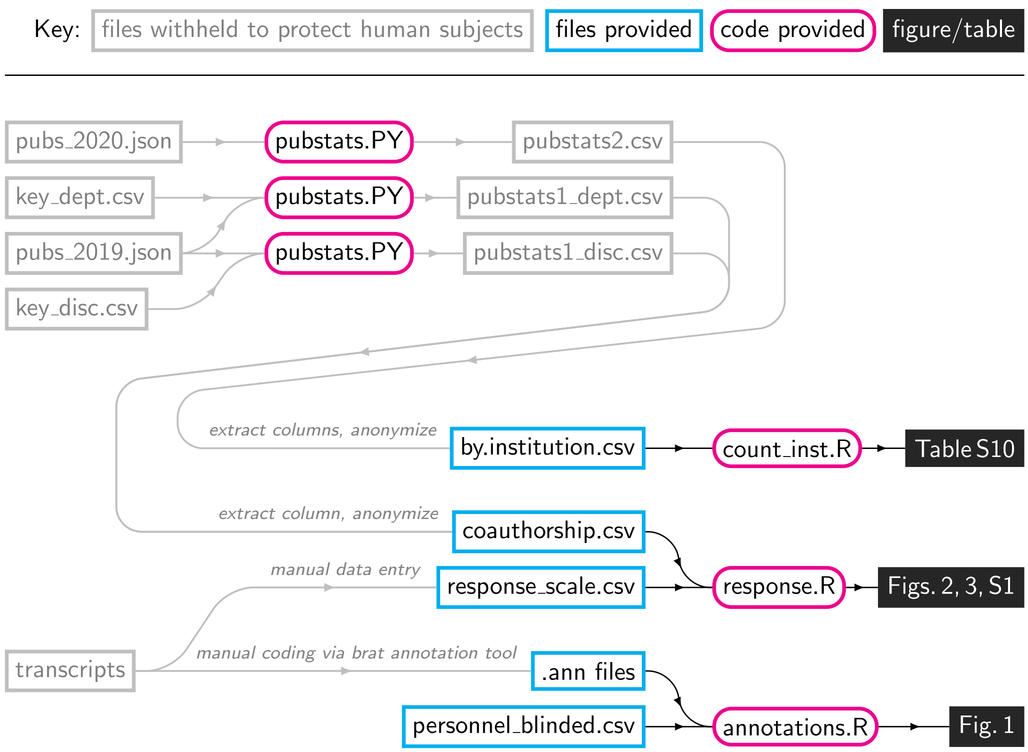

Figure S3: A flow chart of data sources and code used in the analysis. Text along gray lines describes how provided files were generated from withheld files. Interview transcripts are withheld as per the informed consent agreement with study participants. Input and output files used by the pubstats package are withheld because they reveal the identities of our study participants.