

Interdisciplinary collaboration from diverse science teams can produce significant outcomes

Alison Specht^{1¶}, Kevin Crowston^{2¶*}

¹ Terrestrial Ecosystem Research Network, the University of Queensland, Long Pocket Precinct Level 5, Foxtail Building #1019, 80 Meiers Road, Indooroopilly QLD 4068 Australia.

^{2*} School of Information Studies, Syracuse University, NY USA.

Email: crowston@g.syr.edu

Specht A, Crowston K (2022) Interdisciplinary collaboration from diverse science teams can produce significant outcomes. *PLoS ONE* 17(11): e0278043.

<https://doi.org/10.1371/journal.pone.0278043>

¶ These authors contributed equally to this work

Abstract

Scientific teams are increasingly diverse in discipline, international scope and demographics. Diversity has been found to be a driver of innovation but also can be a source of interpersonal friction. Drawing on a mixed-method study of 22 scientific working groups, this paper presents evidence that team diversity has a positive impact on scientific output (i.e., the number of journal papers and citations) through the mediation of the interdisciplinarity of the collaborative process, as evidenced by publishing in and citing more diverse sources. Ironically these factors also seem to be related to lower team member satisfaction and perceived effectiveness, countered by the gender balance of the team. Qualitative data suggests additional factors that facilitate collaboration, such as trust and leadership. Our findings have implications for team design and management, as team diversity seems beneficial, but the process of integration can be difficult and needs management to lead to a productive and innovative process.

keywords: scientific collaboration, interdisciplinary research, diversity, mixed-methods, Input-Mediator-Output (IMO).

Introduction

Complex, global research challenges increasingly require the formation of teams that bring together individuals with diverse skills and perspectives who can work across disciplinary, organisational and national divides [1,2,3,4,5]. One form for these teams is a working group, a group of researchers from different institutions who are brought together and supported to work over a period of a few years on an (often self-identified) compelling research problem. The outcome of the team work may be to

inform policy development, to produce new knowledge and understanding, or the creation of e-infrastructure platforms to support continued advances in science [6]. Such teams should not be mistaken for an established research team given a specific task within their work schedule as part of their employee duties.

Working groups can be assembled in a way that creates several kinds of group diversity. A kind of diversity that is increasingly important to science is disciplinary diversity [7]. In recent years interdisciplinary research has rapidly become the norm [8, 9], that is, research involving scholars from different disciplines collaborating to develop terminology, research approaches, methodologies, or theories that are integrated across multiple disciplines to address broader problems than a single disciplinary approach can address [8]. There is an accepted understanding that interdisciplinary research will foster important ideas beyond the boundaries of a single discipline, even creating new disciplines [7, 11,12, 13]. Further, there is some evidence that increasing the interdisciplinarity of the research team will increase the originality and creativity of the outcomes [7,14, 15]. Certainly an interdisciplinary team is thought likely to be more creative and produce more novel results than conventional research teams due to the variety of disciplines being blended [16, 17, 18, 19, 20, 21].

In addition to disciplinary variety, other kinds of group diversity have impacts on group process and outcomes. Combinations of participants from different kinds of organisations can result in technology fusion, creating new opportunities for each participant [22, 23]. International collaboration in the environmental sciences has often been required due to the nation-boundary-blindness of organisms and processes. In addition, there are often compositional requirements, such as

representative membership, as is the case in most European Union-funded projects, or to comply with societal expectations, such as gender and race representation.

While group diversity can provide the pathway for innovative insights, diverse teams face challenges to collaboration, for example friction, lack of communication and break points, such as those due to differing cultural norms around gender, power, and the preferred level of individual behaviour versus collectivism [24]. Group members in new combinations may find more questions than answers for a particular problem, defeating their purpose. Effective collaboration across disciplinary or national boundaries does not result from simply 'putting people together in a room and shutting the door'. For diverse groups of people to make sense of the vast range of data and information increasingly available and generate an outcome that is innovative and creative, thoughtful group construction and support is required.

The goal of this paper is to identify key aspects of group composition conducive to productive work around multi-disciplinary, multi-layered problems. Is team diversity positively related to outcomes that combine the team's diverse backgrounds and is the productivity of such a team enhanced by the level of diversity within it? If so, what aspects of diversity might be most important? Task-related diversity has been found to be positively related to performance, but bio-demographic diversity has not [25]. Does having a certain gender balance help? Does diversity itself create difficulties of communication, and impede cross-fertilisation across sectors and hence output? How might this work if more than one factor occurs at once?

Theory development

As a basis for our analysis, we develop a model of the factors that may predict the success of interdisciplinary science teams through providing the basis for the development of group collective intelligence [21, 26, 27] and the intellectual fusion or, as some would express it, integration, that is desirable to generate transformative outcomes [28, 29]. We are guided by the Input-Process-Output group framework (Fig. 1). A version of this input-process-output model was used by Stokols et al. [30] in their examination of ‘transdisciplinary scientific collaboration’. In this section, we explain the model and the specific variables we chose for each aspect of the model.

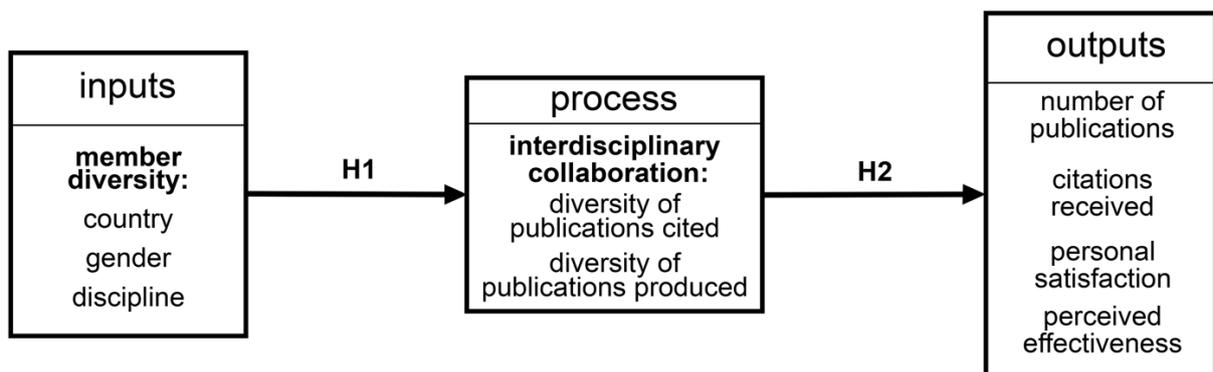


Fig 1. Research model. Inputs, mediators (process and emergent states) and outputs. Hypothesis numbers are shown.

In this model, inputs are the attributes that the team members bring to the teamwork. In our setting, we are focusing on the members themselves and their demographic backgrounds. A working group is composed in response to a problem: people are deliberately chosen with relevant skills and expertise to contribute to a solution, while keeping in mind other factors, such as organisation, country, and life situation. These choices create groups with different levels of diversity.

Outputs of a group are the measurable consequences of team function. For research groups the desired outcomes are new knowledge and understanding expressed usually through refereed publications, deposited data in open data repositories, and new, long-term collaborations across disciplinary, organisation and geographical boundaries. Lynch et al. [31], show a range of types of valued outputs from synthesis centers. As noted also by Hackman [32, p. 128], team outputs can extend beyond specific products to include participant satisfaction with their work and their perceptions of the effectiveness of the team.

Between inputs and outputs is the process that mediates the influence of inputs on outputs. The measurable and intended process in our study is that of collaboration across disciplinary boundaries. Interdisciplinary or trans-disciplinary group membership is often used to support cross-fertilisation to “co-design research, co-produce solution-oriented knowledge, and reintegrate the knowledge into strategies for problem-solving and the development of new scientific insights” [33]. The details of these interactions are subtle and difficult to capture in a comparable way across teams. Research has therefore usually adopted proxy measures. In this paper we assess interdisciplinary research output by the diversity of the authors and of the references they cite [34, 35].

Hypothesis development

Much of the research that has addressed the question of the effect of team composition on outputs has examined diversity directly without postulating an intervening process. Publication output, for example, has been positively related not only to disciplinary diversity but also low seniority of teams [36]. Others have found a

positive effect of diversity of career stages and gender on publication output, but a negative effect of educational (disciplinary) diversity [37]. Gender heterogeneity in a group has been shown to enhance citation rates [38] and link positively with team performance [39]. The novelty of our approach is that rather than hypothesizing direct impacts of diversity on outputs, we expect outcomes to be moderated by the nature of the research process. We consider that different kinds of diversity may contribute in different ways and so develop a set of hypotheses for each of these relationships.

Inputs to process

We first consider how group diversity (the input) may have an impact on the group process, specifically the interdisciplinarity of the collaboration (the process).

Although men and women share common predictors for collaboration, women have been found to tend to engage in more interdisciplinary research collaborations [40], to have more collaborators than men [41] and they have characteristically more social sensitivity [26, 32]. Indeed, in studies of small groups (7-10 members) the collective intelligence of the group has been positively correlated with the percentage of women in a group [26], as has group effectiveness and innovation, particularly when task intensity is high [41, 43]. This result has been explained by improvements in the emotional intelligence of the group and a reduction in conflicts with the addition of a component of women [39, 42]. We therefore hypothesise that interdisciplinary collaboration will be enhanced by the proportion of female participants (Table 1, H1(a)).

Without a range of disciplines present in a group—as assessed by a classification of the group members into their ‘home’ disciplines [2]—the meshing of disciplines could not occur. It has been found that the diversity of citations tends to grow with the diversity of disciplines of the authors [34]. We therefore hypothesise that interdisciplinary collaboration will be enhanced by the interdisciplinarity of the group (Table 1, H1(b)).

Finally, intra-country (regional, state) or international membership may also be required for project performance, depending on the problem being considered or the funding arrangements. In environmental and biodiversity science and management, global participation in some shape or form is essential not the least because organisms do not abide by human political boundaries, but also (i) to get a comprehensive grasp on a problem, (ii) to acquire the requisite data, and (iii) to access appropriate expertise and data. Indeed, international collaboration has become the norm rather than the exception in modern science [3, 35]. Group membership may be mandated or perceived as desirable for funding reasons (e.g. funding from the European Union or the Belmont Forum) along with a perception that the results will be enhanced by such inclusivity.

While there may be a clear need or desirability for international collaboration, there are clear challenges to successful collaboration. The physical dispersion of group members creates logistic problems due to travel approvals and cost, and if collaboration is to be remote, time-zone challenges can be prohibitive [1, 17].

Participants from different countries with different languages and ways of working can have communication and cultural challenges which take time to overcome, reducing the productivity of the group. This can result in less novel outcomes, and

high transaction costs and communication barriers; for instance, data suggest that the more nations involved in a project, the fewer publications and the more conventional the research [3]. We therefore hypothesise that interdisciplinary collaboration will be negatively affected by the diversity of international membership (Table 1, H1(c)).

Process to outputs

We next consider how the process affects the outputs, in the first instance, the number of publications produced. Collaboration in general has been positively related to higher output [44, 45]. Researchers engaging in interdisciplinary research have been found to be less productive but have higher impact [46]. In contrast, authors who publish at moderate levels in disciplinarily diverse journals and with a moderate level of collaborator diversity have been found to publish more, while a moderate level of collaborator diversity is beneficial for authors who are more focussed on output [47]. We draw on these two streams of literature to propose that there is a positive correlation between interdisciplinary collaboration in a group and the number of publications produced (Table 1, H2(a)).

Our second output measure is the impact of the work as indicated by citation rates. Several articles have examined the effect of team diversity on the citations received (e.g., [48, 49, 50]). We conceptualize these factors as having an impact via their effect on the interdisciplinarity of the team process. The effects of interdisciplinarity have mostly been studied at the level of individual articles rather than teams, generally finding a positive impact of citing more diversely (e.g., [51, 52, 53]). In contrast, at the individual researcher and team level, the effect has been found to be curvilinear, with an inverted U-shape, meaning that impact is highest for intermediate

levels of interdisciplinary [54, 55]. The effect has been attributed to the costs of learning to work across disciplines, where too much interdisciplinarity can become a distraction (Table 1, H2(b)).

In the previous two hypotheses, we have argued that teams with an interdisciplinary process will be more successful in publishing. We expect this success to be reflected in the evaluations of the team members of group function. For example, that there will be a positive relationship between the interdisciplinarity of the collaboration in a group and member satisfaction with the group (Table 1, H2(c)), and the perceived effectiveness of the group (Table 1, H2(d)).

Table 1. Hypotheses posed in this paper.

| Hypothesis | Proposed relationship |
|-------------------|--|
| H1 | <p>There is a positive relationship between interdisciplinary collaboration and:</p> <ul style="list-style-type: none"> (a) proportion of female members (b) the interdisciplinarity of the group (c) There is a negative relationship between interdisciplinary collaboration and the diversity of international membership |
| H2 | <ul style="list-style-type: none"> (a) There is a positive correlation between interdisciplinary collaboration in a group and number of publications. (b) There is an inverted-U relationship between interdisciplinary collaboration in a group and impact as measured by the median number of citations. (c) There is a positive correlation between interdisciplinary collaboration in a group and personal satisfaction (d) There is a positive correlation between interdisciplinary collaboration in a group and perceived effectiveness |

Methods

The project follows a mixed methods approach using triangulation of quantitative results with qualitative survey information [56]. The quantitative study examines team demographics, processes and outcomes while the qualitative component examines the opinions proffered by team members about the reasons for group success or otherwise.

Before commencing the detailed description of the data acquisition for the reported project, the limitations related to the ethics statements made to the participants and accepted by the Institutional Review Board of the corresponding author's university and complying sufficiently with the Australian ethics standards were as follows:

“Personal information that is collected will be used solely to enable network analysis of members within a working group and will not be used for any other purpose.

Results of the research may at some future time be published. Although your responses may be identifiable to the researchers, responses will be kept confidential and no individual responses will be reported; only summarized findings will be reported.” and “Your identity will be held in confidence as an invitee to the survey associated with the relevant synthesis centre and group. Your identity will not be published.” This being so, we have anonymised all respondent identities, including identifiable links to their organisations and disciplines, and to the specific publications by each group (which would allow identification of the authors and hence the groups being studied). We have not published here or in the data repository for the paper the full demographic profiles of members.

Research setting

We test these hypotheses through studies of a selection of research working groups designed to have some degree of diversity, with members with a variety of disciplines, skillsets and origins. The research working groups we are focussing on are a facilitated version of the 'self-assembly' to which Twyman and Contractor [57] refer and which are well-discussed in the synthesis center literature [31, 58, 16, 50]. These groups gather, with logistic support, to analyse a specific complex problem, drawing on members from many institutions and countries for intermittent but concentrated periods of time apart from their normal workplace commitments. The staccato nature of the short but concentrated time the groups spend with each other in these 'hot moments' is often termed 'island time' (Eric Garnier pers. comm. in [16]) and has been found to be particularly conducive to creativity [27, 58]. The members of such groups are largely 'volunteers', driven by a common interest in the problem being addressed. They are supported to a limited extent in their endeavours, but not salaried to solve the problem at hand (Figure 2, next page).

The groups we studied were drawn from a large cyberinfrastructure project (DataONE) and from two synthesis centers in the biodiversity and environmental fields. The DataONE project adopted the synthesis working group approach in its participatory engagement model, enabling it to tap into expertise in the wider community as it constructed its e-infrastructure [59].

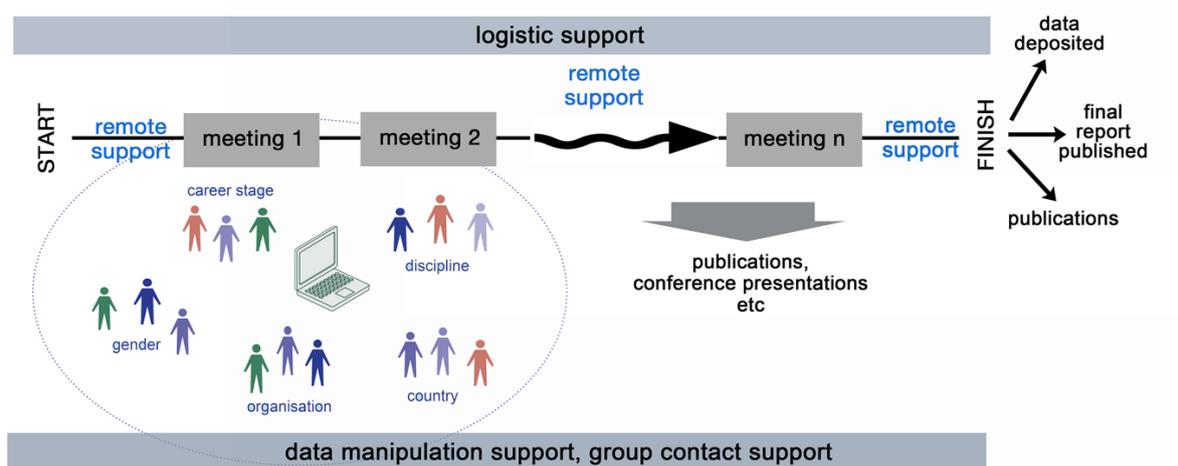


Fig. 2. Working group workflow. A depiction of the way the working groups work together in our scenarios, where groups of diverse people meet together intermittently over a period of time in a supported environment to examine difficult, multi-disciplinary, problems. The meetings might each last up to a week, and are often at six month to yearly intervals.

Synthesis centers are purpose-built organisations designed to enable diverse groups of people (working groups) to synthesise disparate data and information on a particular topic to produce new understanding [18, 50, 60, 61]. They have been likened to the business ‘incubator’ [62, 63]. Synthesis centers support the use of existing information and apply, to greater or lesser extent, many of the elements of team science, open science and data-intensive science [5, 16, 31, 63]. They have been highly successful in facilitating collaborations and are highly productive [16, 18, 50, 61, 64].

The composition of synthesis center groups is largely determined by group leaders (Principal Investigators) within criteria set by the centers, such as a degree of inter- (or less often, trans-) disciplinarity, multi-national and multi-sectoral teams, career and gender balance ([52] and as described on the center web sites accessible through www.synthesis-consortium.org). This approach means that a large

proportion of the group starts with agreement on a common goal. The working groups in our study were supported for a maximum of four years with multiple meetings over that time. For the synthesis centers and for DataONE, good outcomes are critical to continued funding and the organisations post the products of the groups they sponsor on their web sites for public access.

It is relevant to note that the work of these groups is collegiate, as far as communication and the development of trust will allow. Decisions are made as a group, for example, once the groups have examined the data and information to hand, they will determine their work plan, identifying the products (usually articles but some code and conference presentations) that would seem most achievable and valuable, and then separate the tasks among the group members according to the skill sets needed. Occasionally needed skills were not available within the group and additional experts were brought in to complete the specific task. As one survey respondent put it, "The group was very effective at dividing tasks, assigning roles, and then getting their parts done. Everyone wanted to contribute." (C-1, R2). If a group member did not pull their weight they could be isolated. The group leader(s) keep track of these tasks and all group members will revisit their various tasks at the start of each meeting.

By confining our study to this cohort we were able to compare across three similarly-aged organisations all providing similar logistic and infrastructure support (an important factor influencing creativity of output according to [15]). The selection of the specific groups for study within each Center or project were determined by (a) advice of the director of the Center (JWP), (b) maturity of the group, and (c) response level to the qualitative survey (described below).

It should be noted that although our ambitions were to include another center and hence suite of groups in our study, the response from the center approached, although positive, was tardy (one year after initial enquiry) and did not fit our time-frame, especially given we already had results from 2014. The size of the sample we analyze in this paper restricts the analysis possibilities.

Data acquisition

We collected three types of data for the model:

1. demographic profile of group members (inputs: Fig 1)
2. collation and analysis of articles used and produced (process and outputs: Fig 1)
3. group member perceptions (outputs: Fig 1)

For items 1 and 2 the data were available through the group websites. Item 3 was accomplished by a survey sent to the members as published on the group websites.

Working group demographics

The demographic profile of all members of each selected Working Group was collated. This included their country of origin and their gender (binary only; no members with non-binary identities were identified).

The primary scientific discipline of each group member was defined using their own statements to the contributing organisations, and where not available, their self-stated fields on sources such as ResearchGate and Google Scholar, and their pages on their organisation's web sites. An initial list of forty-six primary disciplines was created using the Australian and New Zealand Fields of Research [65] to derive a

smaller controlled vocabulary of 22 categories. This reduced vocabulary included biology and applied biology, chemistry, climate science, communication, computing, data science, earth sciences, ecology, education, engineering, environmental science, evolution studies, freshwater (and marine) studies, geography, health, hydrology, library and information studies, modelling, policy, sociology, and statistics. As mentioned, for confidentiality reasons these discipline fields are not listed to enable linkage with the participant or the group to which the member belongs.

Collation of articles used and produced

A key construct in our model is the interdisciplinarity of the group process. As we were unable to follow each group individually over the several years they were collaborating, we sought a measure that could be applied retrospectively.

Interdisciplinarity can be assessed in many ways [5], of which bibliometric measures are the most developed and have valuable antecedent analyses [e.g. 66, 67, 68, 69, 70]. We used journals as an indication of discipline. Journals are responsive to the communities they serve (they have to have an 'audience' to survive), and several disciplinary themes or 'keywords' are used to describe the journal field, thus encouraging relevant articles and providing appropriately qualified editors and reviewers to assess the submissions. In this way the selection of journals in which to publish is tailored to the subject matter. Of course, some journals are quite general and welcome a wide range of articles, which clouds the specificity of the domain description, although in our case these occur rarely in our study due to their high impact (e.g., Nature and Science). Other measures such as data deposition, grant success and conference presentations are also useful indicators of productivity as well as measurable for disciplinary variety, but they were inconsistently measured

across our study groups, so we have confined ourselves to refereed journal articles produced by the groups.

We based our analysis on the groups' publications. We obtained copies of all papers published by the groups up until August 2020, the vast majority of which were open access. The number of citations to each publication was acquired from CrossREF on 19 August 2020. Not all articles produced by the groups were in journals tracked fully by CrossREF, as CrossREF mainly scans English-language journals, but this is a relatively minor impediment to our analysis as English is the norm for the fields we are studying.

To assess the interdisciplinarity of the working groups' collaborations, we analysed the literature on which the group drew to produce their published articles, the 'inspiration' measure of [9] and used by several authors in whole or in part as a measure of interdisciplinarity [34, 71, 72, 73]. Specifically, we examined the disciplinary diversity of the cited journals [10, 74]. We selected this measure based on the premise that groups working in an interdisciplinary way bring together diverse knowledge [2, 75, 76], which will be reflected in the diversity of the literature cited. To consider the inter-or trans-disciplinary nature of the outputs from a group rather than just their productivity, we also examined the degree to which the outputs appear in journals in a diversity of disciplines [10, 50].

To obtain these measures, we first extracted which journals were cited in the published papers, as well as the journals in which the groups published their own articles. A total of 1007 journals were identified in this process. Each journal was classified according to disciplinary category(ies) using SCOPUS, SCIMAGO and journal-stated discipline fields if not listed in those databases, being generous rather

than reductionist in allocation. In this manner 128 specific journal disciplines were identified across all working groups, from agriculture to ecology to parasitology and water science. Guided by the Australia and New Zealand Fields of Research categorisation (ANZFoR [65]), these specific disciplinary categories were then clustered under 22 ANZFoR disciplinary divisions, with two non-ANZFoR categories created due to no suitable matches in the ANZFoR Divisions. These discipline categories for journals, although they are responsive to the communities they serve, are admittedly relatively general. The disciplinary divisions used were: agriculture, veterinary and food sciences, biological sciences, biomedical and clinical sciences, built environment and design, commerce, management and tourism services, earth sciences, economics, education, engineering, environmental sciences, health sciences, history, heritage and archaeology, human society, information and computing sciences, language, communication and culture, law and legal studies, mathematical sciences, philosophy and religious studies, physical sciences, and psychology. The two additional divisions were 'bio water science' which covered all freshwater and marine biological sciences, and 'general science' which enabled the classification of journals like Science, and Nature that cover multiple disciplines.

Group member perceptions

A third source of information came from a survey of Working Group members who were asked via electronic survey about their perceptions of group performance (satisfaction and perceived effectiveness on a 1–5 point scale, from not at all effective/satisfied to very effective/satisfied, with an invitation to make comments explaining their choice). The initial intent of these questions was to combine them with productivity as a multi-item measure of group success [32], but as the

correlations were not high, we instead analyzed them individually. We also invited open-ended comments from subjects about the 'primary factors that they felt contributed to their working group's effectiveness or lack thereof', for a total of five questions. These questions were part of a longer survey including questions not related to interdisciplinarity that we do not analyse in this paper. As there were many other components in the survey, we considered that having too many items for these scales would negatively impact the return rate. Links to the entire questionnaire were emailed first to the DataONE email distribution list during an All Hands meeting in October 2014 (including to members who were not in attendance at the meeting) at the end of Phase 1 of DataONE. The same survey was subsequently emailed to Working Group members of the two synthesis centers in late 2018. Only groups that had finished or were close to finishing their work were invited to participate and there was little change in group composition at the times of the surveys (2014 and 2018) in relation to when the groups started (and finished). In all cases, 2 weeks was given for return, with an extension of another week. We only included groups in our analysis from which we received survey responses from at least 20% of members.

Analysis

Operationalization of concepts

In this section we explain how we used the collected data to measure our research constructs. First, we computed measures of group diversity along the different demographics. For each group, we started with counts of members in different demographic categories (e.g., number of members from different countries). For

country and discipline, we used the category counts to compute a measure of diversity (entropy) using the Shannon index [50, 77, 78, 79]:

$$\text{Diversity or entropy: } H = - \sum (p_i \ln p_i) \quad (\text{equation 1})$$

where p_i is the proportion of members in group i .

Since the data collected on gender were binary, we simply used the proportion of female members in the group. In our dataset the range of proportion female was 4-67%, meaning we had some nearly all-male groups but no all-female groups.

To assess the extent of the interdisciplinarity of the collaboration process, we measured the diversity of the disciplines of the journals in which the groups published and those they cited in their publications (computed using equation 1 above). For this purpose, we combined the count of journals cited per discipline across all the working group's publications. Diversity can, of course, be assessed in different ways, such as variety, balance and disparity [79]. In a similar manner to assessment of the variety of demographics, we used Shannon diversity as our measure for disciplinary diversity. Although this measures variety it is also affected by balance (lower balance leads to lower diversity). We also note that some studies of interdisciplinarity have assessed not just the variety of disciplines cited but also the atypicality of the combination (e.g., [80]), however, since our classification of journals is based on a different system than in Uzzi [80], we do not have data on typicality.

We measured the output of the group in four ways, first by the number of publications and the impact of the group's work by computing the median number of citations the various publications attracted. We assessed satisfaction with group performance and perceived group effectiveness at an individual level through a

survey of members. We used the mean of the individual scores as the group measure.

In addition to the variables in the model, we included two control variables: (i) the age of the group, with the assumption that groups that have been around longer have had more of an opportunity to publish and to have an impact (as measured by citations), and (ii) the Center (a three-level factor variable), to control for different expectations around publishing in the three settings.

Hypothesis testing

Hypotheses were tested using regression on the data emerging from the demographics and diversity measures. Data other than counts were standardized before regression. One problem arose in carrying out the regressions: as we had only 22 groups, using too many variables in the regressions led to overfitting.

Unfortunately, the small number of data points also meant that we could not test the hypotheses simultaneously, e.g., with a structural equation model.

The problem of potential overfitting arose for the analyses of the demographic input variables. A regression that included all the demographic variables achieved nearly perfect R^2 , an indication that the model was overfit. To avoid this problem, we used a stepwise regression approach, adding variables that were most related to the outcome, but stopping with a small number of variables. We explored reducing the dimensionality of the input variables through factor analysis but did not find a satisfactory solution with a smaller number of factors. We also explored more modern techniques for variable selection such as lasso but did not have enough data to use them.

Where we hypothesised curvilinear relationships, we entered variables both as a linear and a squared term using the R poly function, which computes orthogonal polynomials to avoid multicollinearity. As the use of this function complicates interpretation of the regression coefficients, we present non-linear relationships graphically.

Qualitative analysis

The open-ended responses were subject to thematic analysis related to the research question (“In your opinion, what are the primary factors that contributed to your working group's effectiveness or lack thereof?”) using an inductive semantic approach [81]. The themes thus identified alongside their associated (anonymised) comments were sent to the groups for validation of our interpretation and modified if required. In this process additional insight was often obtained.

Results

Group analysis

Demographic data

Demographic profiles were obtained for the groups that responded to the on-line surveys in sufficient numbers (>20% responses). The resulting population of 389 people came from 28 countries, the majority from the USA (62%), and the next highest from France (13%), the home of the third organisation in the study. The UK and Canada were also relatively well represented. The total population was predominantly male (68% male and 32% female) and 51% of members were from

universities, 19% were from government organisations, and 15% from research organisations.

Ecology was by far the most common primary discipline type (23%), with an equal proportion of people in the computing, data science, statistics and modelling areas. These distributions reflect the focus of the sponsoring organisations and the emphasis of the working groups on working with data. Freshwater biology and ecology were clustered together into freshwater science, which contributed 4.6% of the total.

Group size was variable, ranging from 11 members up to a maximum of 28 (S1 Appendix). The demographics differed between groups, with some groups very international, and others exclusively from one country. The proportion of females in the groups ranged from 0% to 67%, an average of $33\% \pm 4\%$. While there are clear differences among groups, overall the mix of respondents seems representative of scientists who participate in working groups [18, 28, 42] and was consistent with the first author's personal experience coordinating working groups. Groups that had at the time of the survey tended not to respond to the survey (nor to have output to analyse), and so are not included in the study. Nevertheless, we have a range of productivity and impact in the groups included.

Meeting attendance data (not that that is the only measure of fidelity to the group) was not available consistently for all groups, but for the data that were available, fidelity to the group was between 60% and 80% over 2-10 meetings. Fidelity was always greater than 80% for groups that had only two meetings.

Publication analysis

One hundred and fifty-seven journal articles were collectively produced by the groups at the time of our study (August 2020), and 6,749 articles were cited in these articles. The articles appeared in 83 different journals in 22 ANZFoR disciplinary divisions. The articles cited (the 'inspiration' of Gates et al. [9]) were drawn from journals in 18 ANZFoR disciplinary divisions.

Group A-1 had the lowest publication diversity of all the groups at 0.693 (S2 Appendix) with only two publications at the time of the study, one in the ANZFoR Division biological sciences and one in environmental sciences. The citation diversity was similarly low (0.287, S2 Appendix). The articles cited across these two papers were from forty-six journals, with thirty-two journals cited in one paper and seventeen cited in the other. The vast majority of the citations (91% and 93%) were to publications in the ANZFoR Division biological sciences.

In contrast, B-7, with eighteen publications at the time of the study, had the highest publication diversity at 1.874 (S2 Appendix) with publications across seven ANZFoR divisions (agriculture, veterinary and food sciences, biological sciences, earth sciences, environmental sciences, general science, human society, and information and computing sciences). To produce these articles, the group drew on papers from seventeen ANZFoR Divisions (agriculture, veterinary and food sciences, bio water science, biological sciences, biomedical and clinical sciences, built environment and design, commerce, management and tourism services, earth sciences, engineering, environmental sciences, general science, history, heritage and archaeology, human society, information and computing sciences, language, communication and culture,

philosophy and religious studies, physical sciences, and psychology). The citation diversity (and its distribution amongst the papers), however, was not the highest of the groups at 1.301 (S2 Appendix). Group C-4 had the highest citation diversity with an index of 1.656 drawing on articles from ten ANZFoR Divisions (agriculture, veterinary and food sciences, bio water science, biological sciences, earth sciences, engineering, environmental sciences, general science, information and computing sciences, mathematical sciences, and psychology) across their three articles.

Citations per article ranged from a high of ninety-six different journals for paper C of group A-6, to no, or one, citation per article (e.g. B-6 paper C and B-3 paper F), while many groups cited forty or fifty articles per paper (e.g. C-1 paper A, A-9 paper F, B-1 paper B).

The average number of authors per article was 8.75 ± 0.51 . This number alone is greater than the average number of authors for all journal discipline fields found in past studies of scientific collaboration. For example, a maximum of 5.9 authors per article for the environmental sciences was reported in Patience et al. [82] and 5.19 for synthesis centers by Hackett et al. [50]). This evidence suggests the working groups in our study were highly collaborative. The average number of authors per article was highest in group A-4 closely followed by C-1 and A-9, while the minimum average number of authors was 1.7 for group B-6 (S2 Appendix).

The bibliographic data are available in the repository of the Environmental Data Initiative [83]

Qualitative feedback

Response rates to the surveys per group ranged from 56% to a minimum of 20% (our cut-off for including the group in the study). Of the ninety-two respondents who provided feedback to the ranked questions about group effectiveness and satisfaction, only forty-eight provided open-ended responses about effectiveness, and twenty-two, about satisfaction. One hundred and twelve responses were received to the general request for comment across all three survey groups. Detailed commentary on the open-ended responses is included in the qualitative analysis.

Respondents' perception of their group's effectiveness and satisfaction was positive (all measures were above 2.5 on a 5-point scale, S3 Appendix). However, some groups (A-7 and A-8, for example, and B-5) were as low as 3 on the 5-point scale, while others (A-4, B-6 and C-2) thought their groups were particularly effective and were correspondingly satisfied with the group's function. There was not, however, a linear relationship between the two metrics, so we treated them as separate entities.

Hypothesis testing

In this section, we describe the results of testing the proposed hypotheses.

Prediction of interdisciplinary collaboration (H1)

To assess the level of interdisciplinary collaboration in the groups we used the diversity of the disciplinary fields of the references cited by the groups, and the diversity of journals chosen by the groups for their publications. We used stepwise regression to identify influential variables for each dependent variable separately.

The disciplinary diversity of the journals cited by the groups in the production of their articles (cited publication diversity) was positively related to the disciplinary diversity and the proportion of women in the group ($p < 0.01$; Table 2). The diversity of the journals in which the groups published was negatively related to the county diversity of the group ($p < 0.05$; Table 2) and positively related to discipline diversity ($p < 0.05$), but other factors were not selected by the regression.

Table 2. The effect of input variables on citation and publication diversity.

| | Diversity of publications cited | Publication diversity |
|------------------------------|--|------------------------------|
| Intercept | 0.000 (0.122) | 0.000 (0.165) |
| Country diversity | | -0.403* (0.175) |
| Proportion of female members | 0.644*** (0.130) | |
| Discipline diversity | 0.391** (0.130) | 0.445* (0.175) |
| Number of observations | 22 | 22 |
| adjusted R ² | 0.669 | 0.401 |

Mean, standard error (in parentheses) and significance.

+ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Prediction of output (H2)

As the outcome variable was a count, a Poisson regression was used to assess effectors of the number of publications produced. On the other hand, the median number of citations received was over-dispersed (the standard deviation was greater than the mean), with a few clear outliers, so a negative binomial regression was used instead.

The disciplinary diversity of the journals in which the groups published predicted the number of publications ($p < 0.01$, Table 3). The disciplinary diversity of the articles

cited in producing those publications, our measure of interdisciplinary collaboration, also positively predicted the number of publications ($p < 0.05$, Table 3). As hypothesised, the median number of citations received had inverted-U relationships with both measures of interdisciplinary collaboration, as shown in Fig 3. Among the controls, there was a strong linear positive relationship between group age and the number of papers produced ($p < 0.001$; Table 3) and differences were noted among the centers. Specifically, the B groups had significantly fewer publications but more citations than the A groups, while the C groups were not significantly different (note that the choice of the A groups as the baseline to which B and C are compared is arbitrary).

Table 3. Factors predicting output variables.

| | Number of publications | Median number of cites to |
|--|-------------------------------|----------------------------------|
| Intercept | 2.105*** (0.160) | 2.351*** (0.241) |
| Project age | 0.818*** (0.156) | 0.091 (0.220) |
| A groups | – | – |
| B groups | –1.381*** (0.342) | 1.854*** (0.473) |
| C groups | 0.329 (0.287) | 0.623 (0.451) |
| Publication discipline diversity | 0.257** (0.099) | 0.608 (0.767) |
| Publication discipline diversity squared | | –2.434*** (0.636) |
| Cited discipline diversity | 0.372* (0.152) | –1.788* (0.840) |
| Cited discipline diversity squared | | –1.627+ (0.831) |
| Number of observations | 22 | 22 |
| Nagelkerke's pseudo R ² | 0.888 | 0.940 |

Regression weight, standard error (in parentheses) and significance of each variable is shown. “A” groups were used as the baseline for the group factor.

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

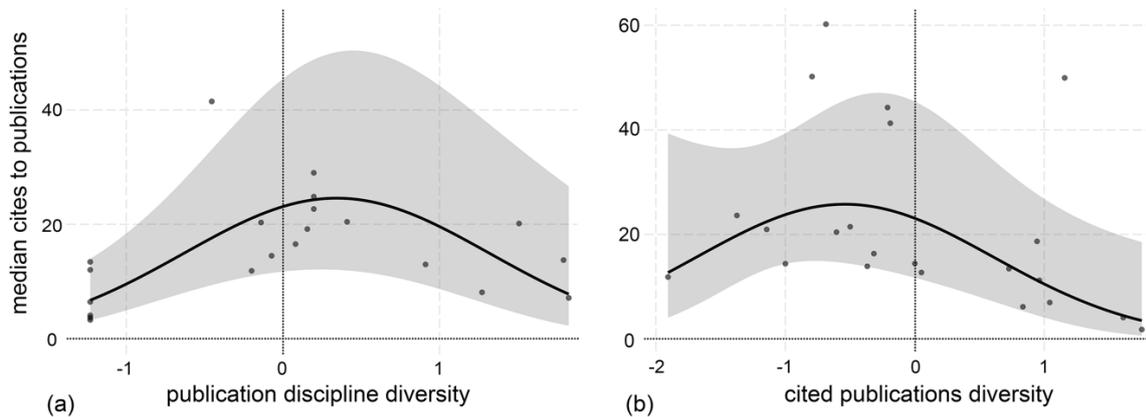


Fig 3. Influences on citations received. The median number of citations received by a group’s publications against (a) publication discipline and (b) the diversity of publications cited. One group is an outlier to the plot in 4(a) and is not visible in the plot. The 95% confidence envelope around the trend is shown. ‘0’ on the axes is the grand mean for all groups for that variable, and the intervals shown are one standard deviation from the mean. The data points show the partial residuals for the groups, i.e., the residual after controlling for the other variables in the regression.

Perception of satisfaction and effectiveness

A model predicting self-reported satisfaction and perceived group effectiveness based only on process and emergent state variables (the hypothesised model) did not reveal any significant predictors. We therefore considered whether these outcomes might be affected also by group composition. To test this *post-hoc* hypothesis, we used stepwise regression to identify independent variables. We also added the number of publications (log transformed because of skew) as a predictor. Satisfaction with the group was negatively related to the age of the group ($p < 0.05$; Table 4) and, as might be expected, positively related to the number of publications the group produced ($p < 0.05$; Table 4). Satisfaction was negatively related to the diversity of publications cited in the production of those papers ($p < 0.05$; Table 4), negatively related to the diversity of countries of the group members ($p < 0.05$; Table

4), and positively related to the proportion of female members in the group ($p < 0.05$; Table 4).

Predictors of perceptions of the effectiveness of the group were more limited, namely that the more diverse the publications sourced by the group the less effective the group was perceived to be ($p < 0.001$; Table 4), the higher disciplinary diversity of the group, the more effective respondents felt the group was ($p < 0.05$; Table 4), and there was a moderate positive influence of the proportion of female members ($p < 0.05$; Table 4).

Table 4. Predictors of satisfaction and perceived group effectiveness.

| | Perceived satisfaction | Perceived effectiveness |
|----------------------------------|-------------------------------|--------------------------------|
| Intercept | 0.000 (0.164) | 0.017 (0.159) |
| Project age | -0.572* (0.192) | |
| In (number of publications) | 0.462* (0.209) | 0.238 (0.168) |
| Publication discipline diversity | -0.404 (0.255) | 0.376 (0.213) |
| Cited discipline diversity | -0.817* (0.313) | -0.512* (0.199) |
| Country diversity | -0.626* (0.242) | -0.260 (0.176) |
| Proportion female | 0.624* (0.289) | |
| Disciplinary diversity | 0.316 (0.247) | |
| Number of Observations | 22 | 22 |
| Adjusted R ² | 0.422 | 0.169 |

Mean, standard error (in parentheses) and significance of each variable shown.

+ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$

Summary

All hypotheses were supported except for Hypotheses 2(c) and 2(d) (Table 5).

Country of origin negatively affected the process of publishing diversely and personal satisfaction with the group (Fig 4). Citing a diverse range of articles in producing output was correlated with lower satisfaction and perceived effectiveness, despite not actually affecting measurable output, which had a strong positive effect on personal satisfaction with the group (Fig 4). The proportion of female members had a positive effect on the diversity of articles cited and on personal satisfaction with the group (Fig 4). The diversity of disciplines represented in the groups had a strong positive effect on both metrics used for interdisciplinary collaboration and had a less positive effect on the perceived effectiveness of the group (Fig 4).

Table 5. Summary of the outcome of the hypothesis tests.

| Hypothesis | Proposed relationship | Comment |
|------------|--|--|
| H1 | <p>There is a positive relationship between interdisciplinary collaboration and:</p> <p>(a) measures of gender diversity</p> <p>(b) the interdisciplinarity of the group</p> <p>(c) There is a negative relationship between interdisciplinary collaboration and the diversity of international membership</p> | <p>The more female members in the group, the greater the cited discipline diversity (Table 2).</p> <p>The more discipline diversity in a group, the greater the cited discipline and publication diversity (Table 2)</p> <p>There is a negative relationship between publication diversity and the diversity of countries involved (Table 2)</p> |
| H2 | (a) There is a positive correlation between interdisciplinary | There is a positive relationship between the number of publications and their |

| | |
|---|--|
| <p>collaboration in a group and number of publications.</p> <p>(b) There is an inverted-U relationship between interdisciplinary collaboration in a group and impact as measured by the median number of citations.</p> <p>(c) There is a positive correlation between interdisciplinary collaboration in a group and personal satisfaction</p> <p>(d) There is a positive correlation between interdisciplinary collaboration in a group and perceived effectiveness</p> | <p>disciplinary diversity and the diversity of the articles cited (Table 3).</p> <p>Median number of citations received by a group's publications is highest for intermediate values of publication and citation diversity (Table 3 and Fig 2)</p> <p>Respondents' satisfaction with the group was <u>positively</u> related to the number of publications the group produced and the proportion of female members in the group, but <u>negatively</u> related to the diversity of the cited publications and the diversity of countries (Table 4).</p> <p>Perceived group effectiveness was <u>negatively</u> related to the diversity of the cited publications. It was, however, <u>positively</u> related to the disciplinary diversity of the group (Table 4)</p> |
|---|--|

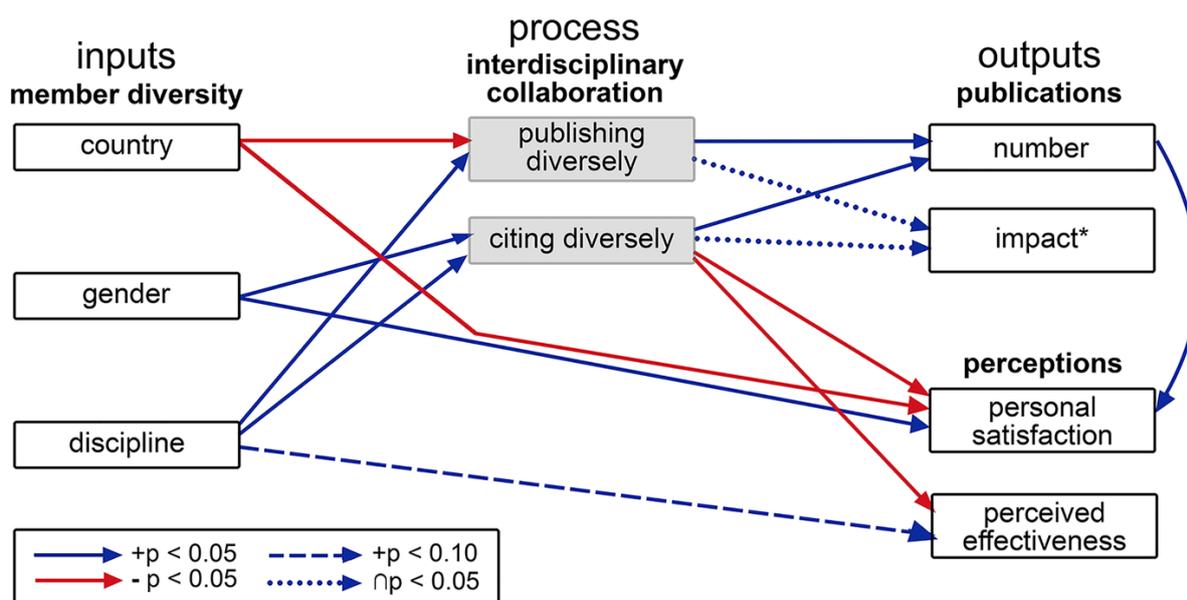


Fig 4. Model output. Graphical summary of the quantitative results. *specifically median citations received by each article.

Qualitative analysis

The outcomes from the open-ended feedback to the survey of member's satisfaction with and perceptions of the function of the group did not always compartmentalise neatly under our hypotheses nor did they directly complement the quantitative results. They added, however, more depth and some additional insights into perceptions of what made the group successful or not. It must be noted that any comments were entirely voluntary. These comments have been collated and can be viewed through the Environmental Data Initiative (link to be inserted on publication).

The quantitative analysis showed that diversity of country of origin, gender and disciplinary representation can affect group process and satisfaction with group function. There were many comments about the benefits and challenges because of disciplinary or skillset differences, while gender balance was never commented on.

Hypothesis 1b posited that there was a positive relationship between interdisciplinary collaboration and the interdisciplinarity of the group, and this was supported in the quantitative analysis. In the qualitative analysis, six respondents mentioned the positive benefits of having diverse disciplines in the groups, and although there were some difficulties, respondents mentioned that these difficulties were worth dealing with. As one commented, "Slightly different objectives and different viewpoints did not always make for the most effective discussions and decisions, but I think that is part of the process and is critical to interdisciplinary groups." (C-3, R14). There was awareness of a need to deal with any challenges arising from disciplinary differences "everyone's willingness to think actively about how to improve the merger between the disciplines" (C-5, R1). Generally, this diversity was regarded as stimulating "I learnt much by simply listening to so many different scientists with different and

complementary skillsets” (A-7, R1), “we brought in a diverse range of disciplines to liven up the stew.” (C-2, R1). “The distance between disciplines was clearly sufficient to create some learning opportunities, but not thought to impede output” (A-7, R2). “I found the functioning of the group very effective despite people coming from different expertises” (A-3, R2).

Disciplinary differences were often spoken of interchangeably with skillset differences, “We also worked on group dynamics so different personality types and skillsets could work together and benefit from the multidisciplinary (as opposed to becoming defensive in our corners).” (C-5, R2), “The diversity of skills represented...We were able to distribute the work effort in a way that made the most sense for moving the project forward.” (C-4, R1). A diversity of skillsets was inherent in every group in the study, and there were reasons for separation according to skillsets but it was felt that such separation should not exclude sub-groups from overall group discussions “Everyone got along just great, and we worked hard. I would gladly work with folks again. The group somewhat naturally by necessity split into two parts; those with the technical skills for the data wrangling and statistical modelling, and those who could not. The latter group thus was able to spend a lot of time conceptualizing, a good and important task, but it left out the technical folks from that part of the project. As one of the technical folks, I would like future projects to allow a more even mix of doing both.” (C-4, R5).

The value of diversity in general was, however, considered positive, “...I think the diversity of approaches and skills hugely contributed to the quality of the final paper. That paper was focused on how to effectively merge findings from fieldwork with the work of the modelers - that was very exciting and couldn't have happened without

the diverse membership of the group.” (C-5, R1). A member of another group, which comprised members from the non-research community, systems analysis, web developers, ecologists and those concerned with management and social engagement, noted that “having a diverse set of viewpoints to contribute to the work” (B-3, R1) was a factor that contributed to the group’s success.

Respondents identified several additional factors beyond those we measured in the quantitative analysis. These included the diversity of background (assumed to be organisation and country of origin) of team members, career stage representation, team fidelity, including having common goals and objectives, within-group trust and respect, and the nature of leadership.

Background was highlighted occasionally as evidenced by four responses: “To put it simply, the researchers and the management agency participants initially had difficulties communicating due to different perspectives and terminology barriers” (C-3, R6). However, this diversity was regarded by another in the group as “We had somewhat different perspectives and backgrounds but mostly saw eye-to-eye, had fun and got along well.” (C-3, R4). Rather interestingly, this group recognised the problem and turned it into a positive “I think the team did overcome this to some extent; indeed, one paper was largely devoted to terminology...” (C-3, R6). Another group observed the tension that can occur as “a lot of the collaboration depends on the diversity in the group, despite a temptation of some of the co-located members to conduct independent discussions and work without communicating with others.” (B-7, R5). This tendency for co-located members (especially if a common language was also involved) to have separate discussions excluding others in the group, was

observed more than once by the authors. Ensuring this division is not disruptive is a particular role for the hosts of the groups.

The presence of early career researchers (including postdocs) was mentioned by three respondents as an important, positive, contributor to group function, "...some members of our group, young postdoctoral students who were invited by their supervisors ... contributed a lot. Less established researchers can be a positive force in this type of working group." (A-7, R1); "The importance of the meeting and postdoc support ... to coordinate efforts and achieve plans jointly designed during meetings" (A-1, R2), "We had a good mix of career stage: senior scientists that served as excellent advisors, we had postdocs and PhD students that have the interest and ability to work and bring new ideas, and we had motivated early career scientists. We learned a lot and some excellent publications and collaborations came out of this working group." (C-5, R3).

Sharing common goals and objectives was considered an important factor in group effectiveness, "Convergence of goals despite different backgrounds, skills and kinds of contributions" (A-5, R7); "Common interest and objectives..." (A-7, R2); "...consensus over the common goals." (A-8, R2); "We had clear and common goals..." (C-1, R5); "...dedication to tasks; belief in goals" (B-27, R6); and "The group self-selected and each member has a vested interest in the successful outcomes and outputs" (B-7,R1).

The importance of good social skills, including within-group trust and respect was raised in several ways: "Enthusiasm, nice people, no intrigue" (A-7, R4); "...mutual trust...and friendly atmosphere" (A-4, R1); "We listened to each other and simply got along well!" (C-1, R5); "mutual respect" (B-1, R1); "compatibility, respect, and the

skills and knowledge of the members” (A-2, R4); “Tremendous individual integrity, trust in teammates, empathy in reading others' feeling, and supporting everyone on the team.” (C-2, R10); and “Collaborative and friendly attitude” (B-7, R6). “It is also a great group of people who were willing to listen to others, lead, let others lead, pull their weight, be responsive, be respectful, etc. Most of these people could have big egos, but they don't. It was an outstanding professional and personal experience.” (C-2, R12).

Several comments were made about the quality of leadership. If leadership was not good, things could become difficult: “...We did not have very strong group leadership, which also contributed to not getting as much done as we could have.” (A-8, R4).

This response incidentally was consistent with the quantitative ranking for effectiveness by group A-8 (section 4.1.3). “Effectiveness resulted from great leadership (encouraging openness to voices and ideas)...” (C-6, R2); “Good leadership” (B-4, R1); “Leadership. Co-chairs provided an encouraging working environment where each member of the WG can thrive.” (B-6, R2); “Strong leadership and vision from group PIs” (C-2, R13). As noted in the comments of R12, C-2 above, leadership need not rest with one person “Everybody's willingness to take a leadership role when needed, take a step back and letting others lead when needed, and never dismissing any idea just because it did not align with an a priori opinion” (C-6, R4).

As a final mention, however, is that these groups, as with many such groups, have a core set of enthusiasts and the face-to-face meetings were pivotal. Many of the non-targeted comments were about participation (those fence-sitters) and of the importance of the face-to-face meetings for getting work done. “A smaller subset of

the group did most of the work. Many folks in the group did little or nothing” (C-5, R4); “except for a couple of fellows” (A-2, R3); and “We had a few ‘doers,’ that is a few key folks kept things moving. A couple of the leaders were integral in keeping a vision up front” (C-2, R3). Getting people’s attention and commitment was easy within the meeting but between meetings (especially when there were a few) was harder: “most people are already overbooked for time so we had some difficulty finding people who wanted to lead projects.” (C-1, R3); “Once we went home after each meeting there was little contact and cooperation on projects. From my perspective, I think I could have made the outcome better if I had time to pursue collaborations after the face-to-face meetings. Unfortunately it was an extremely intense time in my ‘real’ job.” (A-2, R4); “The group was very effective during the week where we were together (at least the persons that assisted to the meetings). However, much of the momentum was lost after some months” (A-1, R1); and “Working sessions went generally well. But only three to four persons were really active between working sessions.” (A-3, R5). Basically “The number of meetings was super important and also helped people to have a sense of accountability. There is no substitute for being in the same room and working on projects together.” (A-4, R4). This conclusion has been argued cogently by Srivastava et al. [84].

Discussion

The goal of the work reported in this paper was to uncover the influences of group composition on the interdisciplinarity of teamwork, where the groups were attempting to find solutions to inter-disciplinary research problems. We posed a number of hypotheses about potential influences that may affect interdisciplinary practice

resulting in articles, particularly articles that have impact. To examine the dynamics experienced by such groups, we used an input-mediator-output model (Fig 1) with a sample of working groups in the environmental sciences.

We found that key aspects of group composition had a largely positive effect on interdisciplinary team processes: the gender balance and the diversity of disciplines represented in the group all increased interdisciplinary collaboration (Table 5 and Fig 4), while the number of countries represented in a group offset this general trend.

The number of countries is only a partial indicator of interpersonal, cultural challenges, but it is a reasonable measure which can imply a range of types of differences related to origin and background [85]. While it is a truism of research that correlation cannot prove causation, in this study the diversity of the team was set when the team was created, ruling out reverse causation, and making spurious correlation less plausible.

Our work has shown that the greater the level of interdisciplinary collaboration the greater the number of resulting publications. The impact of those publications (measured by the median level of citations received) was less straightforward, with greatest impact at moderate levels of interdisciplinary collaboration (Fig 3). It seems that being too interdisciplinary can mute impact, and too little may be 'boring'. We acknowledge that our citation data are relatively short-term (1-10 years after publication). Even though papers are usually cited most heavily shortly after publication, it is possible that more interdisciplinary papers will go on to be more cited in the longer term, as suggested in the literature [51]. The causal direction of these results might be debated, though it is not clear why teams that publish more seem to choose to do so in an interdisciplinary way, and as citations follow the process, reverse causation at least seems ruled out.

Group-member perceptions enriched our understanding of the process. Not surprisingly personal satisfaction increased with the number of publications produced and, more unexpectedly perhaps, with the proportion of women (Fig 4). Most evident from the quantitative study was the negative effect of country and diverse work practice (measured by the level of diversity of articles sourced) on personal satisfaction and perceived effectiveness of the group. There were no comments received about difficulties specifically related to nationality. The opportunity of working with people from different disciplines and backgrounds was generally considered stimulating, and respondents recognised that this discomfort did not relate to the quality of the output.

The qualitative responses highlighted that different backgrounds were often difficult at first, but adjustments could be made in the time-frame available to these groups. Four factors were identified which promoted productivity and offset any difficult collaboration barriers: (i) having early career members in the group, (ii) sharing common goals and objectives, (iii) mutual trust and respect, and (iv) the quality and team-centered nature of group leadership (as per [86]). It must be remembered that these groups worked together in a formal manner (as in having supported collaboration) for up to four years with multiple meetings during that time, and they had ample time to reflect upon and resolve their differences to produce their outcomes.

Our study examined working groups, which are scientific teams, but with differences from managerially-assembled teams employed on a research project. For instance, it is easy for a member of a voluntary working group who is dissatisfied with the team process to discontinue participation. While we believe the hypothesized relationships

between demographics and process and process and outputs should not be affected by the differences, the generalizability of the model should be tested with a broader range of team types.

In summary, it can be said that the number of (relevant) disciplines brought to the table and a 'good' gender balance in a group, an interdisciplinary process can be enhanced. Open-ended feedback suggested that this would be supported by having young members in a group, and ensuring good *esprit de corps*. A strong interdisciplinary process is positively related to having more publications, with the proviso that exploring and publishing too diversely is related to reduced impact. It seems that in the case of deliberately-formed scientific working groups, with sufficient time challenges can be successfully tackled as the same group of people learn to work together.

Conclusion

A fundamental question for our study was whether a diversity of participants in a working group does lead to diverse practices and outcomes. We think we have shown that, with some provisos and a degree of management, it does. We used an original conceptual model, namely positing the interdisciplinarity of collaboration as a moderating factor between team demographics and outcomes, and we think this has provided new insight into how groups function. The overall pattern of the findings is mostly consistent with our expectations, while the contradictory findings point to possible new perspectives on the function of research groups and a consequent management focus. We suggest that group diversity is not just a goal in itself, but is

rather a support for building an interdisciplinary collaboration, the success of which has benefits for output (in moderation).

While our findings point to several factors important for working group success, they are limited by the fact that we had data from only 22 groups and 3 centers. As the survey response rates were not high, the conclusions about satisfaction and effectiveness could be more robust, and this deserves further exploration. We should also note that the groups studied were assembled by organisations that were already sensitive to some of the factors we have been discussing.

The qualitative data point to several factors important to the group participants that we were unable to capture with our quantitative data but were important to the successful function of the groups. Approaches such as the use of sociometers [87] and autoethnography [35] for example, could add greatly to the further understanding of the group dynamics that promote positive research outcomes.

Even with the limitations of our dataset, our findings reinforce the importance and apparent benefits of diversity of various types in research groups. They also indicate the need to pay attention to the interdisciplinarity of the collaboration itself and being more deliberate about the combination of new knowledge, working to ensure that the groups bring together their diverse perspectives. The negative impacts on satisfaction and effectiveness due to diversity due to (i) participant countries and (ii) the diversity of cited publications suggests a need to help participants with the knowledge integration process to make it more enjoyable and ensure fidelity within the group.

While our study is correlational rather than causal, our findings do suggest some guidelines for the management of cross-boundary groups to successfully engage in interdisciplinary work such as:

1. ensure gender balance in a group, as this seems positively related to the interdisciplinarity of the work, the number of articles produced, and to overall group satisfaction;
2. encourage publication (in conferences or journals) throughout the group process, as group satisfaction improves with the number of publications (in contrast to the delayed publication implied in a typical Working Group Workflow);
3. be cautious if high impact articles are desired. Too much interdisciplinarity may be a distraction;
4. allow the group time and give them support to work out ways to deal with personnel differences as well managed differences can be stimulating and rewarding;
5. manage and support the discomfort that can occur when working across disciplinary boundaries, as some tension can result in novel outcomes; and
6. support the group to achieve and maintain focus on the common goal, and ensure there is mutual respect and trust between members.

As stated at the beginning of this paper, global research challenges increasingly require the formation of teams that bring together individuals with diverse skills and perspectives who can work across disciplinary, organisational and national divides. Indeed, research has become so complex that individual scientists cannot achieve meaningful results without collaborating—the so-called collaboration imperative [88].

But while interdisciplinary research is needed, will assembling a diverse team necessarily result in publications that have impact? Our results show that interdisciplinary work is, in fact, positively related to publication rates, but if those publications are themselves spread across too wide a range of disciplinary domains (perhaps reflecting the work itself), their impact may be lessened, and contrary-wise, too narrow a view will diminish impact. At the same time some other aspects of team diversity can enhance or diminish this output. Paying attention to some simple factors in the design and management of such groups can improve the likelihood of a positive research outcome and enhance the satisfaction of group members.

Acknowledgements

Authors would like to acknowledge DataONE for facilitating our participation in the Usability and Assessment Working Group (inter alia) and for their consistent support for this project. We would also like to acknowledge the support and contribution of the synthesis center CESAB of the French Foundation for Research on Biodiversity and Jill Baron, co-Director of the John Wesley Powell synthesis center of the USGS. We are grateful for the open and transparent work practices of all three organisations which enabled this work to be conducted. We also would like to thank the several reviewers who provided invaluable advice on drafts of the article.

References

1. Baker B. The Science of Team Science. *BioScience*. 2015;65: 639–644.
doi:10.1093/biosci/biv077
2. Cheruvellil KS, Soranno PA. Data-Intensive Ecological Research Is Catalyzed by Open Science and Team Science. *BioScience*. 2018;68: 813–822.
doi:10.1093/biosci/biy097
3. Wagner CS, Whetsell TA, Mukherjee S. International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. *Research Policy*. 2019;48: 1260–1270. doi:10.1016/j.respol.2019.01.002
4. Xu J, Ding Y, Malic V. Author Credit for Transdisciplinary Collaboration. Glanzel W, editor. *PLoS ONE*. 2015;10: e0137968. doi:10.1371/journal.pone.0137968
5. Laursen BK, Motzer N, Anderson KJ. Pathway profiles: Learning from five main approaches to assessing interdisciplinarity. *Research Evaluation*. 2022; rvac036.
doi:10.1093/reseval/rvac036
6. Specht A, Corrêa P, Belbin L, Loescher HW. Critical research infrastructure: The importance of synthesis centers. *Elephant in the Lab*. 2020 [cited 12 Jan 2022].
doi:10.5281/zenodo.3660920
7. Fontana M, Iori M, Montobbio F, Sinatra R. New and atypical combinations: An assessment of novelty and interdisciplinarity. *Research Policy*. 2020;49: 104063.
doi:10.1016/j.respol.2020.104063
8. Roy ED, Morzillo AT, Seijo F, Reddy SMW, Rhemtulla JM, Milder JC, et al. The Elusive Pursuit of Interdisciplinarity at the Human—Environment Interface. *BioScience*. 2013;63: 745–753. doi:10.1525/bio.2013.63.9.10
9. Gates AJ, Ke Q, Varol O, Barabási A-L. Nature's reach: narrow work has broad impact. *Nature*. 2019;575: 32–34. doi:10.1038/d41586-019-03308-7
10. Yegros-Yegros A, Rafols I, D'Este P. Does Interdisciplinary Research Lead to Higher Citation Impact? The Different Effect of Proximal and Distal Interdisciplinarity. Glanzel W, editor. *PLoS ONE*. 2015;10: e0135095.
doi:10.1371/journal.pone.0135095

11. Coccia M. The evolution of scientific disciplines in applied sciences: dynamics and empirical properties of experimental physics. *Scientometrics*. 2020;124: 451–487. doi:10.1007/s11192-020-03464-y
12. Sharp PA, Langer R. Promoting Convergence in Biomedical Science. *Science*. 2011;333: 527–527. doi:10.1126/science.1205008
13. Pluchino A, Burgio G, Rapisarda A, Biondo AE, Pulvirenti A, Ferro A, et al. Exploring the role of interdisciplinarity in physics: Success, talent and luck. Hernandez Montoya AR, editor. *PLoS ONE*. 2019;14: e0218793. doi:10.1371/journal.pone.0218793
14. Hall KL, Stokols D, Stipelman BA, Vogel AL, Feng A, Masimore B, et al. Assessing the Value of Team Science. *American Journal of Preventive Medicine*. 2012;42: 157–163. doi:10.1016/j.amepre.2011.10.011
15. Harvey S. Creative Synthesis: Exploring the Process of Extraordinary Group Creativity. *AMR*. 2014;39: 324–343. doi:10.5465/amr.2012.02241.
16. Baron JS, Specht A, Garnier E, Bishop P, Campbell CA, Davis FW, et al. Synthesis Centers as Critical Research Infrastructure. *BioScience*. 2017;67: 750–759. doi:10.1093/biosci/bix053
17. Crowston K, Specht A, Hoover C, Chudoba KM, Watson-Manheim MB. Perceived discontinuities and continuities in transdisciplinary scientific working groups. *Science of The Total Environment*. 2015;534: 159–172. doi:10.1016/j.scitotenv.2015.04.121
18. Hampton SE, Parker JN. Collaboration and Productivity in Scientific Synthesis. *BioScience*. 2011;61: 900–910. doi:10.1525/bio.2011.61.11.9
19. Lee Y-N, Walsh JP, Wang J. Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*. 2015;44: 684–697. doi:10.1016/j.respol.2014.10.007
20. Read EK, O'Rourke M, Hong GS, Hanson PC, Winslow LA, Crowley S, et al. Building the team for team science. Peters DPC, editor. *Ecosphere*. 2016;7. doi:10.1002/ecs2.1291

21. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*. 2010;330: 686–688. doi:10.1126/science.1193147
22. Bachmann R. Trust, Power and Control in Trans-Organizational Relations. *Organization Studies*. 2001;22: 337–365. doi:10.1177/0170840601222007
23. Millar J, Demaid A, Quintas P. Trans-organizational innovation: a framework for research. *Technology Analysis & Strategic Management*. 1997;9: 399–418. doi:10.1080/09537329708524294
24. Verhoeven D, Cooper T, Flynn M, Shuffler ML. Transnational Team Effectiveness. 1st ed. In: Salas E, Rico R, Passmore J, editors. *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes*. 1st ed. Wiley; 2017. pp. 73–101. doi:10.1002/9781118909997.ch4
25. Horwitz SK, Horwitz IB. The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography. *Journal of Management*. 2007;33: 987–1015. doi:10.1177/0149206307308587
26. Woolley AW, Aggarwal I, Malone TW. Collective Intelligence and Group Performance. *Curr Dir Psychol Sci*. 2015;24: 420–424. doi:10.1177/0963721415599543
27. Bernstein E, Shore J, Lazer D. How intermittent breaks in interaction improve collective intelligence. *Proc Natl Acad Sci USA*. 2018;115: 8734–8739. doi:10.1073/pnas.1802407115
28. Hackett EJ, Parker JN. From Salomon’s House to Synthesis Centers. In: Heinze T, Münch R, editors. *Innovation in Science and Organizational Renewal*. New York: Palgrave Macmillan US; 2016. pp. 53–87. doi:10.1057/978-1-137-59420-4_3
29. O’Rourke M, Crowley S, Gonnerman C. On the nature of cross-disciplinary integration: A philosophical framework. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*. 2016;56: 62–70. doi:10.1016/j.shpsc.2015.10.003

30. Stokols D, Fuqua J, Gress J, Harvey R, Phillips K, Baezconde-Garbanati L, et al. Evaluating transdisciplinary science. *Nicotine & Tobacco Res.* 2003;5: 21–39. doi:10.1080/14622200310001625555
31. Lynch AJJ, Thackway R, Specht A, Beggs PJ, Brisbane S, Burns EL, et al. Transdisciplinary synthesis for ecosystem science, policy and management: The Australian experience. *Science of The Total Environment.* 2015;534: 173–184. doi:10.1016/j.scitotenv.2015.04.100
32. Hackman, J.R., 1977. Work design. In: Hackman JR, Suttle JL, editors. *Improving life at work: behavioral science approaches to organizational change.* Santa Monica, Calif: Goodyear Pub. Co; 1977.
33. Bammer G, O'Rourke M, O'Connell D, Neuhauser L, Midgley G, Klein JT, et al. Expertise in research integration and implementation for tackling complex problems: when is it needed, where can it be found and how can it be strengthened? *Palgrave Commun.* 2020;6: 5. doi:10.1057/s41599-019-0380-0
34. Abramo G, D'Angelo CA, Zhang L. 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.* 2018;12: 1182–1193. doi:10.1016/j.joi.2018.09.001
35. Dusdal J, Powell JJW. Benefits, Motivations, and Challenges of International Collaborative Research: A Sociology of Science Case Study. *Science and Public Policy.* 2021;48: 235–245. doi:10.1093/scipol/scab010
36. Stvilia B, Hinnant CC, Schindler K, Worrall A, Burnett G, Burnett K, et al. Composition of scientific teams and publication productivity at a national science lab. *J Am Soc Inf Sci.* 2011;62: 270–283. doi:10.1002/asi.21464
37. De Saá-Pérez P, Díaz-Díaz NL, Aguiar-Díaz I, Ballesteros-Rodríguez JL. How diversity contributes to academic research teams performance: Diversity in academic research teams performance. *R&D Management.* 2017;47: 165–179. doi:10.1111/radm.12139

38. Campbell LG, Mehtani S, Dozier ME, Rinehart J. Gender-Heterogeneous Working Groups Produce Higher Quality Science. Larivière V, editor. PLoS ONE. 2013;8: e79147. doi:10.1371/journal.pone.0079147
39. Meslec N, Graff D. Being open matters: the antecedents and consequences of cross-understanding in teams. *Team Performance Management*. 2015;21: 6–18. doi:10.1108/TPM-10-2014-0055
40. van Rijnsoever FJ, Hessels LK. Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy*. 2011;40: 463–472. doi:10.1016/j.respol.2010.11.001
41. Bozeman B, Gaughan M. How do men and women differ in research collaborations? An analysis of the collaborative motives and strategies of academic researchers. *Research Policy*. 2011;40: 1393–1402. doi:10.1016/j.respol.2011.07.002
42. Xie L, Zhou J, Zong Q, Lu Q. Gender diversity in R&D teams and innovation efficiency: Role of the innovation context. *Research Policy*. 2020;49: 103885. doi:10.1016/j.respol.2019.103885
43. Curşeu PL, Pluut H, Boroş S, Meslec N. The magic of collective emotional intelligence in learning groups: No guys needed for the spell! *Br J Psychol*. 2015;106: 217–234. doi:10.1111/bjop.12075
44. Landry R, Traore N, Godin B. An econometric analysis of the effect of collaboration on academic research productivity. *High Educ*. 1996;32: 283–301. doi:10.1007/BF00138868
45. Fox MFrank, Mohapatra Sushanta. Social-Organizational Characteristics of Work and Publication Productivity among Academic Scientists in Doctoral-Granting Departments. *The Journal of Higher Education*. 2007;78: 542–571. doi:10.1353/jhe.2007.0032
46. Leahey E, Beckman CM, Stanko TL. Prominent but Less Productive: The Impact of Interdisciplinarity on Scientists' Research. *Administrative Science Quarterly*. 2017;62: 105–139. doi:10.1177/0001839216665364

47. Belkhouja M, Fattoum S, Yoon H (David). Does greater diversification increase individual productivity? The moderating effect of attention allocation. *Research Policy*. 2021;50: 104256. doi:10.1016/j.respol.2021.104256
48. Barjak F, Robinson S. International collaboration, mobility and team diversity in the life sciences: impact on research performance. *Soc Geogr*. 2008;3: 23–36. doi:10.5194/sg-3-23-2008
49. Maddi A, Gingras Y. Gender diversity in research teams and citation impact in economics and management. *Journal of Economic Surveys*. 2021; joes.12420. doi:10.1111/joes.12420
50. Hackett EJ, Leahey E, Parker JN, Rafols I, Hampton SE, Corte U, et al. Do synthesis centers synthesize? A semantic analysis of topical diversity in research. *Research Policy*. 2021;50: 104069. doi:10.1016/j.respol.2020.104069
51. Chen S, Qiu J, Arsenault C, Larivière V. Exploring the interdisciplinarity patterns of highly cited papers. *Journal of Informetrics*. 2021;15: 101124. doi:10.1016/j.joi.2020.101124
52. Larivière V, Haustein S, Börner K. Long-Distance Interdisciplinarity Leads to Higher Scientific Impact. Glanzel W, editor. *PLoS ONE*. 2015;10: e0122565. doi:10.1371/journal.pone.0122565
53. Steele TW, Stier JC. The impact of interdisciplinary research in the environmental sciences: a forestry case study. *Journal of the American Society for Information Science*. 2000;51: 476–484. doi:10.1002/(SICI)1097-4571(2000)51:5<476::AID-ASI8>3.0.CO;2-G
54. Belkhouja M, Yoon H (David). How does openness influence the impact of a scholar's research? An analysis of business scholars' citations over their careers. *Research Policy*. 2018;47: 2037–2047. doi:10.1016/j.respol.2018.07.012
55. Whitfield J. Collaboration: Group theory. *Nature*. 2008;455: 720–723. doi:10.1038/455720a
56. Schoonenboom J, Johnson RB. How to Construct a Mixed Methods Research Design. *Köln Z Soziol*. 2017;69: 107–131. doi:10.1007/s11577-017-0454-1

57. Twyman M, Contractor N. Team Assembly. In: Hall KL, Vogel AL, Croyle RT, editors. *Strategies for Team Science Success: Handbook of Evidence-Based Principles for Cross-Disciplinary Science and Practical Lessons Learned from Health Researchers*. Cham: Springer International Publishing; 2019. pp. 217–240. doi:10.1007/978-3-030-20992-6_17
58. Palmer MA, Kramer JG, Boyd J, Hawthorne D. Practices for facilitating interdisciplinary synthetic research: the National Socio-Environmental Synthesis Center (SESYNC). *Current Opinion in Environmental Sustainability*. 2016;19: 111–122. doi:10.1016/j.cosust.2016.01.002
59. Michener WK, Allard S, Budden A, Cook RB, Douglass K, Frame M, et al. Participatory design of DataONE—Enabling cyberinfrastructure for the biological and environmental sciences. *Ecological Informatics*. 2012;11: 5–15. doi:10.1016/j.ecoinf.2011.08.007
60. Halpern BS, Berlow E, Williams R, Borer ET, Davis FW, Dobson A, et al. Ecological Synthesis and Its Role in Advancing Knowledge. *BioScience*. 2020; biaa105. doi:10.1093/biosci/biaa105
61. Specht, A., Synthesis centers: their relevance to and importance in the Anthropocene. In: Chabbi A, Loescher HW. *Terrestrial ecosystem research infrastructures: challenges and opportunities*. 2017. Available: <http://www.crcnetbase.com/isbn/9781498751339>
62. Bergek A, Norrman C. Incubator best practice: A framework. *Technovation*. 2008;28: 20–28. doi:10.1016/j.technovation.2007.07.008
63. Rodrigo A, Alberts S, Cranston K, Kingsolver J, Lapp H, McClain C, et al. Science Incubators: Synthesis Centers and Their Role in the Research Ecosystem. *PLoS Biol*. 2013;11: e1001468. doi:10.1371/journal.pbio.1001468
64. Hackett, E.J., Parker, J.N., Conz, D., Rhoten, D., Parker, A., *Ecology transformed: NCEAS and changing patterns of ecological research*. In: Olson, G.M., Zimmerman, A., Bos, N. (Eds.), *Scientific collaboration on the Internet, Acting with technology*. MIT Press, Cambridge, Mass.; 2008. pp. 277-296.

65. Australian and New Zealand Standard Research Classification (ANZSRC), 2020 | Australian Bureau of Statistics. 30 Jun 2020 [cited 5 Jul 2021]. Available: <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-research-classification-anzsrc/latest-release> (accessed 7.5.21).
66. Wagner CS, Roessner JD, Bobb K, Klein JT, Boyack KW, Keyton J, et al. Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*. 2011;5: 14–26. doi:10.1016/j.joi.2010.06.004
67. Leydesdorff L, Rafols I. Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*. 2011;5: 87–100. doi:10.1016/j.joi.2010.09.002
68. Rafols I, Leydesdorff L, O’Hare A, Nightingale P, Stirling A. How journal rankings can suppress interdisciplinary research: A comparison between *Innovation Studies and Business & Management*. *Research Policy*. 2012;41: 1262–1282. doi:10.1016/j.respol.2012.03.015
69. Marx W, Bornmann L. Change of perspective: bibliometrics from the point of view of cited references—a literature overview on approaches to the evaluation of cited references in bibliometrics. *Scientometrics*. 2016;109: 1397–1415. doi:10.1007/s11192-016-2111-2
70. Leydesdorff L, Wagner CS, Bornmann L. Betweenness and diversity in journal citation networks as measures of interdisciplinarity—A tribute to Eugene Garfield. *Scientometrics*. 2018;114: 567–592. doi:10.1007/s11192-017-2528-2
71. Qin J, Lancaster FW, Allen B. Types and levels of collaboration in interdisciplinary research in the sciences. *Journal of the American Society for Information Science*. 1997;48: 893–916. doi:https://doi.org/10.1002/(SICI)1097-4571(199710)48:10<893::AID-ASI5>3.0.CO;2-X
72. Huang L, Cai Y, Zhao E, Zhang S, Shu Y, Fan J. Measuring the interdisciplinarity of Information and Library Science interactions using citation analysis and semantic analysis. *Scientometrics*. 2022 [cited 26 Jul 2022]. doi:10.1007/s11192-022-04401-x

73. Shu F, Dinneen JD, Chen S. Measuring the disparity among scientific disciplines using Library of Congress Subject Headings. *Scientometrics*. 2022;127: 3613–3628. doi:10.1007/s11192-022-04387-6
74. Craven D, Winter M, Hotzel K, Gaikwad J, Eisenhauer N, Hohmuth M, et al. Evolution of interdisciplinarity in biodiversity science. *Ecol Evol*. 2019;9: 6744–6755. doi:10.1002/ece3.5244
75. Kerr NL, Tindale RS. Group Performance and Decision Making. *Annu Rev Psychol*. 2004;55: 623–655. doi:10.1146/annurev.psych.55.090902.142009
76. Stanley T, Matthews J, Davidson P. Insights from Stimulating Creative Behaviours in a Project-Based Organization Team. *Technology Innovation Management Review*. 2016;6: 26–33. doi:10.22215/timreview/979
77. Spellerberg IF, Fedor PJ. A tribute to Claude Shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon-Wiener’ Index: On species richness and diversity. *Global Ecology and Biogeography*. 2003;12: 177–179. doi:10.1046/j.1466-822X.2003.00015.x
78. Aydinoglu AU, Allard S, Mitchell C. Measuring diversity in disciplinary collaboration in research teams: An ecological perspective. *Research Evaluation*. 2016;25: 18–36. doi:10.1093/reseval/rvv028
79. Wang J, Thijs B, Glänzel W. Interdisciplinarity and Impact: Distinct Effects of Variety, Balance, and Disparity. Smalheiser NR, editor. *PLoS ONE*. 2015;10: e0127298. doi:10.1371/journal.pone.0127298
80. Uzzi B, Mukherjee S, Stringer M, Jones B. Atypical Combinations and Scientific Impact. *Science*. 2013;342: 468–472. doi:10.1126/science.1240474
81. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology*. 2006;3: 77–101. doi:10.1191/1478088706qp063oa
82. Patience GS, Patience CA, Blais B, Bertrand F. Citation analysis of scientific categories. *Heliyon*. 2017;3: e00300. doi:10.1016/j.heliyon.2017.e00300
83. Specht A, Crowston K. Analysis of papers produced and cited by synthesis center-type groups in the USA and France, 2010-2020, and open-source

feedback from those groups. Environmental Data Initiative; 2022.
doi:10.6073/PASTA/C5033B64FC57E2B3DA45C8F8541BD98A

84. Srivastava DS, Winter M, Gross LJ, Metzger JP, Baron JS, Mouquet N, et al. Maintaining momentum for collaborative working groups in a post-pandemic world. *Nat Ecol Evol.* 2021; 1–2. doi:10.1038/s41559-021-01521-0
85. Salazar M, Salas E. Reflections of Cross-Cultural Collaboration Science: REFLECTIONS OF CROSS-CULTURAL COLLABORATION SCIENCE. *J Organiz Behav.* 2013;34: 910–917. doi:10.1002/job.1881
86. Kozlowski SWJ, Mak S, Chao GT. Team-Centric Leadership: An Integrative Review. *Annu Rev Organ Psychol Organ Behav.* 2016;3: 21–54.
doi:10.1146/annurev-orgpsych-041015-062429
87. Parker JN, Cardenas E, Dorr AN, Hackett EJ. Using Sociometers to Advance Small Group Research. *Sociological Methods & Research.* 2020;49: 1064–1102.
doi:10.1177/0049124118769091
88. Bozeman B, Boardman C. *Research Collaboration and Team Science.* Cham: Springer International Publishing; 2014. doi:10.1007/978-3-319-06468-0

Supporting Information

S1 Appendix Demographic profiles of the groups from the contributing organisations.

| group | start | members | female (%) | disciplines | countries |
|-------|-------|---------|------------|-------------|-----------|
| A-1 | 2014 | 11 | 9 | 5 | 5 |
| A-2 | 2012 | 15 | 33 | 6 | 4 |
| A-3 | 2010 | 12 | 25 | 5 | 2 |
| A-4 | 2014 | 14 | 36 | 3 | 7 |
| A-5 | 2010 | 17 | 18 | 2 | 8 |
| A-6 | 2015 | 20 | 15 | 5 | 10 |
| A-7 | 2011 | 13 | 31 | 6 | 7 |
| A-8 | 2011 | 13 | 8 | 4 | 6 |
| A-9 | 2012 | 17 | 18 | 5 | 7 |
| B-1 | 2011 | 17 | 67 | 9 | 1 |
| B-2 | 2010 | 24 | 0 | 4 | 1 |
| B-3 | 2012 | 18 | 42 | 7 | 2 |
| B-4 | 2010 | 19 | 41 | 10 | 2 |
| B-5 | 2010 | 21 | 24 | 4 | 4 |
| B-6 | 2010 | 15 | 60 | 6 | 2 |
| B-7 | 2010 | 16 | 63 | 8 | 4 |
| C-1 | 2014 | 20 | 40 | 7 | 4 |
| C-2 | 2014 | 22 | 24 | 7 | 1 |
| C-3 | 2012 | 17 | 28 | 9 | 7 |
| C-4 | 2015 | 14 | 41 | 7 | 2 |
| C-5 | 2013 | 20 | 38 | 5 | 3 |
| C-6 | 2015 | 28 | 32 | 8 | 2 |

The organisations and groups have been given codes to anonymise the source.

S2 Appendix. Profile of publications produced and cited by each group.

| Group | publications | | | | | articles cited | | |
|-------|--------------|------------------------|-------------------|---------------------|--------------------------------|----------------|------------------------|-------------------|
| | number | disciplinary divisions | Shannon diversity | average no. authors | median citations (August 2020) | number | disciplinary divisions | Shannon diversity |
| A-1 | 2 | 2 | 0.693 | 10.50 | 12.5 | 79 | 2 | 0.287 |
| A-2 | 5 | 2 | 0.673 | 13.20 | 19 | 204 | 9 | 1.249 |
| A-3 | 9 | 2 | 0.530 | 11.44 | 22 | 591 | 7 | 0.740 |
| A-4 | 3 | 1 | 0 | 19.00 | 4.5 | 99 | 6 | 0.716 |
| A-5 | 24 | 2 | 0.377 | 8.54 | 41 | 1294 | 11 | 0.834 |
| A-6 | 3 | 1 | 0 | 15.00 | 4 | 137 | 12 | 1.425 |
| A-7 | 8 | 1 | 0 | 8.25 | 15 | 456 | 13 | 0.750 |
| A-8 | 4 | 3 | 1.040 | 10.25 | 14.5 | 167 | 9 | 0.986 |
| A-9 | 7 | 6 | 1.748 | 17.29 | 26 | 350 | 11 | 0.902 |
| B-1 | 9 | 6 | 1.677 | 5.00 | 8 | 154 | 13 | 1.181 |
| B-2 | 2 | 2 | 0.693 | 2.50 | 175.5 | 13 | 3 | 0.721 |
| B-3 | 6 | 3 | 1.011 | 5.17 | 152 | 101 | 13 | 1.362 |
| B-4 | 2 | 2 | 0.693 | 7.00 | 105 | 53 | 10 | 1.407 |
| B-5 | 4 | 1 | 0.000 | 11.00 | 17.5 | 39 | 10 | 0.883 |
| B-6 | 4 | 1 | 0.693 | 1.5 | 8 | 130 | 10 | 1.408 |
| B-7 | 18 | 7 | 1.874 | 5.61 | 10 | 346 | 17 | 1.301 |
| C-1 | 4 | 3 | 1.039 | 17.50 | 21 | 194 | 10 | 1.245 |
| C-2 | 14 | 7 | 1.907 | 9.42 | 22.5 | 618 | 10 | 1.545 |
| C-3 | 10 | 2 | 0.325 | 8.90 | 25 | 602 | 15 | 1.001 |
| C-4 | 3 | 2 | 0.637 | 11.00 | 11 | 244 | 10 | 1.656 |
| C-5 | 6 | 3 | 1.055 | 5.80 | 17 | 476 | 9 | 1.255 |
| C-6 | 5 | 4 | 1.330 | 8.83 | 13 | 400 | 14 | 1.584 |

The organisations and groups have been given codes to anonymise the source.

S3 Appendix. Responses to questions about perceived effectiveness and satisfaction with the group by respondents.

| Group | response rate (%) | perceived group effectiveness (mean \pm se) | satisfaction with group function (mean \pm se) |
|-------|-------------------|---|--|
| A-1 | 36.36 | 4.4 \pm 0.22 | 4.2 \pm 0.34 |
| A-2 | 33.33 | 4.4 \pm 0.22 | 4.2 \pm 0.34 |
| A-3 | 41.67 | 4.25 \pm 0.42 | 4.25 \pm 0.42 |
| A-4 | 35.71 | 4.8 \pm 0.18 | 4.6 \pm 0.22 |
| A-5 | 41.18 | 4.57 \pm 0.28 | 4.57 \pm 0.28 |
| A-6 | 30.00 | 3.4 \pm 0.36 | 3.6 \pm 0.22 |
| A-7 | 38.46 | 3 \pm 0.56 | 3.8 \pm 0.44 |
| A-8 | 30.77 | 4.5 \pm 0.56 | 3 \pm 0.61 |
| A-9 | 29.41 | 4 \pm 0 | 3.5 \pm 0.25 |
| B-1 | 35.29 | 4 \pm 0.24 | 3.83 \pm 0.15 |
| B-2 | 29.17 | 4.5 \pm 0.25 | 4.13 \pm 0.45 |
| B-3 | 27.78 | 4.2 \pm 0.33 | 4.4 \pm 0.22 |
| B-4 | 31.58 | 4.6 \pm 0.22 | 4 \pm 0.28 |
| B-5 | 20.00 | 3.67 \pm 0.72 | 3.33 \pm 0.54 |
| B-6 | 46.67 | 4.8 \pm 0.18 | 4.5 \pm 0.22 |
| B-7 | 56.25 | 4.5 \pm 0.18 | 4.5 \pm 0.25 |
| C-1 | 25.00 | 4.2 \pm 0.18 | 4.2 \pm 0.34 |
| C-2 | 68.18 | 4.63 \pm 0.25 | 4.75 \pm 0.23 |
| C-3 | 41.18 | 4.57 \pm 0.28 | 4.43 \pm 0.4 |
| C-4 | 53.57 | 4.14 \pm 0.26 | 4.29 \pm 0.26 |
| C-5 | 25.00 | 4.2 \pm 0.34 | 4.4 \pm 0.31 |
| C-6 | 25.00 | 4 \pm 0.31 | 4.43 \pm 0.26 |

The organisations and groups have been given codes to anonymise the source.