INNOVATION & RESEARCH CAUCUS

CULTURAL DIVERSITY AND THE INTERDISCIPLINARITY OF RESEARCH OUTPUTS:

Evidence From UK Research Council Funded Projects

IRC Report No: 077a

REPORT PREPARED BY

Abdullah Gök
University of Exeter

Ali Mammadow
University of Exeter

Delivered with ESRC and Innovate UK

Authors
The core members of the research team for this project were as follows:

- Dr Abdullah Gök - University of Exeter
- Ali Mammadow - University of Exeter

This document relates to IRC Project FFGtR004: Cultural diversity and the interdisciplinarity of research outputs: Evidence from UK Research Council funded projects

Acknowledgements
This work was supported by Economic and Social Research Council (ESRC) grant ES/X010759/1 to the Innovation and Research Caucus (IRC).

We are very grateful to the project sponsors at UK Research & Innovation (UKRI) for their input into this research. The interpretations and opinions within this report are those of the authors and may not reflect the policy positions of UKRI or its constituent councils. We would also like to acknowledge and appreciate the efforts of the IRC Project Administration Team for their support in preparing this report.

About the Innovation and Research Caucus
The Innovation and Research Caucus supports the use of robust evidence and insights in UKRI's strategies and investments, as well as undertaking a co-produced programme of research. Our members are leading academics from across the social sciences, other disciplines and sectors, who are engaged in different aspects of innovation and research systems. We connect academic experts, UKRI, IUK and the (ESRC), by providing research insights to inform policy and practice. Professor Tim Vorley and Professor Stephen Roper are Co-Directors. The IRC is funded by UKRI via the ESRC and IUK, grant number ES/X010759/1. The support of the funders is acknowledged. The views expressed in this piece are those of the authors and do not necessarily represent those of the funders.

Contact
You are also welcome to email us if you have any questions about this report or the work of the IRC generally: info@ircaucus.ac.uk

Cite as: Gok, A. & Mammadow, A. 2026 *Cultural Diversity and the Interdisciplinarity of Research Outputs: Evidence from UK Research Council Funded Projects*. Oxford, UK: Innovation and Research Caucus

CONTENTS

Abstract

Cultural diversity is widely expected to broaden the cognitive and epistemic resources available to research teams, and interdisciplinarity is often assumed to be a pathway through which this potential translates into research outputs. This paper tests that assumption using 2,795 UK Research Council funded projects, linking Gateway to Research Plus (GtR+) records to OpenAlex bibliometric data. Cultural diversity is measured as the country-of-origin diversity of the research-active project team, using a distance-aware Leinster and Cobbold index based on name-based origin inference. Output interdisciplinarity is measured with publication-level Rao-Stirling indicators aggregated to the project level. Generalised additive models allow the relationship to take a non-linear form. The results support neither a positive association at low to moderate diversity nor the hypothesised weakening at higher levels. Cultural diversity is statistically detectable only in simpler specifications, where the association is small and negative rather than positive; it contributes very little to model fit and becomes marginal once the prior interdisciplinarity of principal and co-investigators is controlled. Prior interdisciplinarity is the strongest measured predictor. Realised interdisciplinarity appears strongly path-dependent, and compositional diversity alone appears insufficient to generate it. Implications for diversity research and research funding policy are discussed.

1. Introduction

Cultural diversity is widely expected to matter for research and innovation. Diverse research teams may bring together different cognitive frames, epistemic traditions, methodological repertoires and problem definitions, expanding the range of ideas available for scientific problem solving (Hofstra et al., 2020; Stahl and Maznevski, 2021; Wang et al., 2019). In science and innovation policy, this expectation has become especially important because contemporary research is increasingly organised around complex problems that require knowledge to be combined across disciplinary, institutional and social boundaries (Porter and Rafols, 2009; Specht and Crowston, 2022). Yet the mechanisms through which cultural diversity might shape research outputs remain insufficiently understood. In particular, it is not clear whether culturally diverse research teams are more likely to produce interdisciplinary outputs, or whether realised interdisciplinarity is shaped primarily by pre-existing disciplinary trajectories and prior research orientation.

This paper examines that question by analysing the relationship between cultural diversity in UK research project teams and the realised interdisciplinarity of their publication outputs. It focuses on interdisciplinarity because this is a plausible mechanism through which diversity may matter for research. If culturally diverse teams contain broader cognitive and epistemic resources, they may be better positioned to identify connections across knowledge domains, combine different literatures and methods, and produce outputs that cross disciplinary boundaries. Prior work suggests that diverse scientific teams can generate significant outcomes through interdisciplinary collaboration, but also that integration can be difficult and may depend on trust, leadership and team processes (Specht and Crowston, 2022). The mechanism is therefore not automatic. Cultural diversity may create interdisciplinary potential, but that potential must be elaborated, translated and integrated before it becomes visible in research outputs.

The paper therefore distinguishes between interdisciplinary potential and realised interdisciplinary integration. Interdisciplinarity is not treated simply as the presence of multiple disciplines within a project. A project may include researchers from different disciplinary backgrounds without integrating their knowledge in the outputs it produces. Instead, interdisciplinarity is conceptualised as an output-level property: the extent to which project publications draw together knowledge from different disciplinary domains. This distinction follows bibliometric work that treats interdisciplinarity as a multidimensional form of diversity involving variety, balance and disparity (Stirling, 2007; Leydesdorff, 2018; Porter and Rafols, 2009). It is also important empirically because the analysis does not observe internal team processes directly. It observes whether projects with culturally diverse teams produce outputs that are more interdisciplinary in their realised publication profiles.

To develop this argument, the paper brings together three strands of literature. The first is the bibliometric literature on interdisciplinarity, especially work that conceptualises interdisciplinarity as a form of diversity involving the distribution of knowledge categories and the cognitive distance between them. This literature motivates the use of Rao-Stirling style indicators, which account not only for how many knowledge categories are represented, but also for how evenly they are represented and how distant they are from one another (Porter and Rafols, 2009; Leydesdorff, 2018). At the same time, measurement studies caution that interdisciplinarity indicators are sensitive to classification systems, distance measures and aggregation choices, so measurement should be treated as part of the research design rather than as a purely technical detail (Wang and Schneider, 2020; Nakhoda et al., 2023).

The second strand is research on cultural, demographic and cognitive diversity in teams. This literature suggests that diversity can broaden the pool of knowledge, perspectives and problem framings available to a team, but also that diversity-performance relationships are often weak, conditional or indirect. Meta-analytic evidence on culturally diverse teams suggests that deep-level diversity is more consistently associated with creativity and innovation than surface-level diversity, and that the effects depend on task and team conditions (Wang et al., 2019). Similarly, research on multicultural work groups suggests that cultural diversity does not have a simple direct effect on team performance, but works through process variables such as creativity, cohesion and conflict, and depends on contextual conditions such as task complexity, team tenure and co-location (Stahl and Maznevski, 2021). In research-team settings, Barjak and Robinson (2008) find that the most successful life-science teams have moderate levels of cultural diversity, while Specht and Crowston (2022) show that diversity can support scientific output through interdisciplinary collaboration but may also be associated with lower satisfaction and perceived effectiveness.

The third strand is the categorisation-elaboration model of team diversity, complemented by the too-much-of-a-good-thing perspective. The categorisation-elaboration model explains why diversity can have both positive and negative implications for team outcomes: diversity can support information elaboration by making a wider range of task-relevant perspectives available, but social categorisation can disrupt collaboration and inhibit the use of those perspectives (van Knippenberg, De Dreu and Homan, 2004). Recent meta-analytic work reinforces the need to study team processes rather than assuming a direct diversity effect: demographic diversity is more consistently linked to social categorisation processes than to information elaboration (Traylor et al., 2024). The too-much-of-a-good-thing perspective further suggests that a factor beneficial at low or moderate levels may become less beneficial or harmful at higher levels once associated costs outweigh marginal benefits (Pierce and Aguinis, 2013). Applied to this study, cultural diversity may increase interdisciplinary potential at low to moderate levels, but the relationship may weaken or reverse at higher levels if communication, translation and coordination demands rise faster than the benefits of additional cognitive and epistemic variety.

The empirical setting is UK Research and Innovation (UKRI) funded projects. The analysis links project records from Gateway to Research Plus (GtR+) to bibliometric information from OpenAlex. The unit of analysis is the funded research project. The setting also combines disciplinary breadth across six research councils, curated award-to-output links and an internationally recruited research workforce. This is a useful setting because projects provide an intermediate unit between individual publications and aggregated institutions: they have identifiable teams, funders, partners, timing and publication outputs. The analysis focuses on research grants with more than one team member and matched publication outputs, producing a main analytical sample of 2,795 projects.

Cultural diversity is measured as country-of-origin diversity within the research-active project team, using a distance-aware Leinster-Cobbold style index. This measure captures diversity in inferred origin composition at the aggregate team level and should not be interpreted as self-identified ethnicity, nationality or citizenship. Output interdisciplinarity is measured using publication-level Rao-Stirling indicators derived from OpenAlex disciplinary classifications and then aggregated to the project level. The main specification uses a subfield-level Rao-Stirling measure, with a broader field-level version used as a robustness check. The models also control for prior interdisciplinarity, measured from the prior publications of project PIs and CoIs in the five years before the project. This control is central because projects led by researchers with an established interdisciplinary profile may be more likely to produce interdisciplinary outputs irrespective of current team cultural diversity.

The analysis uses generalised additive models to allow the relationship between cultural diversity and output interdisciplinarity to be estimated flexibly. This modelling strategy is appropriate because the theoretical framework predicts a possible non-linear relationship rather than a simple linear association. The model sequence first estimates a baseline specification with core project controls, then adds cultural diversity, and finally adds prior interdisciplinarity. Supplementary specifications examine whether the diversity-interdisciplinarity relationship varies across selected project conditions.

The results support neither component of the hypothesis developed below. Cultural diversity is statistically associated with output interdisciplinarity before prior interdisciplinarity is introduced, but the association is small and negative rather than positive, the improvement in model fit is minimal, and the estimated smooth is close to linear, providing no evidence of the expected curvature. Once prior interdisciplinarity is added, the cultural diversity association becomes marginal. By contrast, prior interdisciplinarity is the strongest and most stable predictor of project-output interdisciplinarity, accounting for the main increase in model fit. Robustness checks using a broader field-level interdisciplinarity measure and an alternative similarity specification lead to the same substantive conclusion.

The paper makes three contributions. First, it clarifies the distinction between cultural diversity as interdisciplinary potential and interdisciplinarity as realised knowledge integration. This distinction helps avoid the assumption that diverse teams automatically produce interdisciplinary outputs. Second, it provides project-level evidence on one proposed mechanism linking cultural diversity to research outcomes. The findings suggest that interdisciplinarity is not an automatic pathway through which cultural diversity translates into research outputs. Third, it highlights the importance of prior interdisciplinary orientation. Realised interdisciplinarity appears strongly path-dependent: projects led by teams with prior interdisciplinary publication profiles are much more likely to produce interdisciplinary outputs. This finding qualifies optimistic accounts of diversity and interdisciplinarity by showing that diversity alone is not enough. The conversion of heterogeneous resources into integrated interdisciplinary outputs likely depends on team processes, leadership, coordination and institutional conditions that are not directly observed in the present data.

The remainder of the paper proceeds as follows. Section 2 develops the conceptual framework linking cultural diversity to realised interdisciplinarity. Section 3 describes the data, measures and modelling strategy. Section 4 presents the empirical results. Section 5 discusses the implications for research on diversity, interdisciplinarity and science policy, and outlines the limitations of the analysis.

2. Literature Review and Conceptual Framework

2.1 Interdisciplinarity as Realised Knowledge Integration

Interdisciplinarity has become an increasingly important feature of contemporary science because many research problems cannot be addressed from within a single disciplinary tradition. Complex societal, technological and scientific challenges often require researchers to bring together different bodies of knowledge, methods and standards of evidence. However, interdisciplinarity should not be equated with the mere presence of multiple disciplines within a project. A project may contain several disciplinary inputs without integrating them in a meaningful way. The key distinction is therefore between disciplinary variety and realised knowledge integration (Porter and Rafols, 2009; Leydesdorff et al., 2013; Soós and Kampis, 2012).

This distinction is central to the present study because the dependent variable captures the interdisciplinarity of project outputs rather than the disciplinary composition of the team. Interdisciplinarity is treated as an output-level property: the extent to which a project draws together knowledge from different disciplinary domains in its realised publications. Cultural diversity may increase the range of knowledge resources available to a team, but interdisciplinarity is realised only when those resources are elaborated, translated and combined in the outputs produced by the project (Carusi and Bianchi, 2020; Eykens et al., 2022; Nakhoda et al., 2025). The empirical question is therefore not whether culturally diverse teams contain heterogeneous backgrounds, but whether their outputs show greater integration across disciplinary knowledge domains.

The bibliometric literature supports this interpretation by distinguishing between different dimensions of diversity. Stirling (2007) identifies variety, balance and disparity as three core properties. Variety refers to the number of categories represented, balance to their distribution, and disparity to the distance between them. This matters because interdisciplinary integration is not captured by variety alone. A project drawing on several closely related fields may be less interdisciplinary than a project combining fewer but more distant fields, and a project dominated by one field may be less integrated than one in which contributions are more evenly distributed (Stirling, 2007; Leydesdorff, 2018; Zhang and Leydesdorff, 2021).

This logic underpins the use of Rao-Stirling style measures of interdisciplinarity. These indicators combine the distribution of knowledge categories with the distance between them. Porter and Rafols (2009) use this logic to examine disciplinary diversity in science, while Leydesdorff and colleagues show that interdisciplinarity indicators are sensitive to classification systems, distance matrices and levels of aggregation (Leydesdorff et al., 2013; Leydesdorff et al., 2019). Related work similarly shows that choices about field classification, citation direction, missing bibliographic information and distance measures can affect observed interdisciplinarity, which means that measurement should be treated as part of the research design rather than as a purely technical detail (Calatrava Moreno et al., 2016; Cassi et al., 2017; Wang and Schneider, 2020; Nakhoda et al., 2023).

The distinction between disciplinary variety and realised integration also clarifies how interdisciplinarity is related to research outcomes. Interdisciplinary knowledge and boundary-spanning combinations have been associated with novelty, disruption, technological impact and research performance, but these relationships depend on the capacity to combine knowledge across boundaries rather than on the mere presence of diverse inputs (Chen et al., 2022; Petersen et al., 2023; Yang, 2025; Wu et al., 2026). The present study therefore treats interdisciplinarity as realised integration in project outputs and asks whether culturally diverse research teams are more likely to convert heterogeneous cognitive and epistemic resources into such outputs.

This interpretation also clarifies the limits of the empirical test. The study uses publication-based measures of interdisciplinarity, calculated from disciplinary classifications and aggregated to the grant level. The empirical outcome should therefore be read as realised output interdisciplinarity rather than as a direct measure of internal team process. The data cannot directly observe whether team members exchanged ideas, resolved disagreement, built trust or developed shared understanding. Instead, the analysis examines whether the outputs associated with culturally diverse project teams show greater disciplinary integration. This makes the empirical design a reduced-form test of the broader theoretical argument.

2.2 Cultural Diversity as Interdisciplinary Potential

If interdisciplinarity is understood as realised knowledge integration, cultural diversity can be understood as a possible source of interdisciplinary potential. Researchers socialised in different cultural, national, educational or institutional contexts may bring different literatures, methodological preferences, problem framings, standards of evidence and tacit assumptions. These differences can broaden the team’s search space and increase the likelihood of identifying connections across knowledge domains that would be less visible in more homogeneous teams.

This argument should be stated carefully. Cultural diversity is not the same as cognitive diversity, and our measure, country-of-origin diversity, is not a direct measure of education, training, ethnicity, nationality, citizenship or lived experience. In this paper, origin diversity is treated as a proxy for possible heterogeneity in cultural, educational and epistemic exposure. To the extent that origin diversity is associated with different academic systems, institutional histories, intellectual traditions and social experiences, it may broaden the range of problem framings, literatures and interpretive repertoires available to a project team. The theoretical claim is therefore not that origin diversity directly measures cognitive diversity, but that it may capture one route through which research teams acquire heterogeneous cognitive and epistemic resources.

This argument is consistent with research showing that diversity can contribute to novelty and knowledge recombination. Hofstra et al. (2020) show that researchers from underrepresented groups are more likely to introduce novel conceptual linkages, even if those contributions receive less uptake. Mohammadi, Broström and Franzoni (2017) find that ethnic and educational diversity can support radical innovation, while Barjak and Robinson (2008) show that cultural diversity in research teams may have a curvilinear association with performance. These studies are not direct tests of origin diversity and grant-level interdisciplinarity, but they support the broader claim that cultural and demographic diversity can expand the range of ideas entering the research process.

Similar reasoning appears in studies of scientific collaboration and team-based knowledge production. Specht and Crowston (2022) show that diversity in scientific working groups can support interdisciplinary collaboration and publication outcomes, while Adachi et al. (2025) find that disciplinary variety in funded research teams is positively associated with research output in some funding contexts. Yang (2025) connects interdisciplinary knowledge and expertise diversity to disruptive scientific potential, and Liu et al. (2024) finds that research-interest diversity is an important predictor of innovative performance in artificial intelligence teams. Together, these studies suggest that teams with more varied knowledge resources may be better positioned to generate outputs that cross disciplinary boundaries.

The broader team-diversity literature reinforces this interpretation but also cautions against a simple positive account. Meta-analytic evidence on culturally diverse teams suggests that deep-level diversity is more consistently associated with creativity and innovation than surface-level diversity, and that these effects depend on task and team conditions (Wang et al., 2019). Research on multicultural work groups similarly suggests that cultural diversity often has indirect rather than direct effects, mediated by team processes such as creativity, cohesion and conflict and moderated by contextual conditions such as task complexity, team tenure and co-location (Stahl and Maznevski, 2021). These findings support the idea that diversity may provide useful resources, but that the use of those resources depends on collaboration conditions.

For the present study, the key implication is that cultural diversity may increase the raw material for interdisciplinarity. Origin-diverse project teams may contain members exposed to different academic systems, intellectual traditions, methodological norms and social experiences. These differences may influence which questions are considered important, which literatures are seen as relevant, which methods are viewed as legitimate and which disciplinary boundaries appear open to crossing. In this sense, cultural diversity may increase interdisciplinary potential by expanding the pool of cognitive and epistemic resources available to the project team.

The argument, however, remains conditional. Cultural diversity does not directly or automatically produce interdisciplinarity. Several studies show that diversity effects are mixed, conditional or non-linear. Specht and Crowston (2022) find that some forms of diversity support interdisciplinary collaboration, while country diversity can create difficulties for publication diversity and team satisfaction. Barjak and Robinson (2008) suggest that the benefits of cultural diversity may peak at moderate levels, and Liu et al. (2024) shows that diversity effects can depend on combinations of team characteristics. Network exposure may also matter, because culturally diverse teams may have access to wider academic communities, international contacts and institutional experiences. However, the main pathway developed here is cognitive and epistemic variety rather than brokerage alone. Cultural diversity may broaden what the team knows and how it frames problems, but without elaboration and integration this potential may remain latent.

2.3 Categorisation-Elaboration as a Conversion Theory

The previous section argued that cultural diversity may increase the raw material for interdisciplinarity by broadening the cognitive and epistemic resources available to research teams. However, the presence of heterogeneous resources does not by itself produce interdisciplinary outputs. Interdisciplinarity requires a process through which different ideas, methods and assumptions are exchanged, challenged, translated and combined. The categorisation-elaboration model provides a useful theoretical lens because it treats diversity as having both productive and disruptive possibilities (van Knippenberg, De Dreu and Homan, 2004; van Knippenberg, Li and Tu, 2024).

The positive side of this model is information elaboration. Diversity can improve team outcomes when members actively share, discuss and integrate different perspectives. Applied to interdisciplinary research, this means that cultural diversity can contribute to interdisciplinarity when different epistemic traditions, methodological repertoires and problem framings are elaborated into a common research output. Studies of team knowledge sharing and creativity support this logic, showing that diversity can be beneficial when connected to communication, knowledge exchange and inclusive team processes (Bodla et al., 2018; Sun et al., 2017; Kim and Song, 2021). In this sense, elaboration is the process through which diversity is converted from a compositional feature into a knowledge-production resource.

The negative side of the model is social categorisation. The same differences that create cognitive variety can also make collaboration more difficult. Cultural diversity may activate subgrouping, status differences, communication barriers or lower trust, especially when differences align with discipline, institution, seniority or project role. In such cases, diverse knowledge may remain fragmented rather than integrated. Research on diverse teams shows that surface-level diversity can reduce knowledge sharing when it triggers in-group and out-group distinctions, while psychological safety and inclusive climates can help teams avoid these problems (Watson et al., 2002; Settles et al., 2019; Bodla et al., 2018). Recent meta-analytic work also reinforces the need to study team processes rather than assuming a direct diversity effect, showing that demographic diversity is more consistently linked to social categorisation processes than to information elaboration (Traylor et al., 2024).

Categorisation-elaboration is therefore best understood here as a theory of conversion. Cultural diversity may create heterogeneous knowledge resources, but realised interdisciplinarity depends on whether teams convert those resources into integrated outputs. Where elaboration dominates, diversity can support cross-disciplinary recombination. Where categorisation dominates, diversity may remain latent or produce parallel disciplinary contributions rather than synthesis. This distinction is important because interdisciplinarity is not simply the coexistence of different forms of knowledge, but the integration of those forms into outputs that cross disciplinary boundaries (Monteiro and Keating, 2009; Oliver et al., 2018; Lewis et al., 2023; Pößneck et al., 2024).

Knowledge sharing, translation and coordination are central to this conversion process. Diverse teams may contain many useful ideas, but those ideas need to be made visible, debated and connected. Interdisciplinary research also requires translation between disciplinary languages, standards of evidence and expectations about what counts as a contribution. Cultural diversity may add further layers of translation because team members may differ in academic training, communication norms and assumptions about collaboration. Prior work on knowledge sharing, inclusive processes and boundary work therefore supports the argument that the value of diversity lies not only in team composition, but in the process through which teams use their differences (Monteiro and Keating, 2009; Oliver et al., 2018; Pößneck et al., 2024; Figueiredo Belchior and Chagas, 2026).

At the same time, this literature warns against assuming that more diversity is always better. High within-team knowledge diversity can create coordination costs when team members lack a shared cognitive framework, while ethnically diverse teams may face early interpersonal process demands before performance advantages emerge (Aggarwal, Hsu and Wu, 2020; Watson et al., 2002). Power, authorship norms and unequal voice can also shape whether diverse perspectives are incorporated into research outputs (Oliver et al., 2018; Lewis et al., 2023). Diversity can therefore create potential, but that potential can be blocked by weak elaboration, translation costs, dominance or coordination problems.

For the present study, the categorisation-elaboration model provides the central theoretical bridge between cultural diversity and realised interdisciplinarity. The empirical data cannot directly observe elaboration, psychological safety, dominance, trust, translation practices or knowledge sharing. These constructs provide the theoretical mechanism through which the observable association between cultural diversity and output interdisciplinarity is interpreted. The analysis can test whether culturally diverse project teams are associated with more interdisciplinary outputs, but the mechanism itself must be inferred cautiously from theory and supporting literature. This reduced-form interpretation prevents overclaiming and leads to the question of whether the relationship should be linear or non-linear.

2.4 Too Much of a Good Thing and the Non-Linear Relationship

The categorisation-elaboration model explains why cultural diversity can have both productive and disruptive effects, but it does not by itself require a linear relationship between diversity and interdisciplinarity. The too-much-of-a-good-thing perspective suggests that factors beneficial at low or moderate levels may become less beneficial, or even harmful, at higher levels when associated costs begin to outweigh marginal benefits (Pierce and Aguinis, 2013). Applied to the present study, this suggests that the relationship between cultural diversity and realised output interdisciplinarity may be non-linear rather than uniformly positive.

At low to moderate levels, cultural diversity is expected to increase interdisciplinary potential. Additional diversity may add non-redundant cognitive and epistemic resources to the team. Researchers from different cultural, national, educational or institutional backgrounds may bring different literatures, methods, problem framings and standards of evidence. This can widen the set of possible knowledge combinations and increase the likelihood that the team identifies connections across disciplinary boundaries (Hofstra et al., 2020; Barjak and Robinson, 2008; Specht and Crowston, 2022; Adachi et al., 2025; Yang, 2025).

However, the marginal benefit of additional cultural diversity is unlikely to increase without limit. Once a team already contains substantial cognitive and epistemic variety, further diversity may add less new information than earlier increments of diversity. At the same time, the demands of integration may rise. More culturally diverse teams may need to invest greater effort in developing shared language, aligning expectations, translating tacit assumptions and preventing subgroup formation. The same diversity that expands the pool of knowledge may therefore also increase the costs of elaboration and integration (van Knippenberg, De Dreu and Homan, 2004; van Knippenberg, Li and Tu, 2024).

This logic suggests that cultural diversity may have a positive association with interdisciplinarity at low to moderate levels, but that the association may flatten or decline at higher levels. A plateau would suggest saturation or diminishing marginal benefits: additional diversity no longer adds much interdisciplinary value because the team already contains substantial cognitive and epistemic variety. A downturn would provide stronger evidence for a too-much-of-a-good-thing mechanism, because it would suggest that coordination, translation and categorisation costs begin to dominate the benefits of additional diversity.

Prior studies support the expectation that diversity effects are conditional or non-linear. Cultural diversity in research teams may have curvilinear associations with performance, and different forms of diversity can have mixed effects on publication diversity, satisfaction and innovation (Barjak and Robinson, 2008; Specht and Crowston, 2022; Liu et al., 2024). Evidence from innovation and collaboration settings also suggests that the value of cultural distance depends on the type of boundary being crossed and on the conditions under which teams work (Brunetta et al., 2020; Haas and Nüesch, 2012; Byron et al., 2023; Chiang and Su, 2026). These studies reinforce the idea that cultural diversity should not be treated as uniformly beneficial or uniformly harmful.

The distinction between interdisciplinary potential and realised interdisciplinary integration is crucial here. Cultural diversity may increase the raw material for interdisciplinarity, but realised interdisciplinarity depends on whether teams can elaborate and integrate that material into research outputs. At low to moderate levels, the additional cognitive and epistemic resources may be easier to absorb and may increase the likelihood of cross-disciplinary recombination. At higher levels, the same diversity may place heavier demands on communication, coordination and translation. If those demands are not managed, the team may produce parallel contributions rather than integrated interdisciplinary outputs.

This argument does not imply that high cultural diversity is inherently problematic. Rather, it suggests that diversity produces potential whose realisation depends on whether teams can manage the increasing demands of elaboration and integration. The empirical question is therefore not simply whether cultural diversity has a positive or negative effect, but what shape the relationship takes across the observed range of diversity.

This leads to the paper's hypotheses, which separate the two components of the argument:

H1a. Cultural diversity is positively associated with the interdisciplinarity of research outputs at low to moderate levels of cultural diversity.

H1b. The association weakens or reverses at higher levels of cultural diversity, as marginal knowledge benefits diminish and translation, coordination and categorisation costs increase.

3. Data and Methods

3.1 Data source and analytical sample

The empirical analysis links project records from GtR+ with bibliometric information from OpenAlex. GtR+ is the enhanced release of UK Research and Innovation's Gateway to Research and records funded projects, participating organisations, named investigators and reported outcomes. The GtR+ dataset was accessed under the Innovation and Research Caucus (IRC) GtR+ Secondary Data Analysis scheme. Our dataset was constructed from the GtR+ people, participants and publication-outcome tables, together with a local OpenAlex database snapshot released on 29 January 2025. Additional inputs were used for specific parts of the workflow, including NamSor for name-based origin inference, the Research Organization Registry (ROR) for institutional matching, UN M49 subregions and geodesic-distance data for country similarity, and the PySciSci Python module for the calculation of Rao-Stirling interdisciplinarity indicators.

The unit of analysis is the funded research project. This unit is appropriate because the paper asks whether culturally diverse research teams produce more interdisciplinary outputs. Projects provide an intermediate level between individual publications and aggregated institutions: they have identifiable teams, participating organisations, funders, partners, start dates and reported outputs. The dataset therefore allows team composition, prior research orientation and realised publication interdisciplinarity to be linked at the project level.

The analysis includes projects funded by the main UKRI research councils: AHRC, BBSRC, EPSRC, ESRC, MRC and NERC. STFC-funded projects are not included. This reflects the structure of STFC funding rather than a substantive judgement: a large share of STFC support takes the form of facility operation, infrastructure and international subscriptions rather than investigator-led project grants, and outputs in STFC-dominated fields frequently arise from very large multi-institution collaborations, for which the research-active project team defined below is not a meaningful unit of diversity measurement.

UK Research Council funding provides a suitable setting for several reasons. The six councils included in the analysis span the disciplinary spectrum from the arts and humanities to the engineering, environmental and medical sciences, so disciplinary context varies widely within a single funding system while the policy and institutional environment is held broadly constant. GtR+ records publication outcomes reported by grant holders against specific awards, which supplies the explicit project-to-output link the design requires and which few funding systems make available at scale. The UK research workforce is also recruited internationally, so funded teams display meaningful variation in origin composition, as the descriptive statistics in Section 4.1 confirm. The setting is policy-relevant in addition, since UKRI operates cross-council schemes intended to encourage research that crosses disciplinary boundaries and monitors the diversity of the research workforce as a matter of policy.

The source data cover projects with start years between 2021 and 2024. GtR+ publication outcomes are used as the starting point for identifying project outputs, rather than OpenAlex funding-acknowledgement links. This choice is deliberate. GtR+ outcomes are reported against specific awards and therefore provide an explicit project-output link. Funding acknowledgements in OpenAlex, by contrast, may be missing, ambiguously parsed, unevenly covered across publishers and disciplines, or difficult to assign to a specific award when grant identifiers are incomplete. OpenAlex is therefore used downstream to resolve the GtR+ reported outputs to publication metadata, author lists, affiliations and disciplinary classifications.

The sample construction proceeds in stages, summarised here because the resulting restrictions define the scope of inference. First, 31,754 projects met the funding and start-year criteria. Of these, 6,775 (21.3%) recorded at least one publication outcome in GtR+. Publication outcomes were then matched to OpenAlex using cleaned Digital Object Identifiers (DOIs), supplemented by exact matching of cleaned publication titles where DOI matching was unavailable or unsuccessful. Projects without matched publication outputs were excluded because the dependent variable is publication-based output interdisciplinarity and cannot be computed without matched publication data. After publication matching and filtering, 4,662 projects (14.7% of the eligible population) remained in our dataset.

For the present analysis, the sample is further restricted to Research Grant projects, the standard investigator-led project mode, and to projects with more than one research-active team member. Fellowships and studentships are organised around an individual rather than a team, and centre, infrastructure and training grants do not define a bounded project team. Single-person projects are excluded because the theoretical argument concerns team diversity and the conversion of heterogeneous knowledge resources into interdisciplinary outputs. The main models also require complete data on output interdisciplinarity, cultural diversity, prior interdisciplinarity and the control variables. These restrictions leave 3,261 projects. The main models also require complete data on output interdisciplinarity, cultural diversity, prior interdisciplinarity and the control variables; after this complete-case restriction, the main analytical sample contains 2,795 projects, 8.8% of the eligible population.

The resulting sample should be interpreted as the subset of UK Research Council funded projects with matched publication outputs and sufficient bibliometric information for the construction of output and prior interdisciplinarity measures.

The direction of any resulting selection effect on the diversity-interdisciplinarity association cannot be established from the observed data, and inference is therefore confined to the publication-reporting, multi-member Research Grant population defined above.

3.2 Construction of project teams

Project teams are constructed by linking GtR+ participating organisations and publication outcomes to OpenAlex author and affiliation records. The team definition distinguishes between three groups of individuals associated with project outputs.

First, named authors are individuals formally listed on the grant in GtR+, including named applicants and investigators. Second, unnamed authors are individuals who are not listed on the grant but appear as authors on the project’s matched publication outputs and are affiliated with one of the project’s participating organisations. Third, collaborators are individuals who appear as authors on matched project outputs but are affiliated only with organisations outside the set of participating organisations.

The project team is defined as the union of named and unnamed authors. This broader definition is used because formal grant records may understate the set of researchers who contributed to the funded work. Research assistants, postdoctoral researchers, doctoral researchers and other staff may not be named on the original award but may nevertheless conduct project work and appear on resulting outputs. Including unnamed authors affiliated with participating organisations therefore captures the research-active project team supported by the grant more fully than the named-applicant list alone.

Collaborators are not included in the core team-diversity measure. This is because they are affiliated only with organisations outside the project’s participating organisations and therefore represent the wider collaboration environment rather than the internal research-active team at participating organisations. Their presence is relevant to the project’s external reach, but including them in the main team-diversity measure would conflate internal team composition with broader collaboration networks.

Affiliation matching was used to distinguish unnamed team members from collaborators. Authors were first matched to participating organisations using exact matches on Research Organization Registry (ROR) identifiers. For organisations without ROR identifiers, fuzzy matching on organisation names was used after removing generic institutional terms and common words. Authors resolving to a participating organisation were classified as unnamed authors if they were not already named on the grant; remaining output authors were classified as collaborators.

This team-construction strategy is appropriate for the paper's research question because the study is concerned with the diversity of the research-active team associated with a project's realised outputs. However, it also implies an important interpretive caveat. Because some team members are inferred from publication outputs, the team-diversity measure captures the realised research-active project team rather than only the ex-ante applicant team. The analysis should therefore be interpreted as an association between the diversity of realised project teams and the interdisciplinarity of project outputs, not as a causal estimate based solely on team composition fixed at the application stage.

3.3 Measures

3.3.1 Output interdisciplinarity

The dependent variable is the interdisciplinarity of project outputs. This measure captures realised interdisciplinarity at the publication-output level rather than the intended disciplinary scope of the funded project. This distinction is important because the paper conceptualises interdisciplinarity as realised knowledge integration, not simply as the presence of multiple disciplinary inputs. A project may involve researchers from different fields, but its outputs are interdisciplinary only to the extent that they draw together knowledge from different disciplinary domains in the published record (Stirling, 2007; Porter and Rafols, 2009; Leydesdorff et al., 2013).

The analysis uses publication-based interdisciplinarity rather than GtR+ project-topic classifications. GtR+ topic classifications are available more broadly and describe the project at award level, but they are coarse administrative categories and do not provide the disciplinary distance structure required for a distance-aware interdisciplinarity measure. Publication-based indicators, by contrast, capture realised research outputs and allow the disciplinary profile of each publication to be measured using OpenAlex subject information. This choice is consistent with the paper's focus on realised output interdisciplinarity.

Interdisciplinarity is measured using publication-level Rao-Stirling indicators derived from OpenAlex disciplinary information. Rao-Stirling diversity incorporates three dimensions of diversity: variety, balance and disparity. Variety captures the number of disciplinary categories represented, balance captures the distribution of activity across those categories, and disparity captures the distance between them. This is preferable to a simple count of fields because realised interdisciplinarity is not only about how many fields are present, but also about how cognitively distant those fields are from one another (Stirling, 2007; Leydesdorff, 2018; Zhang and Leydesdorff, 2021).

For each matched publication, subject information was retrieved from OpenAlex. OpenAlex organises subject information across three hierarchical levels: fields, subfields and topics. The publication-level Rao-Stirling index was computed using the PySciSci implementation. The disciplinary profile of a publication was defined from its references, so that a publication's disciplinary position is inferred from the literature it draws on. Where a publication was assigned to more than one category, publication-to-field weights were normalised to sum to one. Defining the disciplinary profile from references reflects the conceptualisation of interdisciplinarity as the integration of knowledge inputs: a publication's position is inferred from the literature it draws on rather than from the classification of its venue or its own topic labels alone.

The main outcome measures interdisciplinarity at the OpenAlex subfield level, using cosine-based distances between subfields and a time-invariant distance matrix. The subfield level is preferred because it is granular enough to register integration across substantively distinct bodies of knowledge, while avoiding the fragmentation of the highly specific topic level, at which routine within-specialty combinations would register as interdisciplinarity. This specification corresponds to Scenario 7 of the interdisciplinarity scenarios released with our dataset; the correspondence between the measures used in this paper and the released scenario codes is documented in Appendix Table A1.

As robustness checks, the models are re-estimated using a broader field-level measure constructed under the same non-temporal cosine logic (Appendix A3), and using an alternative index based on a wider similarity matrix (Appendix A4). These provide coarser and alternative tests of whether the findings depend on the level of disciplinary classification or on the similarity structure used. The main interpretation does not rest on the specific subfield-level specification if these checks produce the same substantive conclusion.

Because Rao-Stirling interdisciplinarity is calculated at the publication level, project-level interdisciplinarity is obtained by aggregating publication-level scores across the matched outputs of each project. The main analysis uses the project-level mean. This means the dependent variable should be interpreted as the average interdisciplinarity of a project's realised publication outputs. The number of matched publications entering the calculation is recorded separately because project-level means are less stable for projects with very few matched outputs. This instability is considered in the interpretation of the results and is noted as a limitation in Section 5.

3.3.2 Cultural diversity

The main independent variable is cultural diversity within the project team. Cultural diversity is operationalised using the country-of-origin diversity of research-active project team members. Country of origin is inferred from author names using proprietary tool NamSor and then aggregated to the team level. This measure should not be interpreted as self-identified ethnicity, nationality, citizenship or legal migration background. Rather, it captures a probabilistic origin signal derived from names and used only to construct aggregate project-team diversity indicators. No individual-level claim is made about any author's origin, ethnicity, nationality or identity.

This distinction is important because name-based inference can contain uncertainty and bias, especially across different naming traditions (Kozlowski et al., 2022; Lockhart, King and Munsch, 2023; Hafner, Peifer and Hafner, 2024). The use of inferred origin is justified here only because the analysis is conducted at the aggregate project-team level, because predictions are filtered using confidence rules, and because the selected diversity index incorporates similarity between origin categories. Misclassification between neighbouring or onomastically similar countries is therefore less consequential than it would be under a purely categorical origin measure, although it remains a limitation of the analysis. A fuller treatment of the relative merits and limitations of inference-based, self-identification and administrative approaches to diversity data is provided in Gök et al. (2026).

The name-processing and inference pipeline follows Gök and Macmillan (in press), where it is documented in full and externally benchmarked. OpenAlex author display names are cleaned and parsed into forename and surname components; names without usable forename-surname information, names reduced to initials, and evidently corrupted strings are excluded before inference. Cleaned names are submitted to NamSor, which returns probabilistic country-of-origin predictions and associated confidence scores. Predictions are retained under two rules. First, individuals are retained where the first country-of-origin prediction has a confidence score above 0.50. Second, individuals with a lower first-country score are retained where the first and second predicted countries together exceed 0.66 and both countries fall within the same UN M49 subregion. The rationale for the second rule is that some residual uncertainty concerns neighbouring or culturally proximate countries with similar naming conventions; because the diversity index uses geodesic-distance-based similarity between countries, such cases introduce less spurious diversity than they would under a categorical treatment of origin. Threshold choice trades misclassification against non-classification (Sebo, 2023), and the thresholds used here follow the precision-recall rationale set out in Gök and Macmillan (in press). Corresponding classification and retention counts for the present project population are not reported here; as an indication of the pipeline's behaviour, Gök and Macmillan (in press) report that, in a comparable OpenAlex-derived population of UK research communities, 88.1% of unique author names satisfied the name-quality criterion and 8% of retained individuals were excluded from origin inference by the confidence rules.

Cultural diversity is measured using a Leinster and Cobbold (2012) index that combines the number of inferred origins present in the research-active team, their relative abundance and the similarity between them, with sensitivity parameter $q = 1$. Similarity between origins is based on population-weighted geodesic distance between countries, transformed exponentially with decay parameter $t = 0.05$ so that the measure remains sensitive to differences among relatively proximate origins; relative abundances are reweighted by the inverse log-frequency of each origin in the wider dataset, so that globally rarer origins contribute more to measured diversity, reflecting the disproportionate informational contribution of rare, potentially boundary-spanning perspectives. A distance-aware index is preferred over categorical alternatives on both theoretical and measurement grounds: a team drawn from closely related origin contexts is not equivalent to a team drawn from more distant contexts even if both contain the same number of origin categories, and misclassification between onomastically and geographically proximate countries adds little spurious diversity when such countries are treated as highly similar. The conceptual foundations of these choices are set out in Gök et al. (2026). This specification corresponds to Scenario 9b of the origin-diversity scenarios defined in Gök and Macmillan (in press); it retains the substantive logic of the closely related Scenario 9a ($t = 0.1$) while applying the smaller decay parameter, which further increases sensitivity to differences among proximate origins. The full parameter mapping is given in Appendix Table A1.

3.3.3 Prior interdisciplinarity

Prior interdisciplinarity is included as a key control variable. This is important because projects led by researchers with an established interdisciplinary publication profile may be more likely both to form diverse teams and to produce interdisciplinary outputs. Without controlling for prior interdisciplinarity, the estimated relationship between cultural diversity and output interdisciplinarity could partly reflect endogenous pre-existing research orientation rather than the composition of the current project team.

Prior interdisciplinarity is measured using the previous publications of project PIs and CoIs in the five years before the project start. PIs, CoIs and research CoIs are identified from the GtR+ people table and linked to OpenAlex author records. For each project, prior publications are linked to the relevant publication-level interdisciplinarity scores, and the mean prior interdisciplinarity score is calculated for the PI/CoI team.

The prior score used in each model is constructed under the same interdisciplinarity specification as the dependent variable: models of the subfield-level outcome control for subfield-level prior interdisciplinarity, and the field-level robustness models control for field-level prior scores. This ensures that the prior-interdisciplinarity control is measured on the same disciplinary scale as the output-interdisciplinarity outcome.

Prior interdisciplinarity does not capture all PI- or team-level characteristics that may shape project formation and output interdisciplinarity, such as career stage, institutional position, collaboration history, disciplinary identity or leadership experience. It is therefore interpreted as a focused control for pre-existing interdisciplinary orientation rather than as a comprehensive PI-profile measure. Including it makes the empirical test more conservative by estimating whether cultural diversity is associated with output interdisciplinarity over and above the prior interdisciplinary trajectory of the PI/CoI team.

3.3.4 Controls

The models include a concise set of controls selected to account for observable project, funder and team characteristics while keeping the specification interpretable. The main control set includes delivery body, project start year, team size and non-UK partnership. Additional project characteristics are used in supplementary models to examine whether the cultural-diversity relationship differs across observable project conditions.

Project start year is included as a control. It enters the main models as a linear term; a specification with calendar-year fixed effects is examined as a supplementary model (M7, Appendix Table A2). This accounts for cohort differences between projects starting in different calendar years. Such differences may reflect variation in publication windows, OpenAlex coverage, funding priorities or the timing of output production. This is particularly important because the dependent variable is based on realised publication outputs.

Delivery body is included as a factor variable. Delivery body captures the Research Council context of the award, with categories including AHRC, BBSRC, EPSRC, ESRC, MRC and NERC. This is important because disciplinary scope, collaboration patterns and publication practices vary substantially across Research Councils. EPSRC is used as the reference category in the final model table.

Team size is measured as the number of research-active authors associated with the project team. Because team size is highly skewed, the models use log team size. This accounts for the possibility that larger teams may have greater disciplinary breadth or greater capacity to produce outputs spanning multiple knowledge domains.

The models also include an indicator for whether the project includes a non-UK partner. This captures a basic form of international collaboration or international project reach. It is included because internationally connected projects may differ in both team composition and output interdisciplinarity.

Supplementary specifications examine whether the cultural diversity-interdisciplinarity relationship varies with non-university partnership, first PI grant status, spend per month and project duration. These are not treated as formal moderation hypotheses. They involve several comparisons without formal correction and are therefore read as exploratory checks of whether the estimated diversity relationship is sensitive to observable project conditions. Together, the controls and supplementary specifications allow the analysis to adjust for major observable differences between projects.

3.4 Modelling strategy

The empirical analysis uses generalised additive models to estimate the association between cultural diversity and realised output interdisciplinarity. This modelling strategy is appropriate because the theoretical framework does not assume that the relationship between cultural diversity and interdisciplinarity is necessarily linear. The too-much-of-a-good-thing argument developed in Section 2 suggests that cultural diversity may increase interdisciplinarity at low to moderate levels, but that the association may weaken or flatten at higher levels if translation, coordination and categorisation costs increase. The models therefore include cultural diversity as a smooth term, allowing the shape of the relationship to be estimated from the data rather than imposed through a linear or quadratic specification. The models are Gaussian generalised additive models with an identity link, estimated by penalised likelihood using the mgcv package in R (Wood, 2011).

The main dependent variable is the project-level mean of output interdisciplinarity measured at the subfield level, with the broader field-level measure used as a robustness check (Appendix A3). Each model uses the corresponding prior interdisciplinarity measure as a control, so that outcomes and prior scores are always measured under the same specification.

The models are estimated on the complete-case analytical sample of Research Grant projects with more than one research-active project-team member. Projects without matched publication outputs or without the required prior interdisciplinarity information are excluded. The final main models contain 2,795 projects.

The model sequence is progressive. Model 0 includes the baseline controls: delivery body, project start year entered as a linear term, log team size and non-UK partnership. Model 1 adds cultural diversity as a smooth term. Model 2 adds prior interdisciplinarity as a smooth term. Models 3 to 6 examine whether the cultural-diversity smooth varies across selected observable project conditions: non-university partnership, first PI grant status, spend per month and project duration; these models omit the start-year term, so their fit statistics are not strictly nested comparisons with Model 2. Model 7 re-estimates Model 2 with project start year entered as calendar-year fixed effects. These models are not presented as formal moderation tests. They are used as supplementary specifications to assess whether the estimated cultural-diversity relationship is sensitive to observable project conditions.

The models are interpreted through the estimated smooth terms, effective degrees of freedom, approximate p-values, deviance explained, AIC and diagnostics. Effective degrees of freedom are important because they indicate the complexity of the estimated smooth. If the relationship took the form proposed in H1b, the smooth for cultural diversity would have an effective degrees of freedom clearly above one and would show a substantively meaningful change in slope across the observed diversity range. An effective degrees of freedom close to one indicates an approximately linear relationship, consistent with neither a plateau nor a downturn.

The empirical models are interpreted as associational rather than causal. The data do not directly observe internal team processes such as elaboration, psychological safety, translation, dominance or knowledge-sharing practices. The models instead test whether culturally diverse project teams are associated with more interdisciplinary outputs after accounting for prior interdisciplinary orientation and observable project characteristics. This distinction is important because a weak or non-significant diversity effect does not imply that cultural diversity has no theoretical value. It may instead indicate that cultural diversity creates interdisciplinary potential that is not consistently converted into realised interdisciplinary outputs.

Robustness checks examine whether the main conclusion depends on the disciplinary granularity and similarity structure used to measure interdisciplinarity. Because indicators of this kind are sensitive to assumptions about classification, distance and weighting, and the choice of specification can alter the magnitude, sign and significance of estimated relationships (Gök et al., 2026), the model sequence is re-estimated under a broader field-level Rao-Stirling measure and under an alternative, wider similarity matrix (Appendices A3 and A4). The substantive conclusions are unchanged. The origin-diversity specification is held fixed throughout, and sensitivity to alternative diversity specifications is acknowledged as a limitation in Section 5.

4. Results

4.1 Descriptive statistics

Table 1 reports descriptive statistics for the main analytical sample. The sample contains 2,795 Research Grant projects with more than one research-active team member, matched publication outputs and complete data for the main modelling variables. Output interdisciplinarity has a mean of 0.230, while prior interdisciplinarity has a similar mean of 0.234, indicating that the observed level of project-output interdisciplinarity is close to the prior interdisciplinary profile of the PI/CoI team. Cultural diversity varies meaningfully across projects, with a mean of 1.394 and a range from 1.000 to 2.138.

Team size is strongly right-skewed. The median project team contains 6 research-active members, while the mean is 10.303 and the maximum is 609. The models therefore use log team size. The categorical variables show substantial variation across delivery bodies, international partnership, non-university partnership, first PI grant status, spend intensity and project duration. EPSRC-funded projects form the largest group, but all six included Research Councils are represented. Around half of the projects include a non-UK partner, and just over half include a non-university partner.

Table 1. Descriptive statistics

A. Continuous variables

Variable	N	Mean	SD	Median	Min	Max
Output interdisciplinarity	2,795	0.230	0.047	0.234	0	0.35
Cultural diversity	2,795	1.394	0.283	1.423	1	2.138
Prior interdisciplinarity	2,795	0.234	0.034	0.238	0.082	0.333
Team size	2,795	10.303	20.852	6	2	609
Log team size	2,795	2.046	0.743	1.946	1.099	6.413

B. Categorical variables

Variable	N	Distribution
Delivery body	2,795	AHRC: 103 (3.7%); BBSRC: 541 (19.4%); EPSRC: 1,158 (41.4%); ESRC: 189 (6.8%); MRC: 484 (17.3%); NERC: 320 (11.4%)
Project start year	2,795	2021: 1,144 (40.9%); 2022: 931 (33.3%); 2023: 584 (20.9%); 2024: 136 (4.9%)
Includes non-UK partner	2,795	No: 1,397 (50.0%); Yes: 1,398 (50.0%)
Includes non-university partner	2,795	No: 1,244 (44.5%); Yes: 1,551 (55.5%)
First PI grant	2,795	No: 1,291 (46.2%); Yes: 1,504 (53.8%)
Spend per month	2,795	<£10k: 892 (31.9%); £10k-£20k: 1,133 (40.5%); >£20k: 770 (27.5%)

4.2 Correlations

Table 2 reports bivariate correlations between output interdisciplinarity, cultural diversity and prior interdisciplinarity. Cultural diversity is weakly and negatively correlated with output interdisciplinarity, but the magnitude is very small, $r = -0.048$. This offers no descriptive support for the positive low-to-moderate association anticipated in H1a, although a correlation computed across the full range of diversity cannot by itself speak to the non-linear shape proposed in H1b; the direct test is provided by the models below. By contrast, prior interdisciplinarity is strongly and positively correlated with output interdisciplinarity, $r = 0.683$. The correlation between cultural diversity and prior interdisciplinarity is also small, $r = -0.056$, indicating that the two variables capture different aspects of the project team.

The correlation pattern therefore anticipates the main modelling result. Cultural diversity has only a small bivariate association with realised output interdisciplinarity, while prior interdisciplinary orientation is strongly associated with the interdisciplinarity of project outputs.

Table 2. Correlation matrix

Variable	Output interdisciplinarity	Cultural diversity	Prior interdisciplinarity
Output interdisciplinarity	1		
Cultural diversity	-0.048*	1	
Prior interdisciplinarity	0.683***	-0.056**	1

Note: *p < .05; **p < .01; ***p < .001.

4.3 GAM results

Table 3 summarises the main GAM results for the subfield-level outcome. The full model table, including parametric coefficients and supplementary specifications, is reported in Appendix Table A2. The model sequence evaluates whether cultural diversity is associated with realised output interdisciplinarity and whether this association remains after accounting for prior interdisciplinarity and observable project characteristics.

The baseline model, M0, explains 9.03% of the deviance. Adding cultural diversity provides only a very small improvement in model fit. In the progressive summaries, the diversity-only model explains less than 0.2% of the deviance for both the subfield-level and field-level measures. The cultural-diversity smooth is statistically detectable in the simplest models, but its effective degrees of freedom is approximately 1, indicating an essentially linear rather than non-linear relationship. The estimated association is negative rather than positive: moving across the interquartile range of cultural diversity corresponds to a predicted change of −0.0074 in output interdisciplinarity (Table 3), around 0.16 standard deviations of the outcome. This result therefore does not support the positive low-to-moderate association anticipated by H1a, and the essentially linear smooth provides no evidence of the curvature anticipated by H1b.

The main change occurs when prior interdisciplinarity is introduced. Prior interdisciplinarity is highly significant and accounts for the major improvement in model fit. For the subfield-level outcome, the cultural-diversity smooth is no longer significant at conventional levels once prior interdisciplinarity is included ($p = .099$); the point estimate remains slightly negative but the 95% confidence interval spans zero. For the field-level outcome, the smooth weakens to $p = .062$ in the fully adjusted model. In both cases the residual association is best described as marginal. The robust pattern is therefore not a stable cultural-diversity association, but the dominance of prior interdisciplinary orientation.

Supplementary by-group smooth models are reported in Appendix Table A2. These models examine whether the cultural-diversity smooth varies by non-university partnership, first PI grant status, spend per month and project duration. They provide limited evidence of systematic heterogeneity. Some subgroup smooths are marginal or significant, but the pattern is not consistent and the duration model is affected by high concurvity. As noted in Section 3.3.4, these specifications involve multiple comparisons without formal correction; they are therefore treated as exploratory robustness checks rather than as evidence for a distinct moderation mechanism.

Overall, the results support neither H1a nor H1b. The estimated diversity smooths are close to linear, the explanatory contribution of cultural diversity is very small, and the association is not robust once prior interdisciplinarity is included. The evidence therefore does not support a positive diversity-to-interdisciplinarity conversion mechanism in this setting.

Table 3. Main GAM results (subfield-level outcome)

Term / statistic	M0 Baseline	M1 + Diversity	M2 + Prior interdisciplinarity
Cultural diversity smooth	—	edf = 1.00; p < .001	edf = 1.00; p = .099
Prior interdisciplinarity smooth	—	—	edf = 3.30; p < .001
Log team size smooth	edf = 1.00; p = .360	edf = 1.00; p = .008	edf = 2.21; p = .120
Delivery body control	Yes	Yes	Yes
Project start-year control	Yes	Yes	Yes
Non-UK partner control	Yes	Yes	Yes
Predicted change across IQR of cultural diversity	—	-0.0074 (95% CI: -0.0110 to -0.0038).	-0.0023 (95% CI: -0.0051 to 0.0005).
N	2,795	2,795	2,795
AIC	-9,352.00	-9,366.40	-10,959.70
Deviance explained (%)	9.03	9.57	49.08
Δ deviance explained (%)	—	0.53	39.51

Note: Smooth terms report effective degrees of freedom and approximate p-values. Significance codes: . p < .10; *p < .05; **p < .01; ***p < .001. Full coefficient estimates and supplementary models M3-M6 are reported in Appendix Table A2.

Figure 1 presents the estimated smooth terms for cultural diversity from M1 and M2, with 95% confidence intervals. Consistent with the correlations reported above, the estimated relationship is essentially linear in M1 and attenuates towards flatness in M2.

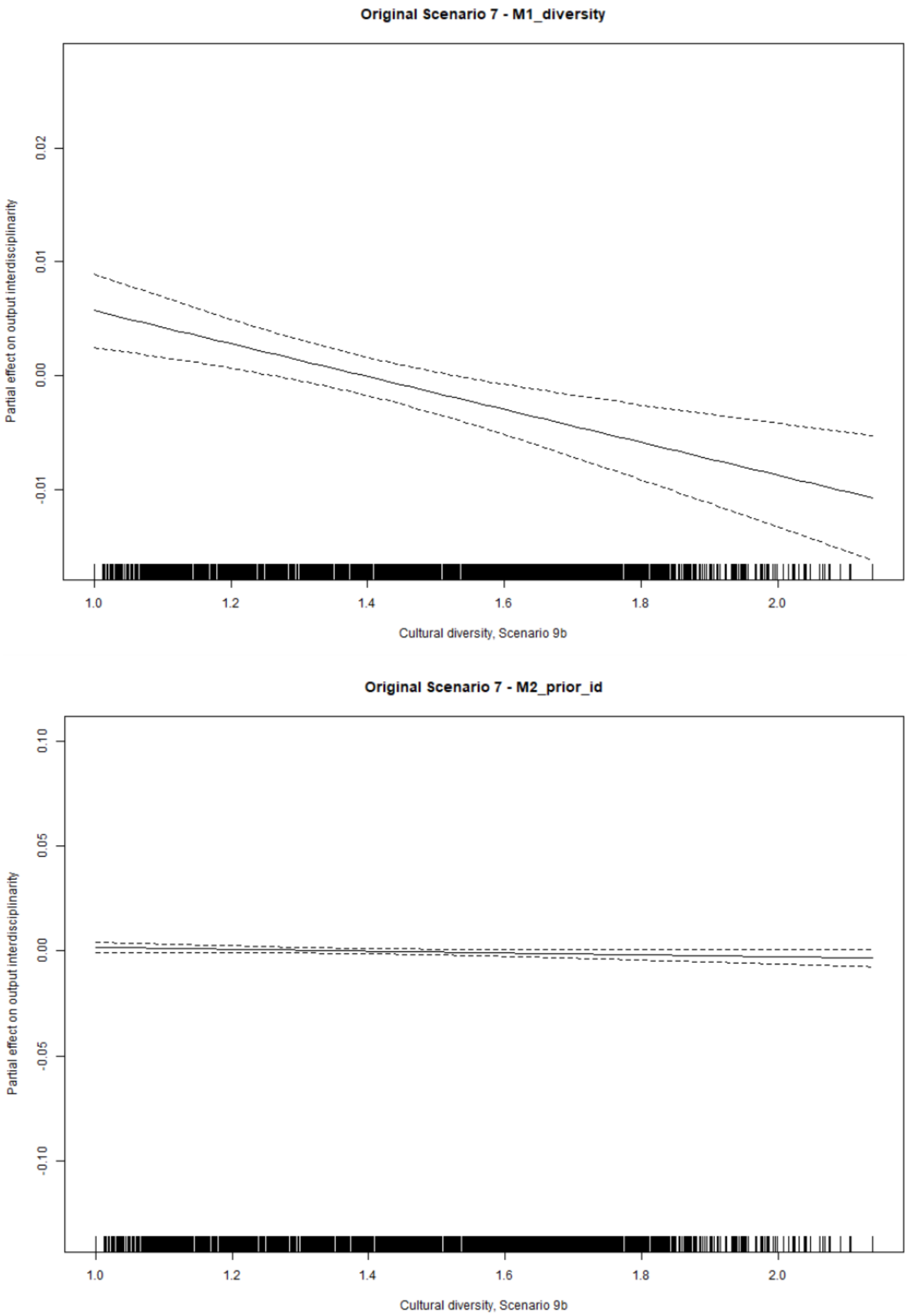


Figure 1. Estimated cultural-diversity smooths with 95% confidence intervals for the subfield-level outcome. (a) M1 before prior interdisciplinarity is included; panel (b) M2 after prior interdisciplinarity is included.

5. Discussion and Conclusion

5.1 Main findings

This study examined whether cultural diversity in research project teams, operationalised as inferred country-of-origin diversity, is associated with the interdisciplinarity of their publication outputs. The theoretical argument distinguished the heterogeneous resources that diversity may make available from their integration into interdisciplinary outputs. It further anticipated that the relationship might weaken or reverse at higher levels of diversity if the demands of communication, translation and coordination began to outweigh the benefits of additional variety. The results support neither H1a nor H1b. In M1, which provides a benchmark estimate without controlling for prior interdisciplinarity, origin diversity has a small negative association with output interdisciplinarity and contributes little additional explanatory power. The estimated smooth is approximately linear, providing no evidence of the expected curvilinear relationship. However, M1 remains vulnerable to confounding by pre-existing research orientation: investigators with established interdisciplinary trajectories may differ systematically in both the teams they form and the outputs they subsequently produce.

M2 is therefore the preferred specification because it controls for the prior interdisciplinarity of the PI/CoI team and thereby addresses one important source of endogeneity. In this model, the estimated association between origin diversity and subfield-level output interdisciplinarity is small and imprecisely estimated, with a confidence interval spanning zero. The field-level robustness model produces similarly weak evidence. The appropriate interpretation is therefore not that origin diversity reduces interdisciplinarity, but that the analysis provides little evidence of the hypothesised positive association once pre-existing interdisciplinary orientation and observable project characteristics are taken into account. The negative association in M1 should not be overinterpreted: its magnitude is small, it attenuates substantially in M2, and the preferred specification does not distinguish it reliably from zero.

The most substantial empirical result is the strong association between prior and subsequent interdisciplinarity. Projects led by investigators with more interdisciplinary prior publication profiles tend to produce more interdisciplinary outputs. This result indicates substantial persistence in interdisciplinary orientation, rather than establishing path dependence in a stronger causal sense. Such persistence may reflect accumulated experience of combining disciplinary knowledge, stable intellectual preferences, established collaboration networks, disciplinary location, publication practices or the selection of particular kinds of projects. Moreover, prior and subsequent interdisciplinarity are measured using closely aligned bibliometric constructions, and the prior measure covers individuals who also contribute to project outputs. Part of the increase in model fit is therefore expected on grounds of temporal and measurement continuity.

These findings nevertheless clarify the relationship between team composition and interdisciplinary output. They suggest that the presence of origin heterogeneity does not, by itself, predict greater bibliometric interdisciplinarity after pre-existing orientation is considered. This is consistent with the wider team-diversity literature, in which diversity effects are generally small, conditional or indirect, and depend on whether heterogeneous perspectives are exchanged and integrated effectively (Stahl and Maznevski, 2021; Wallrich et al., 2024). The categorisation-elaboration model similarly proposes that diversity can expand the information available to a team while also creating social categorisation, communication and coordination challenges that inhibit its use (van Knippenberg, De Dreu and Homan, 2004; Traylor et al., 2024).

The present evidence cannot identify which of these processes occurred. It does not observe information elaboration, psychological safety, leadership, epistemic disagreement, dominance, trust or knowledge-sharing practices. Nor does inferred origin diversity directly measure the cognitive, epistemic or disciplinary resources held by individual team members. The findings are therefore consistent with the proposition that compositional diversity and the capacity to integrate heterogeneous knowledge are distinct, but they do not demonstrate that diverse teams possessed relevant knowledge that subsequently failed to be converted. This distinction is important both theoretically and for the design of research-funding policy.

5.2 Implications for research-funding policy

The first policy implication is that diversity and interdisciplinarity should be treated as distinct objectives. The rationale for promoting diversity in research extends beyond its possible contribution to interdisciplinarity or other measurable research outputs. Equity of access, fair representation, inclusion, legitimacy and the distribution of opportunities provide independent reasons for addressing inequalities in research participation. The weak incremental association observed here does not diminish those rationales. Equally, diversity initiatives should not be judged unsuccessful because origin-diverse teams do not necessarily produce publications with higher bibliometric interdisciplinarity. Treating diversity primarily as an instrument for producing a particular form of output would make its policy legitimacy contingent on performance effects that may be small, context-dependent or expressed through mechanisms not captured by publication-based indicators.

Conversely, policies intended specifically to promote interdisciplinary research should not assume that compositional diversity will automatically produce knowledge integration. Funders frequently encourage applicants to assemble teams spanning different backgrounds, organisations or areas of expertise. Such team formation may broaden the resources available to a project, but the present findings suggest that composition alone is an unreliable indicator of the interdisciplinarity that will appear in its outputs. Where interdisciplinary integration is an explicit programme objective, assessment and support may therefore need to extend beyond who is present in the team to how the proposed collaboration will work.

At the application stage, funders could ask teams to articulate a credible integration strategy rather than relying principally on counts of represented disciplines, organisations or demographic groups. Such a strategy might explain how the research problem will be formulated jointly, how different methods or standards of evidence will be reconciled, how substantive disagreements will be handled, and how intellectual contributions will be recognised. It could also identify the leadership and decision-making arrangements through which team members will participate in shaping the research rather than contributing in parallel. These considerations are consistent with the team-science literature, which emphasises shared goals, communication, leadership, authorship expectations and conflict management as conditions for effective collaboration (Bennett and Gadlin, 2012; Hall et al., 2018).

Programme design may also need to recognise that interdisciplinary integration requires time and resources. Short application periods and funding arrangements that expect fully developed collaborations at the point of submission may advantage teams with established interdisciplinary histories. Funders could instead provide preparatory or seed funding through which researchers develop shared questions, test conceptual compatibility and negotiate methodological differences before undertaking a larger project. Within funded awards, permissible costs could include facilitated problem-formulation activities, collaboration support, dedicated project management or other forms of coordination infrastructure. Where appropriate, review or advisory arrangements during the early stages of an award could focus on whether integration is taking place, rather than waiting until final outputs are available. The present study does not evaluate these interventions directly, but its findings indicate why the practical organisation of integration deserves greater policy attention.

The strong association between prior and subsequent interdisciplinarity creates a particular policy dilemma. Prior interdisciplinary publication records may appear to provide funders with a convenient indicator of a team's capacity to produce interdisciplinary outputs. However, using that record mechanically as a selection criterion could reinforce cumulative advantage. Established interdisciplinary investigators would be more likely to receive further opportunities, while early-career researchers, researchers moving into a new area and teams developing new combinations could be disadvantaged precisely because they lack the prior outputs used to demonstrate capability. The result could be the reproduction of existing interdisciplinary networks rather than the development of new ones.

Prior interdisciplinary experience may therefore be more appropriately treated as an indicator of potential support needs than as a gatekeeping criterion. Teams without an established interdisciplinary record might benefit from longer development periods, access to facilitation, mentoring, brokerage or collaboration-building support. Funding portfolios could combine investment in established teams, where a programme requires a high probability of near-term interdisciplinary output, with explicit capability-building opportunities for less established configurations. This would recognise the predictive value of prior experience without allowing previous publication trajectories to determine access to future interdisciplinary funding.

A further implication concerns how programme success is evaluated. Publication-level interdisciplinarity captures one relevant outcome, but it does not observe the internal quality of collaboration and excludes many outputs through which research may create value. Reports, policy contributions, software, datasets, creative works, professional practices and stakeholder relationships may embody integration that is not well represented in bibliometric classifications. Moreover, relationships between interdisciplinarity and citation impact, novelty or recognition vary across fields, measures and time periods (Larivière and Gingras, 2010; Yegros-Yegros, Rafols and D'Este, 2015; Zhang et al., 2024). Evaluations based narrowly on publication classifications may therefore reward some forms of interdisciplinarity while overlooking others.

Funders should consequently evaluate diversity and interdisciplinarity through separate but complementary dimensions. Diversity monitoring may consider representation, participation, progression and inclusion, using ethically governed self-identified or administrative data where feasible. Interdisciplinary programme evaluation may consider the integration of research questions, concepts, methods and outputs, alongside the collaboration processes through which this was achieved. Bibliometric indicators can contribute to such evaluation at the portfolio level, but they should not be treated as direct measures of inclusive participation or as comprehensive measures of knowledge integration. In particular, name-based origin inference is suitable here for aggregate research analysis; it should not be used to classify individuals or inform award decisions.

The analysis also identifies a data-infrastructure constraint. In the data available for this study, awards could not be linked reliably to the calls or funding opportunities through which they were commissioned. It was therefore not possible to examine whether the wording and design of calls, including their emphasis on interdisciplinarity, equality or collaboration, influenced team formation or subsequent outputs. Stable call identifiers linked to award records would allow funders to evaluate how programme design shapes funded portfolios and outcomes. Preserving information on team composition at application and award stages, subsequent changes in project personnel, forms of integration support and the complete range of reported outputs would further strengthen policy evaluation. Such data would also make it possible to distinguish ex-ante team characteristics from the realised teams reconstructed through publication records.

Taken together, these implications shift the policy question from whether diverse teams automatically produce interdisciplinary research to how funding systems can support both inclusive participation and effective knowledge integration. The results do not warrant reducing support for diversity, nor do they justify selecting only investigators with established interdisciplinary records. They instead suggest a differentiated policy approach: diversity should be pursued as an objective with independent equity and representation rationales, while interdisciplinary funding should invest explicitly in the capabilities and conditions required to integrate knowledge across boundaries.

5.3 Implications for diversity and interdisciplinarity research

The findings qualify a simple diversity-to-interdisciplinarity mechanism without rejecting the broader proposition that diversity may matter for research. Origin diversity may be associated imperfectly with differences in educational experience, institutional socialisation, intellectual traditions or access to empirical contexts. However, it is not equivalent to cognitive, epistemic or disciplinary diversity. Consequently, a weak association with output interdisciplinarity may reflect the distance between the measured construct and the particular knowledge differences required for disciplinary integration. Future studies should distinguish more clearly among origin, disciplinary, educational, institutional, gender and academic-age diversity, rather than treating them as interchangeable manifestations of heterogeneity.

The findings also refine the role of the categorisation-elaboration model in this setting. The model offers a plausible account of why diversity may not have a uniform direct association with research outputs, but the present design tests only its reduced-form implication. The absence of a positive association cannot distinguish between several possibilities: origin diversity may not have introduced relevant cognitive variety; potentially useful differences may not have been expressed; teams may have integrated them through mechanisms not visible in publication classifications; or productive and disruptive processes may have offset one another. Direct evidence on team processes is required before any of these explanations can be preferred.

The results provide no evidence of the anticipated too-much-of-a-good-thing relationship across the observed range of origin diversity. The estimated smooths are approximately linear, and the positive association expected at low to moderate levels is absent. This does not establish that coordination or categorisation costs never increase at very high levels of diversity. The observed range is comparatively compressed and sparsely populated at its upper end. The results therefore reject H1b for the support represented in this sample, rather than the underlying theoretical possibility in all research teams.

More broadly, publication-level interdisciplinarity should not be assumed to be the principal mechanism through which diversity might influence research quality or social value. Diversity may affect problem selection, access to different empirical settings, sensitivity to stakeholder perspectives, methodological reflexivity, international reach or the identification of neglected questions. The present paper does not examine these pathways or research quality directly. Its contribution is narrower: within this project population, inferred origin diversity has little incremental association with bibliometric output interdisciplinarity once prior interdisciplinary orientation is taken into account.

5.4 Limitations and future research

Several limitations qualify this interpretation. First, the observational design does not identify a causal effect of origin diversity. M2 reduces one important source of confounding by controlling for pre-project interdisciplinarity, but endogenous team formation, collaboration history, disciplinary identity, institutional position, career stage and leadership experience remain incompletely observed. In addition, the project-team measure includes unnamed researchers identified from the same matched publications used to construct the outcome. Team composition and output interdisciplinarity therefore share publication provenance. A future design based on the named team recorded at application or award stage would provide a cleaner temporal separation between team composition and subsequent outputs.

Second, origin is inferred probabilistically from names and should not be interpreted as self-identified ethnicity, nationality, citizenship or cultural identity. Classification performance may vary across naming traditions, and the analysis holds the selected origin-diversity specification constant. Although the outcome measure is examined under alternative disciplinary levels and similarity structures, sensitivity to other plausible origin-diversity specifications remains untested. The results are therefore conditional on the particular operationalisation used here. These limitations also reinforce why inference-based measures should remain aggregate research tools rather than instruments for evaluating individuals.

Third, the dependent variable captures the average interdisciplinarity of matched publication outputs. Projects without such outputs are excluded, and publications are not equally representative of research activity across disciplines. Only 8.8% of the initially eligible projects enter the models, and the direction of the resulting selection effect cannot be established from the observed data. Project-level means are also calculated from varying numbers of matched publications and may be less stable for projects with few outputs. Projects beginning later in the observation period had less time to accumulate reported and indexed publications; controlling for start year addresses this only partially. Inference is therefore limited to the publication-reporting, multi-member Research Grant population represented in the analytical sample.

Fourth, the upper range of the origin-diversity measure is sparsely represented. The analysis can conclude that no non-linear relationship is detectable over the observed support, but it cannot exclude a different pattern among teams with substantially higher levels of diversity. Nor can it determine whether the design of individual funding calls affected team composition or interdisciplinarity, because call texts could not be linked reliably to awards.

Future research could address these limitations by linking application-stage team data, funding-call characteristics and longer-term outputs. Surveys, interviews or project observations could measure information elaboration, psychological safety, voice, conflict, leadership, authorship practices and knowledge-sharing directly. Longitudinal research could examine whether teams develop interdisciplinary capability through repeated collaboration and whether funding interventions accelerate that development. Designs comparing different forms and combinations of diversity would help establish which kinds of heterogeneity are most relevant to interdisciplinary integration. Where suitable institutional or programme variation exists, quasi-experimental approaches may also provide stronger evidence about the effects of funding design and integration support.

5.5 Conclusion

This study finds little evidence that inferred origin diversity is positively or non-linearly associated with the bibliometric interdisciplinarity of project outputs once the prior interdisciplinary orientation of principal and co-investigators is considered. Prior and subsequent interdisciplinarity are strongly associated, indicating substantial continuity in interdisciplinary orientation, although this relationship should not be interpreted causally. The findings therefore qualify the assumption that assembling an origin-diverse team will, by itself, produce more interdisciplinary publications.

The policy implication is not that diversity is unimportant. Diversity and interdisciplinarity are distinct objectives: the former has independent equity, representation and legitimacy rationales, while the latter concerns the integration of knowledge across disciplinary boundaries. Funding systems should avoid making diversity policy contingent on bibliometric performance and should not use prior interdisciplinary records mechanically in ways that reproduce cumulative advantage. Where interdisciplinary outputs are sought, greater attention should be directed towards the time, leadership, processes and institutional support through which heterogeneous knowledge can be integrated. This distinction provides a more defensible basis for pursuing both inclusive research participation and interdisciplinary knowledge production.

ETHICS STATEMENT

The analysis uses publicly available bibliometric metadata and administrative records of funded projects. Name-based inference of country of origin is used solely to construct aggregate project-team diversity indicators; no individual-level inferred attribute is reported, released or interpreted as a claim about any person’s origin, ethnicity, nationality or identity.

DATA AVAILABILITY

GtR+ data were accessed under the terms of the IRC GtR+ Secondary Data Analysis scheme and are available under those terms. OpenAlex data are openly available; this study used the full database snapshot released on 29 January 2025. The aggregate project-level dataset underlying the tables will be deposited on publication.

CODE AVAILABILITY

All processing and analysis code was written in R and will be deposited alongside the data.

REFERENCES

Adachi, T., Narimatsu, N., Ogawa, Y., Toriya, M., Fukushi, T., Shirabe, M., Futamura, M., Inomata, T., Kainuma, K., Kan-o, K., Kurashima, Y., Masaki, K., Nakajima, S., Sakashita, M., Sato, S., Tamari, M., Morita, H., & Koizumi, A. (2025). Research team diversity impacts scientific output in allergy and immunology programs. World Allergy Organization Journal. https://doi.org/10.1016/j.waojou.2024.101004

Aggarwal, V. A., Hsu, D. H., & Wu, A. (2020). Organizing knowledge production teams within firms for innovation. Strategy Science. https://doi.org/10.1287/stsc.2019.0095

Barjak, F., & Robinson, S. (2008). International collaboration, mobility and team diversity in the life sciences: Impact on research performance. Social Geography. https://doi.org/10.5194/sg-3-23-2008

Bennett, L. M., & Gadlin, H. (2012). Collaboration and team science: From theory to practice. Journal of Investigative Medicine, 60(5), 768-775. https://doi.org/10.2310/JIM.0b013e318250871d

Bodla, A. A., Tang, N., Jiang, W., & Tian, L. (2018). Diversity and creativity in cross-national teams: The role of team knowledge sharing and inclusive climate. Journal of Management and Organization. https://doi.org/10.1017/jmo.2016.34

Brunetta, F., Marchegiani, L., & Peruffo, E. (2020). When birds of a feather don't flock together: Diversity and innovation outcomes in international R&D collaborations. Journal of Business Research. https://doi.org/10.1016/j.jbusres.2019.08.033

Byron, K., Keem, S., Darden, T., Shalley, C. E., & Zhou, J. (2023). Building blocks of idea generation and implementation in teams: A meta-analysis of team design and team creativity and innovation. Personnel Psychology. https://doi.org/10.1111/peps.12501

Calatrava Moreno, M. C., Auzinger, T., & Werthner, H. (2016). On the uncertainty of interdisciplinarity measurements due to incomplete bibliographic data. Scientometrics, 107, 213-232. https://doi.org/10.1007/s11192-016-1842-4

Carusi, C., & Bianchi, G. (2020). A look at interdisciplinarity using bipartite scholar/journal networks. Scientometrics. https://doi.org/10.1007/s11192-019-03309-3

Cassi, L., Champeimont, R., Mescheba, W., & de Turckheim, É. (2017). Analysing institutions interdisciplinarity by extensive use of Rao-Stirling diversity index. PLOS ONE, 12(1), e0170296. https://doi.org/10.1371/journal.pone.0170296

Chen, S., Song, Y., Shu, F., & Larivière, V. (2022). Interdisciplinarity and impact: the effects of the citation time window. Scientometrics. https://doi.org/10.1007/s11192-022-04338-1

Chiang, H.-L., & Su, H.-N. (2026). The impact of team composition on international and inter-industry innovation: An analysis of spatial and cultural determinants. Journal of Engineering and Technology Management - JET-M. https://doi.org/10.1016/j.jengtecman.2026.101943

Eykens, J., Guns, R., & Vanderstraeten, R. (2022). Subject specialties as interdisciplinary trading grounds: the case of the social sciences and humanities. Scientometrics. https://doi.org/10.1007/s11192-021-04254-w

Figueiredo Belchior, R., & Chagas, B. T. (2026). The interplay between leadership, diversity, creativity, and innovation: A mediation model of team processes. European Journal of Innovation Management, 29(4), 1231-1266. https://doi.org/10.1108/EJIM-04-2024-0431

Gök, A., & Macmillan, G. (in press). Origin, gender and academic age diversity indicators for UK research communities. Scientific Data. https://doi.org/10.1038/s41597-026-07958-x

Gök, A., Macmillan, G., Chen, B., & Karaulova, M. (2026). Measuring cultural and ethnic diversity in research and innovation. Research Evaluation, 35(1), rvag015. https://doi.org/10.1093/reseval/rvag015

Haas, H., & Nüesch, S. (2012). Are multinational teams more successful?. International Journal of Human Resource Management. https://doi.org/10.1080/09585192.2011.610948

Hafner, L., Peifer, T. P., & Hafner, F. S. (2024). Equal accuracy for Andrew and Abubakar—detecting and mitigating bias in name-ethnicity classification algorithms. AI & Society, 39(4), 1605-1629. https://doi.org/10.1007/s00146-022-01619-4

Hall, K. L., Vogel, A. L., Huang, G. C., Serrano, K. J., Rice, E. L., Tsakraklides, S. P., & Fiore, S. M. (2018). The science of team science: A review of the empirical evidence and research gaps on collaboration in science. American Psychologist, 73(4), 532-548. https://doi.org/10.1037/amp0000319

Hofstra, B., Kulkarni, V. V., Galvez, S. M.-N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity-innovation paradox in science. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.1915378117

Kim, J., & Song, C. (2021). The relationship between R&D team diversity and team creativity. Management Decision. https://doi.org/10.1108/MD-12-2019-1727

Kozlowski, D., Murray, D. S., Bell, A., Hulsey, W., Larivière, V., Monroe-White, T., & Sugimoto, C. R. (2022). Avoiding bias when inferring race using name-based approaches. PLOS ONE, 17(3), e0264270. https://doi.org/10.1371/journal.pone.0264270

Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. Journal of the American Society for Information Science and Technology, 61(1), 126-131. https://doi.org/10.1002/asi.21226

Leinster, T., & Cobbold, C. A. (2012). Measuring diversity: The importance of species similarity. Ecology, 93(3), 477-489. https://doi.org/10.1890/10-2402.1

Lewis, H., Biesecker, B., Lee, S. S.-J., Anderson, K., Joseph, G., Jenkins, C. L., Bulkley, J. E., Leo, M. C., Goddard, K. A. B., & Wilfond, B. S. (2023). Promoting equity, inclusion, and efficiency: A team science approach to the development of authorship guidelines for a multi-disciplinary research team. Journal of Clinical and Translational Science. https://doi.org/10.1017/cts.2023.685

Leydesdorff, L. (2018). Diversity and interdisciplinarity: how can one distinguish and recombine disparity, variety, and balance?. Scientometrics. https://doi.org/10.1007/s11192-018-2810-y

Leydesdorff, L., Rafols, I., & Chen, C. (2013). Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal-journal citations. Journal of the American Society for Information Science and Technology. https://doi.org/10.1002/asi.22946

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. Journal of Informetrics. https://doi.org/10.1016/j.joi.2018.12.006

Liu, J., Gong, X., Xu, S., & Huang, C. (2024). Understanding the relationship between team diversity and the innovative performance in research teams using decision tree algorithms: evidence from artificial intelligence. Scientometrics. https://doi.org/10.1007/s11192-024-05183-0

Lockhart, J. W., King, M. M., & Munsch, C. (2023). Name-based demographic inference and the unequal distribution of misrecognition. Nature Human Behaviour, 7(7), 1084-1095. https://doi.org/10.1038/s41562-023-01587-9

Mohammadi, A., Broström, A., & Franzoni, C. (2017). Workforce composition and innovation: How diversity in employees' ethnic and educational backgrounds facilitates firm-level innovativeness. Journal of Product Innovation Management. https://doi.org/10.1111/jpim.12388

Monteiro, M., & Keating, E. (2009). Managing misunderstandings: The role of language in interdisciplinary scientific collaboration. Science Communication. https://doi.org/10.1177/1075547008330922

Nakhoda, M., Whigham, P., & Zwanenburg, S. (2023). Quantifying and addressing uncertainty in the measurement of interdisciplinarity. Scientometrics. https://doi.org/10.1007/s11192-023-04822-2

Nakhoda, M., Whigham, P., & Zwanenburg, S. (2025). How relevant is collaboration in interdisciplinary research? Bibliometric evidence from an institution's publications. Scientometrics. https://doi.org/10.1007/s11192-025-05492-y

Oliver, S. K., Fergus, C. E., Skaff, N. K., Wagner, T., Tan, P.-N., Cheruvelil, K. S., & Soranno, P. A. (2018). Strategies for effective collaborative manuscript development in interdisciplinary science teams. Ecosphere. https://doi.org/10.1002/ecs2.2206

Petersen, A. M., Arroyave, F., & Pavlidis, I. (2023). Methods for measuring social and conceptual dimensions of convergence science. Research Evaluation. https://doi.org/10.1093/reseval/rvad020

Pierce, J. R., & Aguinis, H. (2013). The too-much-of-a-good-thing effect in management. Journal of Management, 39(2), 313-338. https://doi.org/10.1177/0149206311410060

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics. https://doi.org/10.1007/s11192-008-2197-2

Pößneck, J., Kabisch, S., Kowaltowski, D. C. C. K., van Oel, C. J., Soliman-Junior, J., Tzortzopoulos, P., & Koolwijk, J. S. J. (2024). Knowledge sharing through scenario development: Experiences of an interdisciplinary and international research project. Interdisciplinary Science Reviews. https://doi.org/10.1177/03080188241291828

Sebo, P. (2023). How well does NamSor perform in predicting the country of origin and ethnicity of individuals based on their first and last names? PLOS ONE, 18(11), e0294562. https://doi.org/10.1371/journal.pone.0294562

Settles, I. H., Brassel, S. T., Soranno, P. A., Cheruvelil, K. S., Montgomery, G. M., & Elliott, K. C. (2019). Team climate mediates the effect of diversity on environmental science team satisfaction and data sharing. PLoS ONE. https://doi.org/10.1371/journal.pone.0219196

Soós, S., & Kampis, G. (2012). Beyond the basemap of science: Mapping multiple structures in research portfolios: Evidence from Hungary. Scientometrics. https://doi.org/10.1007/s11192-012-0713-x

Specht, A., & Crowston, K. (2022). Interdisciplinary collaboration from diverse science teams can produce significant outcomes. PLoS ONE. https://doi.org/10.1371/journal.pone.0278043

Stahl, G. K., & Maznevski, M. L. (2021). Unraveling the effects of cultural diversity in teams: A retrospective of research on multicultural work groups and an agenda for future research. Journal of International Business Studies, 52(1), 4-22. https://doi.org/10.1057/s41267-020-00389-9

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. Journal of the Royal Society Interface, 4(15), 707-719. https://doi.org/10.1098/rsif.2007.0213

Sun, H., Teh, P.-L., Ho, K., & Lin, B. (2017). Team diversity, learning, and innovation: A mediation model. Journal of Computer Information Systems. https://doi.org/10.1080/08874417.2016.1181490

Traylor, A. M., Dinh, J. V., Ng, L. C., Reyes, D. L., Cheng, S. K., Croitoru, N. C., & Salas, E. (2024). It's about the process, not the product: A meta-analytic investigation of team demographic diversity and processes. Organizational Psychology Review. https://doi.org/10.1177/20413866241245312

van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. Journal of Applied Psychology, 89(6), 1008-1022. https://doi.org/10.1037/0021-9010.89.6.1008

van Knippenberg, D., Li, J., & Tu, Y. (2024). Team informational resources, information elaboration, and team innovation: Diversity mindset moderating functional diversity and boundary spanning scouting effects. Journal of Occupational and Organizational Psychology. https://doi.org/10.1111/joop.12541

Wallrich, L., Opara, V., Wesołowska, M., Barnoth, D., & Yousefi, S. (2024). The relationship between team diversity and team performance: Reconciling promise and reality through a comprehensive meta-analysis registered report. Journal of Business and Psychology, 39(6), 1303-1354. https://doi.org/10.1007/s10869-024-09977-0

Wang, J., Cheng, G. H.-L., Chen, T., & Leung, K. (2019). Team creativity/innovation in culturally diverse teams: A meta-analysis. Journal of Organizational Behavior, 40(6), 693-708. https://doi.org/10.1002/job.2362

Wang, Q., & Schneider, J. W. (2020). Consistency and validity of interdisciplinarity measures. Quantitative Science Studies, 1(1), 239-263. https://doi.org/10.1162/qss_a_00011

Watson, W. E., Johnson, L., & Zgourides, G. D. (2002). The influence of ethnic diversity on leadership, group process, and performance: An examination of learning teams. International Journal of Intercultural Relations. https://doi.org/10.1016/S0147-1767(01)00032-3

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B, 73(1), 3-36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wu, K., Leng, R. I., & Hou, W. (2026). Weak ties matter for scientific innovation: The dual perspectives of knowledge recombination network and collaboration network. Journal of Informetrics. https://doi.org/10.1016/j.joi.2026.101779

Yang, A. J. (2025). Understanding the relationship between interdisciplinary knowledge and disruption in science. Journal of Information Science. https://doi.org/10.1177/01655515251330614

Yegros-Yegros, A., Rafols, I., & D'Este, P. (2015). Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. PLOS ONE, 10(8), e0135095. https://doi.org/10.1371/journal.pone.0135095

Zhang, L., & Leydesdorff, L. (2021). The scientometric measurement of interdisciplinarity and diversity in the research portfolios of Chinese universities. Journal of Data and Information Science. https://doi.org/10.2478/jdis-2021-0027

Zhang, Y., Wang, Y., Du, H., & Havlin, S. (2024). Delayed citation impact of interdisciplinary research. Journal of Informetrics, 18(1), 101468. https://doi.org/10.1016/j.joi.2023.101468

APPENDIX

Appendix A1. Measurement specifications and scenario correspondence

Table A1 documents the correspondence between the measures used in this paper and the scenario codes under which they are released in the GtR+-Diversity dataset and defined in Gök and Macmillan (in press). Scenario codes are retained in the released data for reproducibility; the main text refers to the measures substantively.

Table A1. Measurement specifications and scenario correspondence

Measure	Role	Construction	Scenario code
Output interdisciplinarity (main)	Dependent variable	Publication-level Rao-Stirling; OpenAlex subfield level; disciplinary profile defined from references; cosine distance; time-invariant distance matrix; aggregated to the project level using the mean	Interdisciplinarity Scenario 7 (GtR+-Diversity dataset)
Output interdisciplinarity (robustness)	Robustness dependent variable	As above, at the broader OpenAlex field level	Interdisciplinarity Scenario 11 (GtR+-Diversity dataset)
Output interdisciplinarity (robustness)	Robustness dependent variable	Alternative index based on a wider similarity matrix, under both the subfield-level and field-level specifications	See Appendix A4
Prior interdisciplinarity	Control	Mean publication-level Rao-Stirling of PI/CoI publications in the five years before project start; same specification as the corresponding outcome	Matches the outcome scenario
Cultural diversity	Independent variable	Leinster-Cobbold index with $q = 1$; NamSor country-of-origin inference with confidence filtering; population-weighted geodesic distance between countries; exponential transformation with decay parameter $t = 0.05$; global rarity reweighting of relative abundance	Origin Scenario 9b (Gök and Macmillan, in press); Scenario 9a is the $t = 0.1$ variant

Appendix A2. Full GAM results (subfield-level outcome)Appendix A2. Full GAM results (subfield-level outcome)

The main analysis measures output interdisciplinarity at the subfield level using cosine distance with the temporal option switched off. Two robustness checks were conducted. First, the models were re-estimated using a coarser field-level interdisciplinarity measure (Appendix A3). Second, the models were re-estimated using an alternative interdisciplinarity index based on a wider similarity matrix (Appendix A4). In both robustness checks, publication-level output and prior interdisciplinarity scores were aggregated to the grant level using the mean before re-estimating the GAM models, and the same modelling pipeline was applied throughout, covering descriptives, correlations, GAM models, reporting tables and diagnostics for both the subfield-level and field-level outcomes.

Table A2. Full GAM results for the subfield-level outcome, including supplementary robustness models

Term	M0 Baseline	M1 + Diversity	M2 + Prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration	M7 Start-Year FE
Parametric terms								
Start year: 2022	–	–	–	–	–	–	–	−0.001 (0.002)
Start year: 2023	–	–	–	–	–	–	–	−0.001 (0.002)
Start year: 2024	-	-	-	-	-	-	-	−0.001 (0.003)
Intercept	0.243*** (0.002)	0.243*** (0.002)	0.235*** (0.001)	0.234*** (0.001)	0.234*** (0.001)	0.237*** (0.002)	0.248*** (0.004)	0.236*** (0.001)
AHRC	0.018*** (0.005)	0.016*** (0.005)	0.009** (0.004)	0.009* (0.004)	0.008* (0.004)	0.008* (0.004)	0.008* (0.004)	0.009** (0.004)
BBSRC	-0.029*** (0.002)	-0.030*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)	−0.011*** (0.002)
ESRC	0.004 (0.004)	0.002 (0.004)	0.003 (0.003)	0.004 (0.003)	0.003 (0.003)	0.003 (0.003)	0.003 (0.003)	0.003 (0.003)
MRC	-0.023*** (0.002)	-0.024*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	−0.014*** (0.002)
NERC	-0.023*** (0.003)	-0.026*** (0.003)	-0.010*** (0.002)	-0.010*** (0.002)	-0.010*** (0.002)	-0.010*** (0.002)	-0.010*** (0.002)	−0.010*** (0.002)
Includes non-UK partner	-0.002 (0.002)	-0.002 (0.002)	-0.001 (0.001)	-0.002 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)	−0.001 (0.001)

Term	M0 Baseline	M1 + Diversity	M2 + Prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration	M7 Start-Year FE
Includes non-university partner	—	—	—	0.002. (0.001)	—	—	—	—
First PI grant	—	—	—	—	0.002 (0.001)	—	—	
Spend per month: £10k-£20k	—	—	—	—	—	-6E-06	—	
Spend per month: >£20k	—	—	—	—	—	-0.003. (0.002)	—	
Duration: 7-12 months	—	—	—	—	—	—	-0.00005	
Duration: 13-24 months	—	—	—	—	—	—	-3.6E-05	
Duration: 25-36 months	—	—	—	—	—	—	-0.012** (0.004)	
Duration: 37-48 months	—	—	—	—	—	—	-0.014*** (0.004)	
Duration: 49+ months	—	—	—	—	—	—	-0.015*** (0.004)	
Smooth terms								
s(log team size)	edf = 1.00; p = 0.360	edf = 1.00; p = 0.008**	edf = 2.21; p = 0.120	edf = 2.24; p = 0.147	edf = 2.35; p = 0.086.	edf = 2.40; p = 0.070.	edf = 2.45; p = 0.052.	
s(cultural diversity)	—	edf = 1.00; p < 0.001***	edf = 1.00; p = 0.099.	—	—	—	—	
s(prior interdisciplinarity)	—	—	edf = 3.30; p < 0.001***	edf = 3.23; p < 0.001***	edf = 3.21; p < 0.001***	edf = 3.36; p < 0.001***	edf = 3.30; p < 0.001***	
s(cultural diversity) × non-university partner = FALSE	—	—	—	edf = 1.00; p = 0.136	—	—	—	
s(cultural diversity) × non-university partner = TRUE	—	—	—	edf = 1.00; p = 0.306	—	—	—	

Term	M0 Baseline	M1 + Diversity	M2 + Prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration	M7 Start-Year FE
s(cultural diversity) × first PI grant = FALSE	—	—	—	—	edf = 2.02; p = 0.289	—	—	
s(cultural diversity) × first PI grant = TRUE	—	—	—	—	edf = 2.19; p = 0.089.	—	—	
s(cultural diversity) × spend <£10k/month	—	—	—	—	—	edf = 1.01; p = 0.088.	—	
s(cultural diversity) × spend £10k-£20k/month	—	—	—	—	—	edf = 1.00; p = 0.103	—	
s(cultural diversity) × spend >£20k/month	—	—	—	—	—	edf = 1.00; p = 0.708	—	
s(cultural diversity) × duration 0-6 months	—	—	—	—	—	—	edf = 1.96; p = 0.065.	
s(cultural diversity) × duration 7-12 months	—	—	—	—	—	—	edf = 1.90; p = 0.413	
s(cultural diversity) × duration 13-24 months	—	—	—	—	—	—	edf = 1.01; p = 0.004**	
s(cultural diversity) × duration 25-36 months	—	—	—	—	—	—	edf = 1.64; p = 0.586	
s(cultural diversity) × duration 37-48 months	—	—	—	—	—	—	edf = 1.00; p = 0.225	
s(cultural diversity) × duration 49+ months	—	—	—	—	—	—	edf = 2.32; p = 0.452	
Fit statistics								
N	2,795	2,795	2,795	2,795	2,795	2,795	2,795	
AIC	-9,352	-9,366.4	-10,959.7	-10,958.6	-10,960	-10,959.9	-10,974.1	

Term	M0 Baseline	M1 + Diversity	M2 + Prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration	M7 Start-Year FE
Deviance explained (%)	9.03	9.57	49.08	49.13	49.28	49.24	49.92	
Δ deviance explained (%)	—	0.53	39.51	0.05	0.2	0.16	0.84	

Note: Standard errors in parentheses. Smooth terms report effective degrees of freedom and approximate p-values. Significance codes: . p < .10; *p < .05; **p < .01; ***p < .001. EPSRC is the reference category for delivery body; base categories for the supplementary terms are FALSE (non-university partner, first PI grant), <£10k/month (spend) and 0-6 months (duration). M7 adds project start-year fixed effects to M2. The reference category for project start year is 2021. Start-year fixed effects are not included in M0–M6.

Appendix A3. Field-level robustness check

The first robustness check re-estimates the model sequence using the coarser field-level interdisciplinarity measure, with publication-level output and prior scores aggregated to the grant level using the mean. The results are broadly consistent with the main subfield-level analysis. Cultural diversity is significant before prior interdisciplinarity is added but becomes marginal after prior interdisciplinarity enters the model. Prior interdisciplinarity remains highly significant and accounts for the main improvement in model fit.

The field-level results reproduce the main pattern. Cultural diversity is significantly associated with output interdisciplinarity in M1 (edf = 1.00, p = .009), but the association weakens to marginal significance after prior interdisciplinarity is included in M2 (edf = 1.00, p = .062). Prior interdisciplinarity is strongly associated with the field-level outcome (edf = 1.03, p < .001). Deviance explained rises from 2.84% in M1 to 43.95% in M2, indicating that prior interdisciplinarity accounts for a substantial proportion of variation in subsequent output interdisciplinarity.

Table A3. Full GAM results using the field-level outcome

section	term	M0 Baseline	M1 Baseline + diversity	M2 Baseline + diversity + prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration
Parametric terms	Intercept	0.130*** (0.001)	0.131*** (0.001)	0.128*** (0.001)	0.127*** (0.001)	0.128*** (0.001)	0.129*** (0.001)	0.136*** (0.003)
Parametric terms	Delivery body: AHRC (base: EPSRC)	0.012** (0.004)	0.011** (0.004)	0.008** (0.003)	0.007* (0.003)	0.007* (0.003)	0.008** (0.003)	0.007* (0.003)
Parametric terms	Delivery body: BBSRC (base: EPSRC)	-0.004* (0.002)	-0.004* (0.002)	-0.004** (0.001)	-0.004** (0.001)	-0.004** (0.001)	-0.004** (0.001)	-0.004** (0.001)
Parametric terms	Delivery body: ESRC (base: EPSRC)	0.006* (0.003)	0.005. (0.003)	0.005* (0.002)	0.005* (0.002)	0.005* (0.002)	0.005* (0.002)	0.005* (0.002)
Parametric terms	Delivery body: MRC (base: EPSRC)	-0.012*** (0.002)	-0.012*** (0.002)	-0.005*** (0.001)	-0.005*** (0.002)	-0.005*** (0.001)	-0.005*** (0.002)	-0.005*** (0.002)
Parametric terms	Delivery body: NERC (base: EPSRC)	-0.006** (0.002)	-0.008*** (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)
Parametric terms	Includes non-UK partner: TRUE (base: FALSE)	-0.002. (0.001)	-0.002 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Parametric terms	Includes non-university partner: TRUE (base: FALSE)				0.002. (0.001)			

section	term	M0 Baseline	M1 Baseline + diversity	M2 Baseline + diversity + prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration
Parametric terms	First PI grant: TRUE (base: FALSE)					0.001 (0.001)		
Parametric terms	Spend per month: 10k-20k (base: <10k)						-0.001 (0.001)	
Parametric terms	Spend per month: >20k (base: <10k)						0.000 (0.001)	
Parametric terms	Duration: 7-12 (base: 0-6)							-0.006 (0.004)
Parametric terms	Duration: 13-24 (base: 0-6)							-0.005 (0.004)
Parametric terms	Duration: 25-36 (base: 0-6)							-0.008* (0.003)
Parametric terms	Duration: 37-48 (base: 0-6)							-0.009** (0.003)
Parametric terms	Duration: 49+ (base: 0-6)							-0.007* (0.003)
Smooth terms	s(log team size)	edf = 1.31; p = 0.635	edf = 1.84; p = 0.126	edf = 1.00; p = 0.971	edf = 1.00; p = 0.937	edf = 1.00; p = 0.953	edf = 1.00; p = 0.937	edf = 1.00; p = 0.929
Smooth terms	s(cultural diversity)		edf = 1.00; p = 0.009**	edf = 1.00; p = 0.062.				
Smooth terms	s(prior interdisciplinarity)			edf = 1.03; p = <0.001***	edf = 1.02; p = <0.001***	edf = 1.16; p = <0.001***	edf = 1.03; p = <0.001***	edf = 1.03; p = <0.001***
Smooth terms	s(cultural diversity) by non-university partner = FALSE				edf = 1.00; p = 0.214			
Smooth terms	s(cultural diversity) by non-university partner = TRUE				edf = 1.00; p = 0.123			
Smooth terms	s(cultural diversity) by first PI grant = FALSE					edf = 1.01; p = 0.371		
Smooth terms	s(cultural diversity) by first PI grant = TRUE					edf = 1.00; p = 0.053.		
Smooth terms	s(cultural diversity) by spend per month = <10k						edf = 1.00; p = 0.095.	

section	term	M0 Baseline	M1 Baseline + diversity	M2 Baseline + diversity + prior	M3 Non-university	M4 First PI	M5 Spend	M6 Duration
Smooth terms	s(cultural diversity) by spend per month = 10k-20k						edf = 1.00; p = 0.220	
Smooth terms	s(cultural diversity) by spend per month = >20k						edf = 1.00; p = 0.581	
Smooth terms	s(cultural diversity) by duration = 0-6							edf = 1.87; p = 0.007**
Smooth terms	s(cultural diversity) by duration = 7-12							edf = 1.00; p = 0.472
Smooth terms	s(cultural diversity) by duration = 13-24							edf = 1.00; p = 0.083.
Smooth terms	s(cultural diversity) by duration = 25-36							edf = 1.87; p = 0.374
Smooth terms	s(cultural diversity) by duration = 37-48							edf = 3.16; p = 0.108
Smooth terms	s(cultural diversity) by duration = 49+							edf = 1.00; p = 0.837
Fit statistics	N	2795	2795	2795	2795	2795	2795	2795
Fit statistics	AIC	-10690.1	-10694.5	-12232.7	-12232.1	-12230.6	-12226.7	-12239.0
Fit statistics	Deviance explained (%)	2.56	2.84	43.95	44.02	44.00	43.99	44.70
Fit statistics	Δ Deviance explained (%)		0.28	41.11	0.07	0.05	0.04	0.75

Note: Parametric terms report coefficient estimates with standard errors in parentheses. Smooth terms report effective degrees of freedom and approximate p-values. Scenario 11 measures output interdisciplinarity at the field level. All models use the same common complete-case sample of 2,795 projects. Significance codes: *** p < .001; ** p < .01; * p < .05; . p < .10.

Note: Standard errors in parentheses. Smooth terms report effective degrees of freedom and approximate p-values. Significance codes: . p < .10; *p < .05; **p < .01; ***p < .001.

Appendix A4. Wider similarity matrix robustness check

The alternative wider similarity matrix robustness check uses two new paper-level files: one for project-output interdisciplinarity and one for prior PI/CoI interdisciplinarity. Both were aggregated to the grant level using the mean. The merge coverage was good: among 3,261 projects in the model dataset, 3,166 had alternative output interdisciplinarity scores, 3,107 had alternative prior interdisciplinarity scores, and 3,017 had both alternative output and prior scores for the subfield-level and field-level specifications. Among the 3,261 Research Grant projects with more than one team member in the modelling dataset before the scenario-specific complete-case restriction, 3,166 had alternative output-interdisciplinarity scores, 3,107 had alternative prior-interdisciplinarity scores, and 3,017 had both under the Scenario 7 and Scenario 11 specifications. The common complete-case sample used in the reported models contains 2,795 projects.

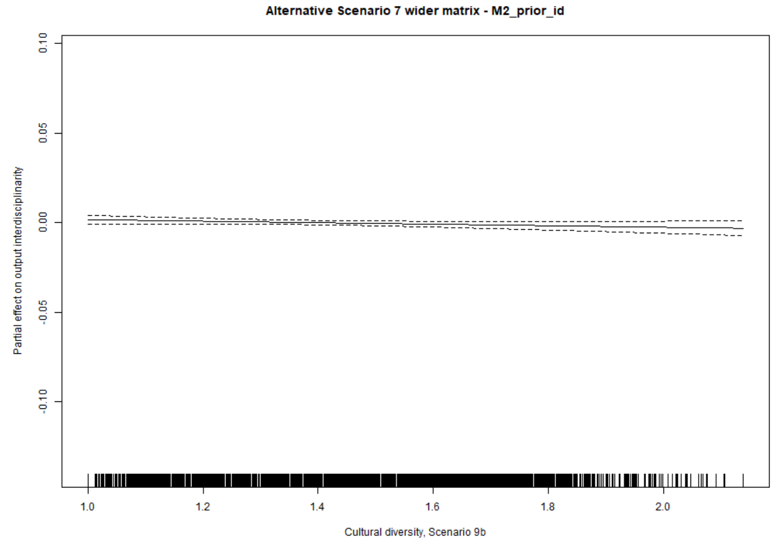
The robustness results are substantively unchanged. For the subfield-level outcome, cultural diversity remains significant in the diversity model, but becomes non-significant after prior interdisciplinarity is added. The M2 cultural-diversity p-value changes only slightly from 0.099 in the original specification to 0.109 in the wider-matrix specification. Prior interdisciplinarity remains highly significant in both specifications. The shape of the diversity smooth also does not materially change: the edf remains close to 1, and the smooth plot is nearly flat after prior interdisciplinarity is controlled.

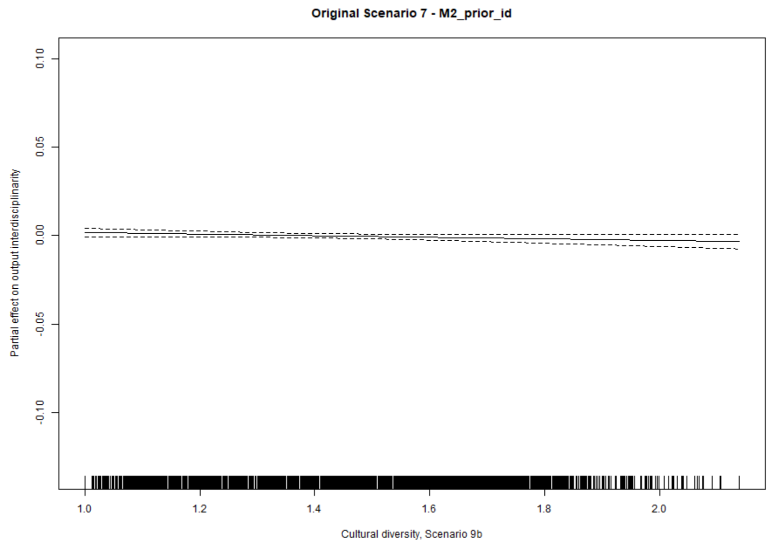
For the field-level outcome, the same pattern appears. The M2 cultural-diversity p-value changes from 0.062 in the original specification to 0.092 in the wider-matrix specification, while prior interdisciplinarity remains highly significant. Deviance explained changes only slightly for the field-level models, from 43.95% to 43.71% in M2. For the subfield-level models, deviance explained is modestly lower under the wider matrix, falling from 49.08% to 46.94% in M2. Overall, the wider similarity matrix does not alter the main interpretation: cultural diversity is weak or marginal after prior interdisciplinarity is included, the smooth remains close to linear, and prior interdisciplinarity remains the dominant predictor.

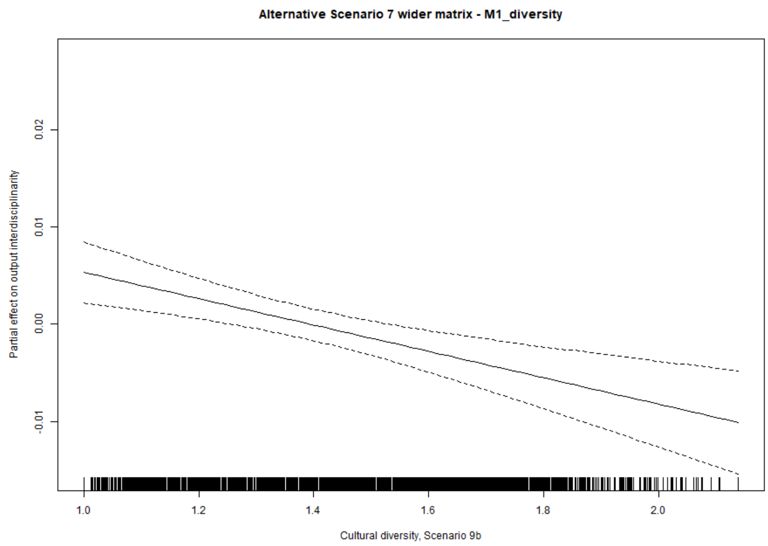
Table A4. Wider similarity matrix robustness summary

Outcome	Model	Original diversity p	Alternative diversity p	Original diversity edf	Alternative diversity edf	Original deviance explained (%)	Alternative deviance explained (%)	Prior p-value
Subfield (S7)	M1 Diversity	<0.001	<0.001	1	1	9.57	8.58	—
Subfield (S7)	M2 + Prior	0.099	0.109	1	1	49.08	46.94	<0.001
Subfield (S7)	M3 Non-university	0.136	0.204	1	1	49.13	46.95	<0.001
Subfield (S7)	M4 First PI	0.089	0.13	2.19	2.32	49.28	47.17	<0.001
Subfield (S7)	M5 Spend	0.088	0.12	1.01	1	49.24	47.08	<0.001
Subfield (S7)	M6 Duration	0.004	0.011	2.32	2.64	49.92	47.72	<0.001
Field (S11)	M1 Diversity	0.009	0.017	1	1.01	2.84	2.16	—
Field (S11)	M2 + Prior	0.062	0.092	1	1	43.95	43.71	<0.001
Field (S11)	M3 Non-university	0.123	0.108	1	1	44.02	43.77	<0.001
Field (S11)	M4 First PI	0.053	0.064	1.01	1	44	43.74	<0.001
Field (S11)	M5 Spend	0.095	0.095	1	1	43.99	43.73	<0.001
Field (S11)	M6 Duration	0.007	0.014	3.16	2.73	44.7	44.37	<0.001

Figure A1. Cultural diversity smooths under the original and wider similarity matrix specifications







www.ircaucus.ac.uk

Email info@ircaucus.ac.uk Twitter @IRCaucus

INNOVATION & RESEARCH CAUCUS

Delivered with ESRC and Innovate UK