technopolis group

June 2026

Evaluation of the Cross Research Council Responsive Mode Pilot Scheme

Final report

Prepared by Technopolis, the Innovation Research Caucus and UKRI

June 2026

Evaluation of the Cross Research Council Responsive Mode Pilot Scheme

Final report

Prepared by Technopolis, the Innovation Research Caucus and UKRI

Accompanying documents can be found on the UKRI website:

- Summary Report
- Technical Annex

Technopolis: Anoushka Davé, Kalle Nielsen, Rima Martin, Bruno Raabe, Diogo Machado, Rosa Parker, Carlos Vergara, James Bates, Tia-J'Nae Murray, Cristina Rosemberg.

Innovation & Research Caucus (IRC): Jen Nelles, Michalis Papazoglou, Lauren Tuckerman, Kevin Walsh. The Innovation and Research Caucus (IRC) is supported by Economic and Social Research Council (ESRC) grant ES/X010759/1.

Table of Contents

UKRI Foreword 6

1 Introduction 12

1.1 Overview of the CRCRM pilot scheme 12

1.2 Evaluation 28

2 Process evaluation findings 30

2.1 Overview of application data 30

2.2 Quality and effectiveness of assessment processes 34

2.3 Effectiveness of the IAC approach 53

2.4 Successes and challenges 64

2.5 Assessing IDR in applications – key learnings 74

3 CRCRM portfolio 76

3.1 Quality of applications 76

3.2 Characteristics of collaborations being supported 77

4 Impact on portfolio 101

4.1 The CRCRM pilot addresses a gap in the UK's IDR funding landscape 101

4.2 Benchmarking to other UKRI schemes 104

4.3 Cross-disciplinary research challenges being tackled 107

4.4 Enablers and barriers 113

4.5 Impact on the research community 113

5 Conclusions and Recommendations 116

5.1 Conclusions 116

5.2 Recommendations for IDR opportunities and assessment processes 121

Appendix A Theory of Change 124

Appendix B Methodology 125

B.1. Evaluation questions 125

B.2. Interviews 127

B.3. Surveys 128

B.4. Interdisciplinarity and novelty analysis 130

B.5. Signals of Interdisciplinary Focus in Text (SIFT) 131

B.6. AI Policy Concierge (AIPC) analysis of application summaries 133

Appendix C Supplementary data 135

C.1. Interdisciplinarity and novelty analysis 135

C.2. UKRI analysis 135

C.3. Clarification Statements sent to Round 2 outline stage applicants 143

Tables

Table 1 Changes between Round 1 and Round 2 Outline Stages 22

Table 2 Changes between Round 1 and Round 2 Full Stages 25

Table 3 CRCRM applications by remit area and application stage for Round 1 (a) and Round 2 (b) 31

Table 4 Proportion of applications per band 44

Table 5 Statistical (t-test) comparisons of CRCRM application indicators against benchmark projects 106

Table 6 Aggregate means of key indicators per funding programme 107

Table 7 Breakdown of interviews 128

Table 8 Survey response rates 129

Table 9 Semantic fixed-effect regressions with CRCRM award dummy (Round 2) 135

Table 10 Project Lead gender data for each assessment stage and awards for Round 1 (a) and Round 2 (b). Data is illustrated in Figure 20. 135

Table 11 Project Lead ethnicity data for each assessment stage and awards for Round 1 (a) and Round 2 (b). Data illustrated as binary in Figure 21. 136

Table 12 Project Lead disability data for each assessment stage and awards for Round 1 (a) and Round 2 (b). 137

Table 13 Project Lead age range for each assessment stage and awards for Round 1(a) and (b) 138

Figures

Figure 1 What excellent interdisciplinary research looks like 17

Figure 2 CRCRM interdisciplinary panel interfaces 18

Figure 3 Round 1 applications' scores distribution for (a) outline and (b) full stage 36

Figure 4 Round 2 applications scores distribution for (a) outline and (b) full stage 39

Figure 5 Distribution of final agreed scores by band 42

Figure 6 Stakeholder views on communication about CRCRM aims, assessment processes and assessment criteria, Round 1 and Round 2 45

Figure 7 Applicants' view on guidance provided, Round 1 and Round 2 47

Figure 8 Stakeholder views on the extent to which the multistage assessment process adds value, Round 1 and Round 2 50

Figure 9 Applicants' views on the usefulness and clarity of feedback 52

Figure 10 IAC members' motivations for applying to join the assessment college, Round 1. 54

Figure 11 Overall, the training and guidance materials provided were sufficient in preparing you for your role in the CRCRM scheme full stage panels (n = 63) 56
Figure 12 The training materials for the full stage were of high quality and informative 56
Figure 13 Differences between applications' and assessors' remits across rounds 60
Figure 14 Questions comparing applicants' experiences with Round 1 (n=163) 67
Figure 15 Network maps showing disciplinary combinations present in outline, full and awarded applications in CRCRM Round 1 79
Figure 16 Network maps showing disciplinary combinations present in outline, full and awarded applications in CRCRM Round 2 80
Figure 17 Mean SIFT score by band and outcome with 95% confidence intervals, Round 1 full stage applications 82
Figure 18 Mean SIFT score for CRCRM grants by round and stage using the application summary sections and the fit to opportunity section (Fit) for Round 2 applications compared with summary text from previously UKRI funded research available on Gateway to Research (GtR) 84
Figure 19 Mean SIFT scores for dimensions of interdisciplinarity in Round 1 Full stage and Round 2 86
Figure 20 Project Leads gender at each assessment stage and awards for Round 1 (a) and Round 2 (b). 88
Figure 21 Project Leads Ethnicity (binary) at each assessment stage and awards for Round 1 (a) and Round 2 (b) 88
Figure 22 Success rates associated with protected characteristics at outline (a) and full (b) stages for Rounds 1 and 2 90
Figure 23 Invitation and success rates by organisation for project leads (a) and all applicants (b) 95
Figure 24 Number of organisations on applications 97
Figure 25 Number of team members on applications 98
Figure 26 Percentage of individuals associated with mission groups (a) and success rates (b) for project leads and all applicants by round and stage 99
Figure 27 Bar charts showing the key problem categories addressed in outline, full and awarded applications in Round 1 (a) and Round 2 (b) 108
Figure 28 Bar charts showing the key methods used in outline, full and awarded applications in Round 1 (a) and Round 2 (b) 110
Figure 29 Bar charts showing the main impact areas targeted in outline, full and awarded applications in Round 1 (a) and Round 2 (b) 112
Figure 30 Round 2 full stage applicants' views on the aspects that add value to the CRCRM assessment process (n=158) 113
Figure 31 Stakeholder perceptions about the scheme's assessment process (n=63). 114
Figure 32 Project Lead disability for each assessment stage and awards for Round 1 (a) and (b) 137
Figure 33 Project Lead age range for each assessment stage and awards for Round 1(a) and (b) 138
Figure 34 CRCRM application invitation and success rates by region by round and stage 139
Figure 35 Regional representation of project leads and applicants 140
Figure 36 Number of regions on applications 142

Technical Annex

The technical annex document is accessible via the UKRI website.

Appendix A Survey analysis for CRCRM Round 1 outline stage ____ 1
Appendix B Survey analysis for CRCRM Round 1 full stage ____ 36
Appendix C Survey analysis for the CRCRM Round 2 outline stage ____ 83
Appendix D Survey analysis for CRCRM Round 2 full stage ____ 123
Appendix E CRCRM pilot: novelty and interdisciplinarity assessment ____ 185
Appendix F Analysis of Panel Scores ____ 202
Appendix G Score Metrics ____ 212
Appendix H AIPC analysis of applications: Methods, problems and impacts ____ 213
Appendix I AIPC analysis of applications: Topic co-occurrence ____ 247

UKRI Foreword

Professor Alison Park
SRO for UKRI Cross Research Council Responsive Mode Scheme
Deputy Executive Chair of the Economic and Social Research Council

Dr Alexandra Amey
Strategic Lead
Interdisciplinary Research UKRI

The Cross Research Council Responsive Mode (CRCRM) pilot scheme represented an ambitious and important step forward in UKRI's approach to supporting interdisciplinary research (IDR).

At a time when many of the most pressing societal and scientific challenges increasingly demand insights, methods and collaborations that span traditional disciplinary boundaries, CRCRM set out to test how UKRI could more effectively enable, identify and support genuinely interdisciplinary research.

The value and outcomes of the CRCRM pilot

The findings of this evaluation demonstrate that the scheme has been highly effective in achieving its aims. It has successfully funded high-quality, innovative and well-integrated interdisciplinary projects, while simultaneously generating a rich evidence base on how such research can be assessed more fairly and effectively. Stakeholder feedback from applicants, assessors and observers consistently highlights the value of the scheme—not only in enabling new research, but in legitimising and strengthening interdisciplinary approaches across the research community.

Across the two funding rounds, the CRCRM pilot has supported 68 ambitious projects exemplifying the power of integrating perspectives to generate novel solutions and reframe complex problems. The portfolio includes work combining digital art and clinical practice to improve children's healthcare experiences, bio inspired robotics for targeted cancer drug delivery, and new economic models to provide early warning of climate tipping points. It spans research drawing on historical and scientific approaches to inform modern medicine, alongside projects addressing complex systems challenges such as lung disease, global resource use, and children's agency in an AI driven

world. The funded projects demonstrate both the intellectual diversity and real-world relevance of interdisciplinary research supported by the scheme.

Defining interdisciplinary research and why it matters

A central achievement of the CRCRM pilot has been the development of a clearer, shared understanding of what constitutes interdisciplinary research. While "multidisciplinary" research typically involves contributions from multiple disciplines working in parallel on a shared problem, "interdisciplinary" research goes further—requiring genuine integration of concepts, theories, data, methods and perspectives to create new frameworks, insights or approaches that would not be achievable within a single discipline alone. This integration is often characterised by co-creation, reciprocity between disciplines, and the development of shared language and understanding across traditionally distinct research cultures.

UKRI has a considerable track record in funding multi and interdisciplinary research, but interdisciplinary research introduces additional complexity and uncertainty in the assessment process, particularly in how research quality and integration are assessed. Interdisciplinary proposals may challenge established disciplinary norms, draw on diverse forms of evidence, and adopt novel or emerging methodologies. As a result, they can be more difficult to evaluate using standard peer review processes, which are often rooted in disciplinary expertise and expectations.

The CRCRM pilot has addressed these challenges by developing a robust definition of excellent IDR, supported by practical guidance and training materials. Importantly, the scheme has demonstrated that assessing interdisciplinary research requires not only appropriate criteria, but also cultural and procedural change in how assessment is conducted.

Innovations in assessment and delivery

Equally significant has been the innovation in assessment processes. The creation of the Interdisciplinary Assessment College (IAC), bringing together a trained and diverse pool of experts, has enabled more holistic, balanced and confident assessments of interdisciplinary proposals. A key factor underpinning this success has been the comprehensive training provided to IAC members. Through multiple structured training sessions, supported by practical guidance and opportunities for discussion, members developed a shared understanding of interdisciplinary research and how to assess it effectively. This was further reinforced through peer-to-peer learning within panels, where assessors were able to exchange perspectives, challenge their own assumptions, learn from different disciplinary perspectives, and build collective confidence in their judgements. Together, this combination of

formal training, practical experience and collaborative learning has not only strengthened the consistency and robustness of assessment, but also contributed to a deeper, more reflective understanding of interdisciplinary research within the research community. Panel-based deliberation, structured feedback mechanisms, clarification statements for applicants, and the introduction of Readers have all contributed to more robust and transparent decision-making. The use of clearer scoring definitions and banding approaches has also improved consistency and comparability across panels.

The scheme has also piloted a number of practices with wider relevance for UKRI funding processes. These include the use of partial randomisation to support fair allocation of funding where demand exceeds available resources, clearer guidance for the scoring definitions, enhanced panel practices to ensure robust assessments, and approaches to demand management that reduce burden on the research community while maintaining quality. Collectively, these innovations represent important advances not only for interdisciplinary research, but for funding practices more broadly and are already being included in UKRI policies.

Building the evidence base

The evaluation has also developed new analytical approaches to understanding interdisciplinarity within UKRI's portfolio. In particular, the Signals of Interdisciplinary Focus in Text (SIFT) tool offers a promising method for systematically identifying and analysing interdisciplinary characteristics within applications and awards. This has the potential to strengthen UKRI's ability to understand, monitor and support IDR across its wider funding landscape.

Looking ahead

While the CRCRM pilot scheme will not continue in its current form, UKRI remains committed to supporting interdisciplinary research as a vital component of the UK's research and innovation system. The evidence presented in this evaluation reinforces the importance of increasing opportunities for both multi- and interdisciplinary research, minimising research bureaucracy, and addressing administrative barriers—particularly those that can arise between research councils. To ensure that the insights and innovations from the CRCRM pilot are fully realised, UKRI has established a Task and Finish group to review the evaluation findings in detail and make recommendations for how these can be embedded across UKRI's wider portfolio of funding opportunities.

Acknowledgements

We would like to express our sincere thanks to all those who have contributed to the success of the CRCRM pilot and to this evaluation. This includes the

various governance groups, the IRM team and council delivery teams who designed and delivered the scheme; the members of the Interdisciplinary Assessment College, whose expertise and commitment underpinned the quality of assessment; Catherine Lyall and Samantha Aspinall for their leadership in developing IAC training; and all applicants to the scheme, particularly those who provided valuable feedback through surveys and interviews. We are also grateful to the evaluation partners—Technopolis, the Innovation Research Caucus—and to the UKRI team responsible for the analytical work that has contributed to the insights presented in this report.

The CRCRM pilot has demonstrated both the demand for and the value of dedicated support for interdisciplinary research. The challenge now is to ensure that its lessons are realised across UKRI, so that the benefits of this work continue to be felt across the research and innovation system.

This evaluation is presented through two complementary outputs. A full report, supported by a detailed technical annex, provides comprehensive analysis, while a high-level summary report distils the key findings and recommendations. The summary report and technical annex can both be found on the UKRI website.

It should be noted, at the time of the pilot's design, the term "responsive mode" was used across UKRI to describe funding that supported bottom-up, curiosity-driven research led by applicants, in contrast to more directive, top-down strategic calls. UKRI now refers to this approach as "applicant-led" funding, but for the purposes of this report we have continued to refer to this as responsive mode (RM).

Stakeholders

Project Board and Evaluation Advisory Group members

Alison Park, SRO, Chair of Project Board and Evaluation Advisory Group
Kedar Pandya, EPSRC, Project Board until November 24
Kirsty Grainger, UKRI Research culture, Project Board until September 23
Linda Holliday, UKRI Research culture, Project Board from September 23
Pete Burlinson, BBSRC, Project Board and Evaluation Advisory Group
Stephen Meader, UKRI Talent and Skills team, Project Board and Evaluation Advisory Group
Jane Nicholson, EPSRC, Project Board from December 24
Glenn Wells, MRC, Project Board from December 24
Brooke Rogers, King's College, London, Evaluation Advisory Group
Kim Douglas, Social Sciences and Humanities Research Council, Canada, Evaluation Advisory Group
Rachael Foy, NERC, Evaluation Advisory Group
Rick Hamilton, UKRI Evaluation and Performance team, Evaluation Advisory Group
Alex Hulkes, ESRC, Evaluation Advisory Group and data analysis
Callum Day, UKRI Talent and Skills Team, Evaluation Working Group
Luis Lopez-Bracey, UKRI Evaluation and Performance Team, Evaluation Working Group until August 2024
Ishara Kamalanathan, UKRI Evaluation and Performance Team, Evaluation Working Group from August 2024 to June 2025
Lydia Yeomans, UKRI Evaluation and Performance team, Evaluation Working Group from June 2025

Cross Council Working Group

Barry Knowles, AHRC
Inbal Ringel, BBSRC, from January 24 until October 24
Ruth Betterton, BBSRC, from January 24
Amelia Townley-Taylor, BBSRC, from October 24
Charlotte Cooper, BBSRC, from October 24
Luke Moody, ESRC
Claire Dyball, ESRC
Katherine Rooke, EPSRC, until January 25
Rebecca Williams, EPSRC, from September 25
Ethan Tull, EPSRC
Holger Apitz, MRC, until October 23
Natalia Kapel, MRC, from October 23

Lisa Hill, NERC, until July 23
Beth Taylor, NERC, until Sept 23
Vicki Durbridge, NERC, from Sept 23
Dave O'Gorman, STFC
Leen Van Broeck, STFC, from November 24 until Aug 2025
Phil Edwards, UKRI FLF
Louise Brown, UKRI Evaluation and Performance Team, until January 2024
Darren O'Keefe, UKRI Communications, until March 2025
Jess Phanwises, UKRI Policy

Catherine Lyall, Independent Consultant and Professor Emeritus of Science and Public Policy, University of Edinburgh – IAC training

Members of the Interdisciplinary Assessment College

Interdisciplinary responsive mode team

Alexandra Amey
Jenny Morris
Hilary McNeill
Samantha Aspinall, Horizons Institute University of Leeds, seconded to UKRI 2022-2025
Emma Gerster, from December 2023 to December 2025
Gemma Lupton-Weeks
Adam Woodall
Michala Pulley, until August 2025
Montana Walcott
Paulina Rzepa
Olatunji Adeniji, until August 2024
Caroline Parkes
Joey Tuffin
Joe Jackson, until August 2024
Lauren Millichip, from January 2025 until May 2025
Alicia Flight from February 2025

1 Introduction

This report presents the overall findings for the evaluation of the Cross-Research Council Responsive Mode (CRCRM) pilot scheme. It combines findings from evaluation activities undertaken by UKRI and two independent contractors, Technopolis and the Innovation Research Caucus (IRC).

The report is structured as follows:

- Chapter 1: Introduction including an overview of the CRCRM pilot, its assessment processes and the evaluation questions
- Chapter 2: Process evaluation findings, including feedback on processes from stakeholder surveys, interviews and panel observations
- Chapter 3: CRCRM portfolio, including analysis of application characteristics and interdisciplinarity
- Chapter 4: Impact on the portfolio, including how the projects funded by CRCRM compare to other benchmark programmes
- Chapter 5: Conclusions and recommendations

This report also includes appendices covering the methodology, Theory of Change and other supplementary information.

Technopolis collated the findings across the different partners. Findings from work undertaken by IRC and UKRI are clearly highlighted. The remainder of the findings are based on analysis conducted by Technopolis. Technopolis conducted stakeholder interviews, survey analysis, panel observations and analysed applications including their scores, interdisciplinarity and novelty.

1.1 Overview of the CRCRM pilot scheme

The CRCRM pilot was a £65m scheme launched in May 2023 as part of UKRI's corporate plan 2022-2025.[1] The scheme has now closed but awards funded through it are expected to run till July 2028.[2] It was managed centrally, by a UKRI team, rather than a specific research council. The scheme supported objective 3 of the UKRI corporate plan 'World-class ideas', which included a commitment to fund new ideas and enhance funding mechanisms for interdisciplinary research (IDR) schemes. As such, the CRCRM had four overarching objectives:

[1] https://www.ukri.org/wp-content/uploads/2022/08/UKRI-190822-CorporatePlan2022to2025.pdf

[2] Based on Round 2 funding call, which states grants will be awarded by December 2025 and can start between 1 January and 1 July 2026. Projects will run for up to 24 months, though project start date can be delayed up to 12 months.

- to fund novel IDR ideas that do not have a clear lead UKRI Research Council for responsive mode funding
- to support breakthrough or disruptive ideas; projects incorporating new concepts, approaches and methodologies; and unexpected interdisciplinary combinations
- to understand the demand of the research community for this type of scheme
- to test and refine the processes involved in assessing it[3]

The CRCRM scheme wasn't the only UKRI opportunity for funding IDR. It was complementary to existing UKRI schemes, such as UKRI Councils' existing responsive-mode schemes, supported by the Cross Council Research Agreement (CCRA) policy, and strategic / challenge-led IDR funding opportunities. However, the key feature of the CRCRM pilot was that it had a tailored approach to assessing IDR applications as well as specific guidance, unlike other UKRI funding opportunities.

Round 1 funding was launched in May 2023, and grants were awarded in June 2024 with projects expected to run up to February 2027. Round 2 funding launched in September 2024, with grants announced in January 2026. The total funding available was £65m to be awarded over two Rounds. The full economic cost of projects ranged from £200k to £1.2m, with UKRI covering 80% of the costs. It was expected that up to 36 awards could be funded in each Round.

1.1.1 Rationale and Theory of Change

IDR combines knowledge and/or methodologies from two or more disciplines to create new knowledge, technology, processes or art.[4] IDR plays an important role in enabling discoveries and breakthroughs, sometimes resulting in the development of new fields of study over the longer term (for example development studies, biochemistry and health economics).

Research funders worldwide have been increasingly encouraging IDR as a way to facilitate scientific and technological breakthroughs and tackle complex societal problems,[5] and thereby create more impact. Research funders often pursue IDR to tackle complex societal problems because monodisciplinary approaches may not always fully address them.[4] Moreover,

[3] Scoping Interviews conducted in July and August 2024

[4] Technopolis (2016). Case Study Review of Interdisciplinary Research in Higher Education Institutions in England. Available at https://emps.exeter.ac.uk/media/universityofexeter/emps/research/cee/news/Technopolis-Case_study_of_interdisciplinary_research_in_HEIs_in_England.pdf

[5] Technopolis (2016). Landscape Review of Interdisciplinary Research in the UK. Report to HEFCE and RCUK.

breaking out of the so-called 'disciplinary silos' is seen to facilitate knowledge production and innovation, which has led UK research funders (incl. the Research Councils, Government Departments and charities) to support IDR both in their individual capacities, and collaboratively (for example Newton Fund, Global Challenges Research Fund (GCRF), Strategic Priorities Fund (SPF)). Funders have helped to legitimise and promote IDR in institutions by financially supporting it through challenge-based strategic funding opportunities or applicant-led 'response-mode' applications.

However, two key challenges are seen to negatively affect funding of IDR:

First, IDR applications do not always fit the remit of one specific funder and/or its departments which are usually organised on a disciplinary basis, UKRI being a case in point. While a cross-research council mechanism exists in UKRI, applicants are required to identify a lead research council to submit to, which must account for at least 51% of the remit.[6]

Second, it can be difficult to find peer reviewers with appropriate expertise and knowledge of IDR to holistically assess IDR applications and monodisciplinary perspectives of reviewers are seen to adversely affect the funding of IDR. For example, anecdotal evidence suggests that even if an IDR application is reviewed by reviewers with expertise in the individual disciplines involved in the proposed project, reviews tend to be from the perspective of a single discipline which means IDR applications often fall between the "cracks" of the individual disciplines involved, with their added value (more than the sum of the parts) and reciprocal benefits (to involved disciplines) underappreciated.

In light of these challenges, the 2015 Nurse Review[7] recommended dedicated funding for IDR applications that cross boundaries between UKRI Research Councils and Innovate UK. The review also highlighted that "reviewing of inter-disciplinary research applications should be improved by using reviewers experienced in judging this type of research". Its recommendations to strengthen peer-review processes included provision of guidance and training for reviewers, development of a recognition system for high quality reviewers and working with reviewers from both the UK and abroad.

UKRI has successfully invested in programmes that cut across disciplines and sectors (for example, SPF, Industrial Strategy Challenge Fund (ISCF)). However,

[6] https://www.ukri.org/apply-for-funding/get-support-for-your-project/if-your-research-spans-different-disciplines/

[7] https://assets.publishing.service.gov.uk/media/5a801c6fe5274a2e87db7eaa/BIS-15-625-ensuring-a-successful-UK-research-endeavour.pdf

the 2022 Grant Review[8] noted that existing UKRI structures still limited cross-council working.

Thus, the CRCRM pilot was conceptualised drawing upon the insights from to the 2015 Nurse Review and the 2022 Grant Review, which recognised the key challenges and opportunities around IDR funding in the UK. UKRI's Theory of Change for the pilot can be found in Appendix A.

1.1.2 Definition of IDR

For the purposes of the scheme, UKRI adopted the following definition of IDR from the REF2021 Interdisciplinary Advisory Panel final report: "IDR is understood to achieve outcomes (including new approaches) that could not be achieved within the framework of a single discipline. IDR features significant interaction between two or more disciplines and / or moves beyond established disciplinary foundations in applying or integrating research approaches from other disciplines."

As such, multidisciplinary and transdisciplinary research were not within the scope of CRCRM. However, IDR involving non-academic stakeholders (which could also be considered transdisciplinary) was in scope.

1.1.3 The Interdisciplinary Assessment College (IAC)

The IAC is a key feature of the CRCRM scheme. It aimed to bring together IDR experts to assess IDR applications in a more appropriate way than the standard application assessments used by single council responsive mode schemes that use external written review and moderation panels. To enable this, UKRI's Interdisciplinary Responsive Mode (IRM) Team and the CRCRM working group recruited a cohort of experts and provided training focused on improving the assessment of IDR applications.

Webinars were delivered by UKRI to inform the research community about the recruitment call for IAC members. The main criterion for recruitment was expertise in IDR, and both academics and non-academic professionals were eligible to apply to the IAC. The recruitment process required the completion of a survey with questions based on the questions outlined in the Resume for Research and Innovation (R4RI).[9] However, as per interviews, this format proved challenging for a minority of the IAC members who described it as being overly long and unclear on how to demonstrate their experience in IDR and its assessment. In addition, the UKRI assessment team also noted that it

[8] https://www.gov.uk/government/publications/independent-review-of-uk-research-and-innovation-ukri/independent-review-of-uk-research-and-innovation-ukri-final-report-and-recommendations

[9] https://www.ukri.org/apply-for-funding/develop-your-application/resume-for-research-and-innovation-r4ri-guidance/

took time to be able to identify relevant IDR expertise within the IAC applications.

Some rare struggles with showcasing IDR expertise in the R4RI format notwithstanding, over 1,300 applications were received for the approximately 300 places in the IAC, comprising a panel member pool (~230) and a separate Chairs and Rovers pool (~70). The recruitment effectively attracted applications from academics, the private sector, clinicians and university professional services. IAC interviewees reported a good balance of UKRI council remit areas within the IAC.

IAC members were remunerated for their attendance at the virtual panel meetings and were still eligible to apply for funding through the CRCRM scheme. To avoid potential issues in the assessment process, the standard UKRI conflict of interest process was implemented.

1.1.3.1 IAC Training

Training for the IAC was delivered by the IRM Team and two external experts, Samantha Aspinall (Head of Interdisciplinary Research Development, Horizons Institute, University of Leeds) and Catherine Lyall (Independent Consultant and Professor Emeritus of Science and Public Policy, University of Edinburgh). The training aimed to provide IAC members with a clear understanding of the definition of IDR and an understanding of its meaning in the context of the CRCRM. It covered:

- theoretical aspects of assessing IDR including how to identify excellent IDR approaches in applications
- good practice in assessing IDR including providing feedback to applications
- learning through peer-to-peer discussions

Chairs training covered managing divergent opinions and building consensus. A separate technical briefing was held to support the IAC members with access to the application platforms and submitting their reviews.

The training sessions were delivered prior to the outline and full application assessment processes for both rounds.

1.1.4 Application and assessment process

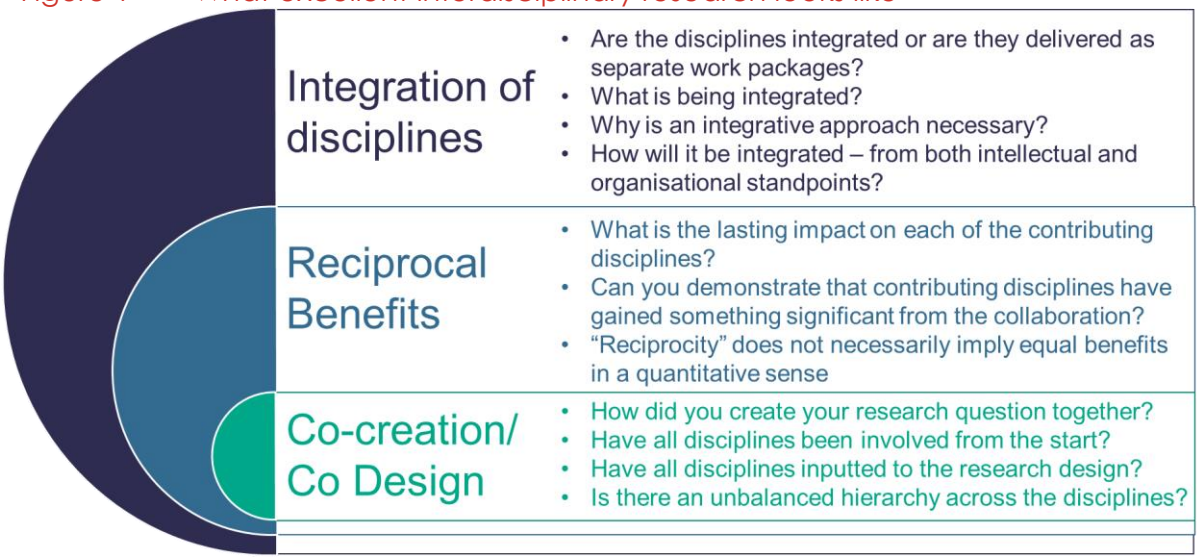
A two-stage process (outline and full application) was developed to reduce burden on the research community (responding to 2022 Tickell's Review on Research Bureaucracy). This marks a departure from the conventional responsive-mode funding process which involves a one stage application process.

The assessment process consisted of panel-only review to manage the issues of individual untrained reviewers giving less integrative perspectives in response to the recommendation in the Nurse review of using reviewers experienced in judging IDR.

1.1.4.1 Funding call

The call text for both funding rounds of the CRCRM pilot were drafted such that they clearly communicated the assessment process and the scope to the research community.[10] The text was mostly standard UKRI funding text covering the timeline, total funding and award range, eligibility of applicants and other standard funding call text elements. The unique elements of funding call text included explaining the purpose and objectives of the CRCRM pilot and assessment criteria with respect to IDR.

Figure 1 What excellent interdisciplinary research looks like



Source: UKRI

Learning from Round 1, in Round 2, the UKRI IRM team provided more detailed guidance for applicants[11], which included more clarity around the pilot's definition of IDR with the three components of integration of disciplines, reciprocal benefits and co-creation/ co-design (see Figure 1). A webinar targeted how to develop strong interdisciplinary research applications and

[10] Scoping Interviews conducted in July and August 2024, interview with the IRM team in September 2025

[11] https://www.ukri.org/wp-content/uploads/2024/04/UKRI-160824-Funding-Opp-CrossResearchCouncilResponsiveModePilotScheme-Round2GuidanceApplicantsResearchOrganisations.pdfhttps://www.ukri.org/wp-content/uploads/2024/04/UKRI-160824-Funding-Opp-CrossResearchCouncilResponsiveModePilotScheme-Round2GuidanceApplicantsResearchOrganisations.pdf

expanded in detail on the three components of excellent IDR for the Round 2 full stage applicants.[12]

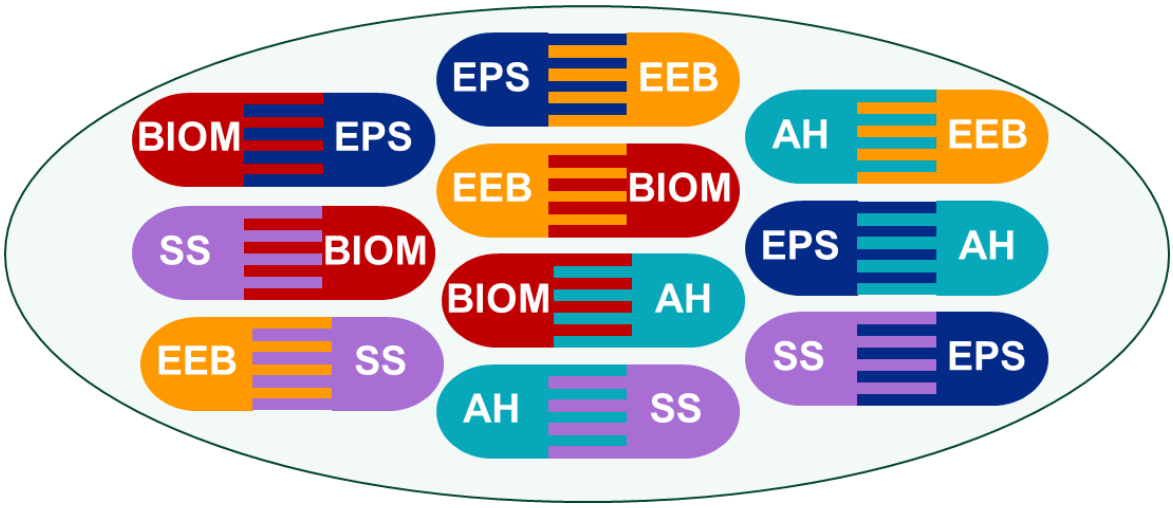
In addition, the councils provided information on how to access UKRI supported facilities and services in a separate webinar.[13] In response to feedback, Round 2 applicants were allowed to claim for an additional £250k, and delay the grant start date by up to 12 months, to enable them to access specific NERC facilities which the funding limitations prohibited in Round 1.

1.1.4.2 Panel structures

To manage the assessment of a large number of applications across the breadth of UKRI remit CRCRM applications were clustered into panels of a practical size. A panel structure was developed based on interfaces between five key research domains:

- Engineering and Physical Sciences (EPS)
- Earth, Environment and Biological Sciences (EEB)
- Biomedical and Health (BIOM)
- Arts and Humanities (AH)
- Social sciences (SS)

Figure 2 CRCRM interdisciplinary panel interfaces



Source: UKRI

The 10 interdisciplinary interfaces created a flexible framework and CRCRM applications were assigned to the most relevant interface according to their

[12] https://www.youtube.com/watch?v=qeGzYwEK8tA

[13] https://www.youtube.com/watch?v=_ORmTo7Poqw

research focus area. Applications that addressed more than two interfaces were put in the most appropriate panel with the assessors with the appropriate expertise. Where a high number of applications were submitted for an interdisciplinary interface, it was split into multiple panels.

In order to ensure consistency of the panel recommendations roving college members, or rovers, from the IAC chair pool joined at least three panels at each stage and round. Rovers did not actively participate in the panels but observed how each panel operated.

A tensioning meeting was held at the end of the panel week attended by the chairs, deputy chairs, rovers and UKRI conveners. Panel chairs reported back to UKRI on the outcomes of the panel assessments and in a separate session rovers reported back on their observations, highlighting if there were any differences between panels in the way in which final scores and bandings were agreed. The chairs and rovers also provided feedback on the assessment process in the tensioning meeting, highlighting areas for improvement.

UKRI made the final funding decision for applications across the panels based on the recommendations from the panel chairs and feedback from the rovers.

1.1.4.3 CRCRM application scoring system

The standard UKRI assessment panel scoring system was used. It consists of 10-point system in which 10 indicates application is exceptional and meets all assessment criteria, and '1' indicates the application is unsatisfactory and does not meet any assessment criteria.[14]

Feedback from the panel members and rovers indicates that assessors struggled to agree what would be considered a weakness in an application and be able to score an application above a 7. Therefore, the IRM team developed some clear guidance on what is meant by minor to clear weaknesses based on consensus arrived at in peer-to-peer discussions during the training sessions. This guidance was included in the updated panel guidance for the Round 2 full stage panels.

Each application was scored on the 10-point scale with an additional high (H), medium (M), and low (L) qualifier. After scoring, applications were assigned to 'assessment bands':

- **Band A* (Full stage only)** - these applications are judged to be of the highest standards and very strongly align with the CRCRM scheme

[14] CRCRM ROUND 1 Full stage Quick Reference Guide

objectives. They are outstanding or exceptional and can be clearly distinguished from the other applications. **They are the top priority for funding.**

- **Band A:** These applications are judged to be of the highest standards and strongly align with the CRCRM scheme objectives. **They are a priority for funding/ inviting through to full stage.**
- **Band B:** These applications are of a high standard and have good alignment with the CRCRM scheme objectives. **They should be considered for funding/inviting through to full stage, but are not prioritised.**
- **Band C:** These applications meet the standards required and have some alignment with the CRCRM scheme objectives. **They are fundable/worth inviting through to full stage, but are the lowest priority.**
- **Band D:** These applications of did not meet the standards required and have limited alignment with the CRCRM scheme objectives. **They are not recommended for funding/ inviting through to full stage.**
- **Unbanded (Outline stage only):** These applications received consistently low scores by the panel members and were agreed not be discussed further at the meeting by the panel. **They are not to be invited through to full stage.**

The banding process was based on that used by the Future Leaders Fellowship (FLF) scheme. CRCRM used the FLF banding definitions but evolved them over the course of the outline/full stages of both rounds of the scheme in response to feedback where panels struggled to allocate applications to the bands.

The banding of applications enables comparisons across multiple panels to assess applications under the same scheme and budget. It mitigates against potential differences in how panels have operated and scored. Applications that can be considered of equal quality are banded together. Banding was then used to determine the awarding of funding.

The next sections discuss how the two application stages were delivered.

1.1.4.4 Outline application stage

In this stage, applicants were asked to complete short outline applications. The applications for Round 1 were submitted in the Joint Electronic Submission (Je-S) system, whereas applications for Round 2 were submitted via its replacement, the UKRI Funding Service (TFS).

Once outline applications were received, UKRI conducted a 'remit check' to determine if the proposed research included disciplines from more than one Research Council. Each outline application that passed the remit check was

tagged with the remits involved and assigned to one of the interdisciplinary interfaces (Figure 2) and was matched to three IAC members based on their expertise. The three assigned IAC members were referred to as 'introducing members' (IMs), whose role is to lead the assessment of the application and IM1 will introduce the application and its assessment at the panel meeting.

To manage the high volume of applications, UKRI was supported by Digital Science to assign the IMs from across the college. Using their Dimensions platform, Digital Science identified up to 10 potential college members for each application, matching their expertise with the work proposed. UKRI then selected the three IMs from the Dimensions suggestions, using the applicant-assigned keywords and remit tags as additional information to make appropriate matches.

In Round 1, each IM independently submitted scores and comments for two separate criteria per application: (1) interdisciplinarity and (2) advancement, novelty and timeliness. Due to the high volumes of applications, IMs did not meet as panels. Rovers viewed a selection of the scores and comments, to check that scores were consistent with comments, and to compare consistency across panels. Chairs and deputy chairs met in groups of three in moderating meetings for each interface to discuss applications with divergent scores and agree a consensus score based on the IM scores and comments. They then assigned all applications to the four bands to allow progression to the next stage. 13 moderating meetings were run over two days with duplicate panels for three interfaces.

For Round 2, applicants were required to select keywords from a predefined list in their application to improve efficiency, reliability and consistency in assigning applications to the most appropriate panel. Assignment of the IM roles for each application was confined to members within the assigned panel rather than the wider college, which made the process of matching more efficient for UKRI. In Round 2, a single unified score was used against four criteria: (1) Fit to opportunity, (2) Vision, (3) Approach and (4) Applicant and team capability to deliver, to align with the new harmonised UKRI assessment criteria. Assessment panel meetings were held with all assessors to discuss all applications to arrive at overarching consensus scores and assignment to the bands. Prior to the meetings applications with consistently low scores were sifted out and assigned as unbanded. 17 panels across 10 interfaces were held over four days with a tensioning meeting held on the fifth day. Two panels each were convened for five interfaces, three panels for one interface and one panel meeting each for the remaining four interfaces.

The main differences between the Round 1 and Round 2 outline stage assessment processes are outlined in Table 1 below.

Table 1 Changes between Round 1 and Round 2 Outline Stages

	Round 1 Outline Stage	**Round 2 Outline stage**
Application route	JeS (outlines only)	UKRI Funding Service (TFS)
Timing	Timing of key application dates were considered tight.	Timings were improved for applicants and college members
Number of applications	979 applications submitted, 77 out of scope, 3 withdrawn. 899 applications assessed. No demand management was imposed.	576 applications submitted, 31 out of scope, 10 withdrawn/rejected. 535 applications assessed. Demand management was put in place, with institutional caps applied. 1060 applications were submitted through the research organisation internal sifts.
Remit Areas and matching of assessors to applications	IAC members provided free keywords to define their expertise, on application to the college. Applicants to the scheme provided free keywords to indicate the disciplines involved in their proposed research.	IAC members provided their areas of expertise from a list of standardised research areas, and also indicated their preferences of interface panels to participate in. This also enabled a refresh of IAC expertise, including filling gaps. Applicants used the same list of research areas to select from at the point of application.
Training of IAC	Training was held for all IAC members, with additional training for Chairs. CRCRM case studies of awarded applications were not available at this point but adapted FLF applications were used as case studies. Technical information was provided in addition to the training sessions.	Further training was held, which included peer-to-peer learning and exercises based on CRCRM case studies from ROUND 1. Technical briefings were held separately from main training.
Guidance	Assessor guidance was provided for IAC members explaining how to assess IDR and finding consensus for Chairs	Assessor guidance provided improved details on IAC roles in panels. Applicant guidance provided improved clarity on how applicants should use R4RI and how team capability to deliver is assessed.

Assessment Criteria	Two equally weighted criteria: Advancement, Novelty and Timeliness and Interdisciplinarity.	Four equally weighted criteria: Fit to Scheme (replaced interdisciplinarity score); Vision; Approach Team Capability to deliver. Further guidance was provided under each criterion.
Scores and Comments	Panel members (introducers) submitted scores and comments on the Extranet. Chairs and rovers could view these in advance of meeting.	Panel members (introducers) submitted scores and comments on the Extranet. All panel members could view these in advance of meeting.
Pre panel sift	No pre panel sift	UKRI sifted out applications with consistently low scores prior to the meeting for effective use of meeting time
Panel Meetings	13 panels in total (3 interfaces were split into 2 panels each). Of the 13 panels, 10 ran over 2 days. Panel members did not attend panel meetings. Chairs met virtually to discuss and moderate divergent scoring applications only. The decision on final scores and banding rested with Chairs. Rovers observed multiple panels.	17 panels in total all as 1-day meetings over 4 days (1 interface was split into 3 panels, 5 interfaces were split into 2 panels, 4 interfaces had a single panels). Panel members attended virtual one-day panel meetings. The decision on final score and banding was made by the panel members, led by the Chair. Rovers observed multiple panels.
Scoring	Two scores were provided, one for each criterion; reconciling of these during banding was found to be challenging	A single score was provided to cover all equally weighted criteria.
Banding	Final agreed scores were banded.	Final agreed scores were banded. Banding definitions were clarified, to differentiate them from the scoring definitions and focus on alignment with scheme objectives and not on the added value of an award.
Feedback	Generalised feedback was provided for all applicants, after their outcomes were communicated. Applicants considered this to be a lack of transparency.	Generalised feedback was provided. In addition, clarification statements, chosen from pre-written set of statements, were selected by Deputy Chairs during the panel meetings, and shared with applicants.

1.1.4.5 Full application stage

In this stage, applicants were asked to complete a full application building on the outline application, including sections such as project vision and approach, team capability to deliver, ethics and responsible research and innovation (RRI), resources and cost justification. In Round 2, an additional section on fit to opportunity had to be completed. The applications were submitted via the TFS.

Full applications were assigned to one of the 10 interdisciplinary interfaces and reviewed by three IAC members from the panel (IMs).

The three assessors (IMs) discussed feedback on discussion boards and agreed a set of questions that could clarify any concerns before sharing with applicants as a single combined set of feedback for applicants to respond to. This avoided the risk of three disparate pieces of feedback potentially viewed through separate disciplinary lenses but also still allowed the process to include a project lead response stage. IMs submitted scores and comments prior to the panel meeting based on the full application and the applicant response. All applications were discussed at the virtual panel meeting and the IMs agreed on final scores. In Round 2, each application was also assigned to two readers who contributed to the panel discussions and were available to replace IMs who could not attend the meeting at short notice.

After all applications had been scored the panel allocated applications into the five assessment bands. With the exception of band A*, there were no limitations on the number of applications assigned to each band. Panels were given an allocation that could be assigned to band A* which was equivalent to half of the overall success rate anticipated for the round. The definition of a A* application states that these applications are outstanding or exceptional and therefore applications scoring only 9 or 10 could be put into this band. Allocating an A* band indicates that the Panel would not want the applications to take their chances in the randomisation process because of the strength of endorsement from the review process. If a panel could not distinguish between a larger number of applications that meet the A* definition than their allocation, all the applications that the panel agreed met the criteria of band A* would be randomised after panels to create a ranked list to determine the A* allocation for that panel. The remaining applications would then be put into Band A.

The main differences between the Round 1 and Round 2 full stage assessment processes are outlined in Table 2 below.

Table 2 Changes between Round 1 and Round 2 Full Stages

	Round 1 Full Stage	**Round 2 Full stage**
Application route	TFS – Vision and Approach in attached pdf document	TFS – all sections
Timing	IMs had 4 weeks for initial assessment and discussion boards	Longer IM initial assessment and discussion boards due to summer holidays (IMs had 6 weeks to review over summer and 2 weeks for discussion boards)
Number of applications	175 invited, 170 submitted; 19.5% success rate between outline and full stage in Round 1. 21% success rate at full stage	Fewer applications invited to the full stage: 100 invited, 98 submitted; 17% success rate. 33% success rate at full stage.
Remit Areas and matching of assessors to applications	Applications assigned to the same panel interfaces as the outline stage.	As with Round 1.
Training of IAC	Focused on the language of IDR applications, distinguishing between excellent and good applications, and how these map onto the assessment criteria. Ran a separate Chairs session on the challenges of chairing IDR panels, how to manage divergent opinions and building consensus.	More peer-to-peer learning, IAC members discussing good practice in the reporting process, using Padlet to gather good practice ideas. Combined Chairs and panel member training on same day (panel members left and Chairs stayed on). Chairs session focused on the effective facilitation of consensus building and best practice for the banding process.
Guidance	Applicant guidance provided was limited to call text	Further applicant guidance was provided through a briefing webinar on: - how to demonstrate interdisciplinarity in applications with examples of integration, co-creation and reciprocal benefits - how risk and ethics are assessed - the project lead response stage - what costs are and are not permitted - the roles of project partners vs sub-contractors and the requirements for collaboration agreements and IP management.
Assessment Criteria	5: Vision, Approach, Capability to Deliver, RRI, Resources and justifications	6: Fit to opportunity, Vision, Approach, Capability to Deliver, RRI, Resources and justifications.

		Additional question on IP management created but not assessed to mitigate the need for letters of support from project partners.
Scores and Comments	Panel member conducted initial assessments then discussed on Discussion Boards and compiled feedback and questions for the Project Lead. IMs submitted pre-scores and comments once having reviewed the PL response without further discussion before the panel meeting	Additional guidance on defining weaknesses in the scoring definitions.
Panel Meetings	10 panel meetings over 4 days with tensioning meeting on day 5. Some interfaces had two-day panel meetings. Panel meetings attended by all panel members. Each application assigned 3 introducers. 15-minute discussion time. Panel meetings on Zoom. Rovers observed multiple panels.	10 panel meetings over 4 days with tensioning meeting on day 5. Considered panel size and split interfaces where application numbers too large (EPS-EEB; 1a and 1b) and combined where too small (panels 7 and 9, AH-BIOM-EEB). All panels were one day meetings. Panel meetings attended by all panel members. Each application also assigned 2 Readers. 20-minute discussion time. Panel meetings on Teams. Rovers observed multiple panels. Most panels had 2 external rovers in attendance
Banding	After scores agreed applications were banded. Used the same banding definitions used at the outline stage. In banding process, bands A-D were agreed before A*s were selected from the A band.	Updated banding definitions to refer to standards and meeting the criteria, rather than quality of the application. In the banding process the decision on Band A* was made at the same time as A-D banding.
Feedback	Individual feedback written by IM1	Individual feedback written by IM1 Guidance given to ensure that the post-panel feedback is consistent with the pre-panel feedback.
Evaluation		Interdisciplinary panel ratings provided by the Deputy Chair, for evaluation purposes

1.1.4.6 Governance of decision making and awarding of funding

Based on the feedback from the tensioning process the CRCRM Project Board reviewed and signed off the banding outcomes from the panels to decide on applications to be invited through to the full stage or to be funded.

At the outline stage, the Project Board conducted cross-panel "tensioning" to address variability in scoring across panels and ensure a balanced portfolio and consistent opportunity for success regardless of assessment panel. Using panel feedback, success rate comparisons, and outcomes data, the Board reviewed and, where appropriate, uplifted high-quality applications near the sift threshold to balance success rates and maintain equitable opportunities for applicants. This process also incorporated equality, diversity and inclusion (EDI) considerations, with monitoring of outcomes across protected characteristics and applying proportionate uplifts to advance equality of opportunity in line with the Equality Act 2010, without compromising overall quality. In Round 1, 44 outline applications were uplifted from Band B and invited through to full stage; of these, 33 uplifts were made to tension success rates across the panels and a further 11 Bs were uplifted to balance success rates for protected characteristics and regional spread. In Round 2, due to the change in process, the panel outcomes were more consistent. The Project Board agreed to uplift 21 Band B applications and invite them through to full stage; 12 uplifts were made to tension success rates across the panels and a further 9 Bs were uplifted to balance success rates for protected characteristics.

For the full stage applications, once the final bandings were approved, the Project Board signed off the band A* applications for funding. The Board then awarded the remaining budget, initially to band A applications and implemented a partial randomisation process where necessary. In Round 1 the number of applications in bands A* and A was greater than the funds available. The Project Board conducted two randomisation processes; one where a panel exceeded their A* cap, and one to randomise all the Band A applications to create a rank-ordered list. The Project Board awarded funding down the list of Band A applications as far as the budget would allow (18 out of 27 applications). In Round 2 the budget allowed for all Band A* and Band A applications to be awarded but was not able to award any Band B applications. Therefore, a partial randomisation process was not required.

After each round a joint meeting of the Project Board and panel chairs was held to present the outcomes and data analysis to the chairs, and for the Project Board to hear reflections from the chairs on the applications and assessment process.

1.2 Evaluation

The scheme was evaluated in real time to understand whether the pilot achieved its objectives and also to understand how UKRI can assess IDR most effectively. Interim evaluation findings and lessons learned from Round 1 were incorporated into the Round 2 processes.

The primary aim of the evaluation was to assess the appropriateness (process evaluation) of the CRCRM scheme and its effectiveness (impact evaluation) to understand in the short term:

- the demand from the research community for the CRCRM scheme
- whether CRCRM is meeting a funding gap in UKRI funding opportunities
- to generate evidence and learning to inform future interdisciplinary funding opportunities and assessment processes
- to identify wider learnings that apply generally to UKRI processes or establishing good practice

This evaluation report covers the following components:

1. Process evaluation components undertaken by Technopolis Ltd. and UKRI looking at whether the approach to scheme design and process contributed to accomplishing its stated objectives
2. Strategic project component undertaken by the IRC which focuses on learning more about how researchers and assessors approach IDR with an aim to understand what challenges, biases and advantages/ disadvantages remain and may be inherent to IDR environments

An additional separate mixed-methods study of the CRCRM awarded grants is being undertaken by a research team at the Science Policy Research Unit (University of Sussex Business School), King's College London Policy Institute, Vitae and RAND Europe, funded by the UK Metascience Unit.

1.2.1 Evaluation questions

The evaluation sought to answer the following **three evaluation questions** (EQs):

- EQ1. *Whether* and how the scheme and its processes are of high quality, efficient and effective in the assessment of IDR applications across all disciplines to inform best practice *and* approaches for the assessment of IDR moving forward (Process evaluation)
- EQ2. Understand how and to what extent the scheme's processes have enabled the effective assessment and funding of new IDR and impacted perceptions of IDR support in the UK funding landscape (Impact of the process on the portfolio)

- EQ3. Understand the impact of the scheme on the nature of the applications and the portfolio of awards supported, and the R&I communities' perceptions and behaviours with respect to funding routes for IDR (Additionality of the CRCRM pilot)

We adopted a mixed-methods approach combining both quantitative and qualitative data collection and analysis, followed by triangulation across these various methods to address all the EQs in turn. Details of the EQs and sub-EQs addressed and the methodology, including surveys (2274 responses) and interviews (n = 67) with applicants, IAC members and UKRI staff, are provided in Appendix B.

2 Process evaluation findings

The process evaluation questions sought to understand if and how the scheme and its processes are of high quality, are efficient and effective in the assessment of IDR applications across all disciplines to inform best practice and approaches for the assessment of IDR moving forward. The findings are presented in the remainder of the section.

Summary findings

Discerning IDR is challenging. The definition of interdisciplinarity was a persistent source of confusion for applicants and assessors across both rounds. Interdisciplinarity is often used as an umbrella term or interchangeably with multi- or trans-disciplinarity. The scheme is the first UKRI scheme to specifically ask applicants to demonstrate how their application is interdisciplinary through articulation of integration, co-creation and reciprocal benefits between disciplines, which some applicants found challenging. The scheme has successfully created legacy material (including guidelines and training material) which should be considered on future schemes.

Panel-based deliberation appears to allow for fairer and more robust assessment of IDR. The move from moderated scoring (Round 1 outline) to full panel discussion (Round 2) produced fairer, more differentiated outcomes. This indicates that assessing IDR requires collective sense-making, not aggregated solo judgements. This also aligns with reports that single-council responsive-mode processes involving individual written review disadvantage IDR applications.

The IAC played a key role in the success of the scheme and is a valuable asset for UKRI. Building a trained college of IDR assessors took considerable effort, and the quality of assessment improved markedly as the college gained experience. There is a strong case for UKRI to retain and develop this resource for future IDR schemes rather than starting from scratch.

2.1 Overview of application data

Table 3 provides an overview of the numbers of outline and full stage applications as well as awards by remit area for Round 1 and Round 2.

Analysis of the applications (excluding those that were withdrawn or rejected) indicates that across the two rounds, the most popular interfaces in terms of the number of outline applications received were:

- Engineering and Physical Science - Earth, Environment and Biological Sciences (EPS-EEB) – 20% of all outline applications.
- Biomedical and Health - Engineering & Physical Sciences (BIOM-EPS) – 20%

- Social Sciences - Engineering & Physical Sciences (SS-EPS) – 13%

The proportion of applications for these and the other interfaces were largely similar across the two rounds, the different application stages and awards.

Table 3 CRCRM applications by remit area and application stage for Round 1 (a) and Round 2 (b)

(a) Round 1

Full stage panel no. (outline stage panel no.)	**Remit area ('primary' interface)**	**Number of outline applications**	**Number of full stage applications**	**Number of awards**
1 (1, 2)	EPS-EEB: Engineering and Physical Sciences & Earth, Environment and Biological Sciences	183	41	6
2 (3, 4)	BIOM-EPS: Biomedical and Health and Engineering & Physical Sciences	182	27	6
3 (5, 6)	SS-EPS: Social Sciences and Engineering & Physical Sciences	118	24	7
4 (7)	SS-BIOM: Social Sciences (SS) & Band Biomedical and Health (BIOM)	56	8	1
5 (8)	EEB-SS: Earth, Environment and Biological Sciences & Social Sciences	89	17	2
6 (9)	EEB-BIOM: Earth, Environment and Biological Sciences & Biomedical and Health	80	16	5
7 (10)	BIOM-AH: Biomedical and Health & Arts and Humanities	23	6	2
8 (11)	AH-SS: Arts and Humanities & Social Sciences	59	9	2
9 (12)	AH-EEB: Arts and Humanities & Earth, Environment and Biological Sciences	42	11	3
10 (13)	EPS-AH: Engineering and Physical Sciences & Arts and Humanities	67	11	2
Rejected/ withdrawn		80	5	
Total		**979**	**175**	**36**

(b) Round 2

Full stage panel no. (outline stage panel no.)	Remit area ('primary' interface)	Number of outline applications	Number of full stage applications	Number of awards
1a-1b (1a–1c)	EPS-EEB	103	20	6
2 (2a–2b)	BIOM-EPS	90	14	4
3 (3a–3b)	SS-EPS	69	10	4
4 (4a–4b)	SS-BIOM	45	12	3
5 (5a–5b)	EEB-SS	71	13	4
6 (6)	EEB-BIOM	29	6	4
7_9 (7)	BIOM-AH	18	4	2
8 (8a–8b)	AH-SS	44	6	1
7_9 (9)	AH-EEB	25	6	3
10 (10)	EPS-AH	41	7	1
Rejected/ withdrawn		41	2	
Total		**576**	**100**	**32**

Note: The full stage panel meetings for the BIOM-AH and AH-EEB interfaces were combined because of the small number of full applications to be covered.

Compared to Round 1 where 49% of the applications in both application stages corresponded to Science, Technology, Engineering and Mathematics (STEM) interface panels (EPS-EEB, BIOM-EPS and EEB-BIOM), this represented around 40% in both application stages in Round 2. It is possible that improved understanding of the scheme's scope and demand management at research organisations resulted in a lower proportion of applications for the STEM-STEM interfaces in Round 2 compared to Round 1.

IRC analysis also noted some disparities in the number of "pure" arts, humanities and social sciences (AHSS) and "pure" STEM combinations. "Pure" collaborations are projects involving only AHSS or only STEM disciplines. IRC analysis also noted some disparities in the number of non-STEM collaborations (i.e., AH-SS remit area) compared to STEM (i.e., EPS-EEB, BIOM-EPS, EEB-BIOM remit areas). IRC found that only three pure non-STEM collaboration projects (AH-SS) were funded across all rounds, compared to 31 pure STEM. The remaining 34 projects include both STEM and AHSS disciplines in some combination.

The shorter cognitive distance between STEM areas where IDR may be easier and/or more well-established because of previous UKRI strategic funding opportunities may have led to a higher number of STEM-led IDR applications. It should be noted that this does not imply no social science (SS) or arts and humanities (AH) disciplines were included in STEM-STEM applications, since the allocation to panels is based on assessment of the two primary disciplines being combined in the proposed project during the remit checks.

Applications with AH disciplines at the 'primary' interface were less well represented among the applications received (about a fifth and a quarter of applications in Round 1 and Round 2 respectively). The difference in application numbers between AH and other disciplines may reflect the total number of researchers active in the corresponding fields in the UK. For instance, REF2021 data included submissions from about 38,000 STEM researchers (FTE under Panel A and B) compared to around 23,000 SS researchers (Panel C) and 14,000 AH researchers (Panel D).[15]

It is possible that the two-council requirement may have limited applications from funding councils that represent more disciplines. Two AHSS councils represent 19 disciplines each, while the STEM disciplines are represented by five councils responsible for an average of seven disciplines each. The objective of the CRCRM scheme was to address the funding gaps between the single councils and therefore, to be in scope, applications had to cover the remits of at least two councils. With five STEM-based councils and two AHSS councils, the scheme would therefore expect to receive more purely STEM applications. Applications falling into the three purely STEM interfaces accounted for 47% of outline applications compared to only 7% of applications in the single purely AHSS interface. The other 46% of applications were in the six STEM-AHSS interfaces.

Cognitive distances between disciplines within a single non-STEM council (AHRC or ESRC) can be quite large compared to *between* STEM councils, for example: the disciplinary difference between music/visual arts and languages and literature, all within AHRC, could be seen as wider than between chemistry (EPSRC) and atmospheric physics and chemistry (NERC) but as the CRCRM scheme's purpose was to address the funding gaps between councils only the latter was eligible under the two-council requirement.

[15] https://2021.ref.ac.uk/results-analysis/submissions-data/index.html

2.2 Quality and effectiveness of assessment processes

2.2.1 Analysis of panel scores

Summary findings

- Panel discussions generally increased differentiation between applications beyond initial individual scores, suggesting the consensus-building process added real value compared to, for example, the process undertaken in Round 1 outline stage, where only applications with divergent scores were discussed in a moderation meeting with three panel chairs.
- A single scoring scale as applied from Round 1 full application stage onwards seems to be more suited for IDR assessment compared to the two separate scores used in Round 1 outline stage. In the latter, interdisciplinarity scores showed more variation across panels than scores for Advancement/Novelty/Timeliness, suggesting assessors found it harder to apply the IDR criterion.
- Band B and C applications dominated the full application stage in both rounds, accounting for 55–60% of banded applications. Band A typically accounted for 15% of outline applications with 26-32% of applications graded as Band A or A* at full application stage; in parallel the proportion of applications graded as Band D fell between the outline and full application stages as expected, highlighting the value of the two-stage application process.

2.2.1.1 Heterogeneity in score distributions between and within panels

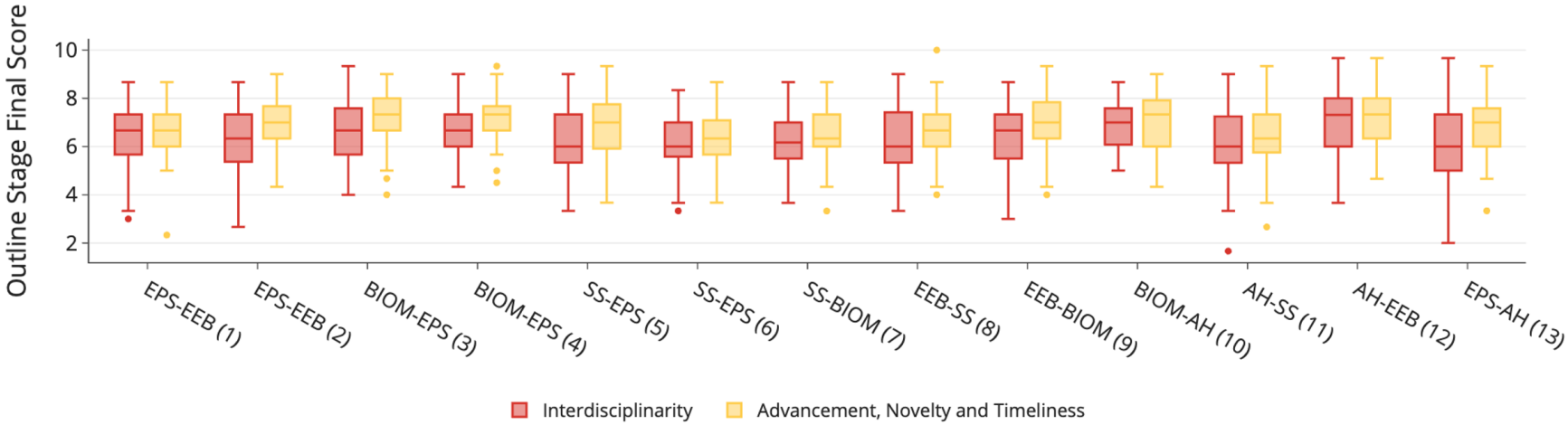
Anonymised panel assessment scores for outline and full-stage applications were shared with the evaluation team for analysing the distribution of scores within and between panels.

Figure 3 presents the distribution of the Round 1 scores. Figure 3(a) shows the final moderated scores awarded during the Outline Stage for two assessment criteria, Interdisciplinarity (red) and Advancement, Novelty, and Timeliness (yellow), across each of the thirteen panels. Figure 3(b) shows the distribution of average of three individual introducing member (IM) scores per application vs the final consensus scores after panel discussion at full application stage. It should be noted that for 24 applications, only two out of the three IM scores were available and for two applications only one score was available.

Figure 3(a) shows that interdisciplinarity scores for outline applications tend to show greater variation within several panels, particularly Panel 13 (EPS-AH), as reflected by the wider interquartile ranges (IQR) and longer whiskers. This may indicate that evaluators within this panel applied more diverse interpretations

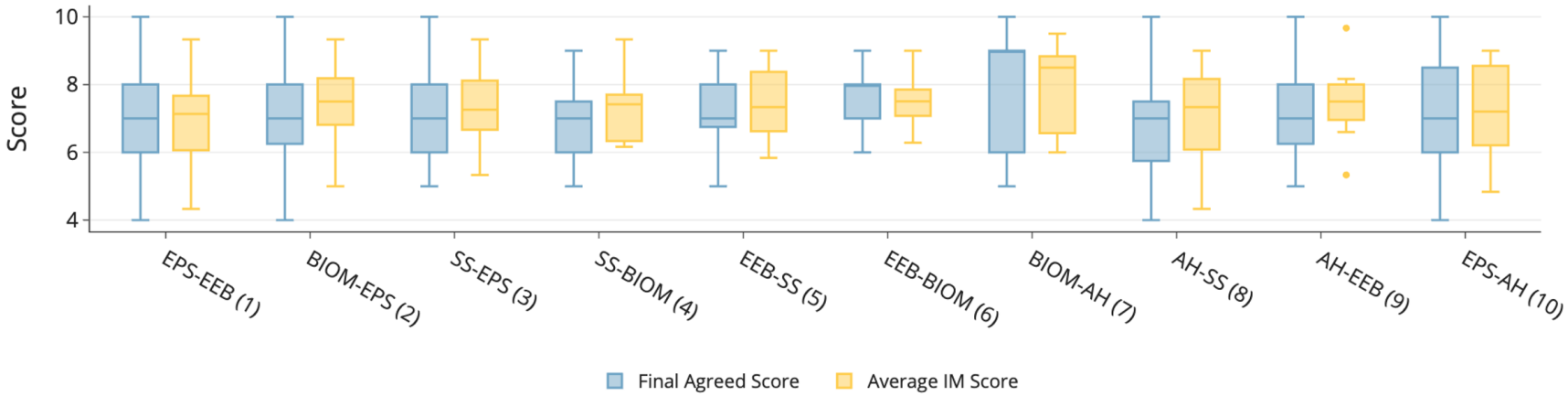
of what constitutes interdisciplinarity, potentially reflecting disciplinary differences between fields or differing levels of confidence in assessing the criterion.

Figure 3 (a) Round 1 applications' scores distribution for (a) outline and (b) full stage



Source: Technopolis analysis of CRCRM Round 1 application data. Note: Numbers in brackets represent the panel numbers.

(b)



Source: Technopolis analysis of CRCRM Round 1 application data. Note: Numbers in brackets represent the panel numbers.

In contrast, Figure 3(a) also shows that scores for Advancement, Novelty, and Timeliness appear less variable across most panels, suggesting greater agreement among panel members when evaluating this criterion. This could be the result of clearer guidance or greater familiarity with the concepts being assessed.

When interpreting the Round 1 outline stage scores, it is important to remember that the CRCRM assessment criteria and processes were being applied for the first time and that the outcome decisions for applications were made through moderation meetings with three panel chairs and not discussed by the IMs.

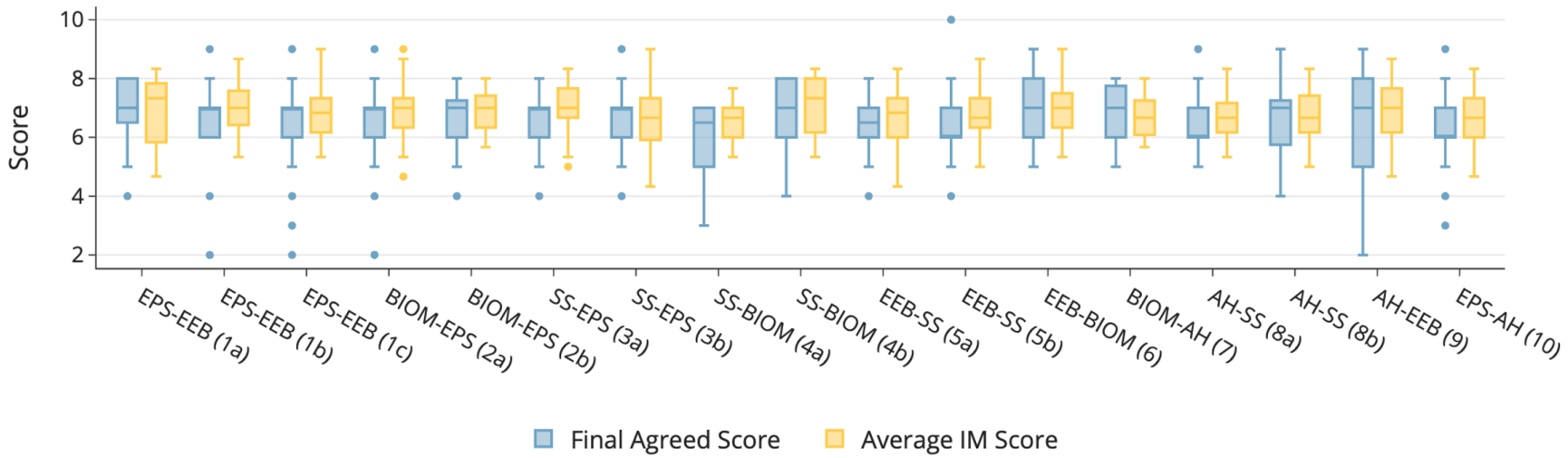
The distribution of average IM scores and final agreed scores across panels in Figure 3(b) shows broadly similar scoring patterns, with most panels exhibiting median scores in the range of 7–8. Average IM scores tend to have lower standard deviations and narrower interquartile ranges in most panels (see Table 12 in the Technical Annex for descriptive statistics). On the other hand, final agreed scores generally show greater variability than the average IM scores, with wider score ranges and higher standard deviations in several panels. Moreover, the final agreed scores following deliberation span a broader range and include the maximum score of 10. The wider dispersion suggests that panel consensus processes introduced additional differentiation between applications beyond the initial reviewer averages.

Figure 4 displays the distribution of the Round 2 scores before and after deliberation and consensus building, that is, the average IM scores and final agreed consensus scores respectively. Figure 4 (a) shows the scores awarded at the outline stage, while Figure 4 (b) shows the scores at full application stage.

Findings for Figure 4 (a) and Figure 4 (b) are similar to those in Figure 3(b). In Figure 4 (a), both score distributions show relatively similar central tendencies across panels, with most median scores in the 6–7 range, suggesting broad alignment between the average IM assessments and the final agreed panel scores. Figure 4 (b) shows a broadly consistent relationship between average IM scores and final agreed scores across panels, with most panels exhibiting median scores in the 7–8 range.

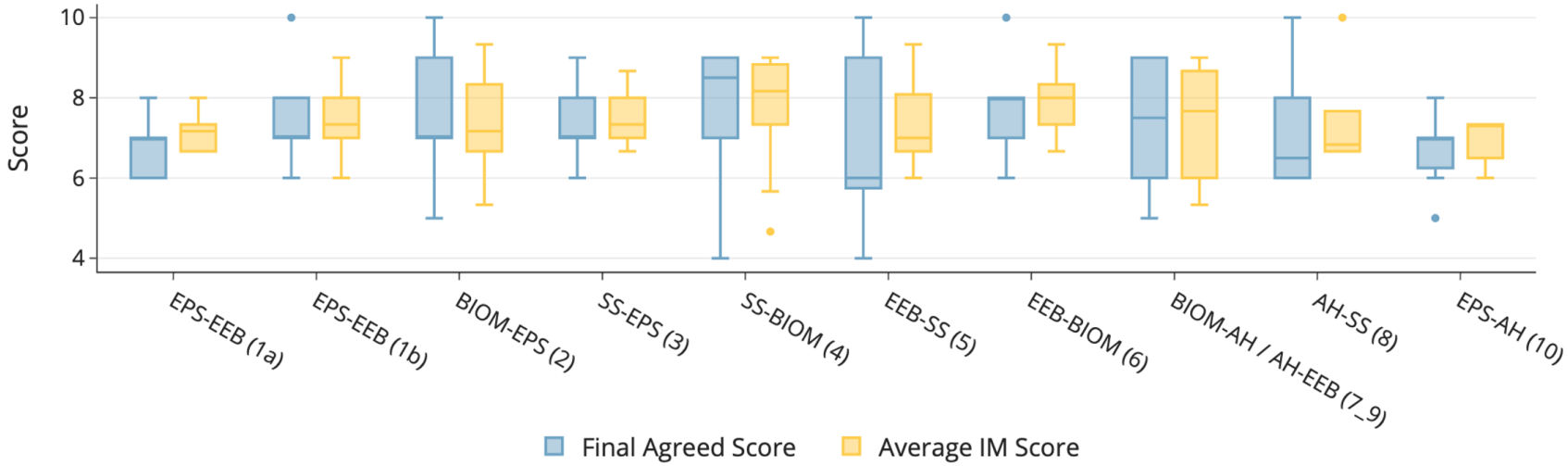
However, the final agreed scores for both outline and full applications in Round 2 generally display greater variability than the average IM Scores, as reflected in the wider interquartile ranges and broader whiskers observed for several panels. Average IM scores have lower standard deviations and narrower score ranges across most panels than final agreed scores (see Table 2 in Technical Annex). This reinforces the finding that the panel discussions introduced additional differentiation between applications, more so for the

Figure 4 (a) Round 2 applications scores distribution for (a) outline and (b) full stage



Source: Technopolis analysis of CRCRM Round 1 application data. Note: Numbers in brackets represent the panel numbers.

(b)



Source: Technopolis analysis of CRCRM Round 2 application data. Note: Numbers in brackets represent the panel numbers.

full applications than the outlines. This may be because of more project details available in the full applications, more time available for deliberation and/or the fact that the scoring will affect funding decisions at this stage of the assessment.

Literature shows that greater variability in the scores on a panel is not necessarily an indication of low validity or low legitimacy.[16] It may indicate that the panel is highly competent with a wide variety of views on what is good and valuable research, which are duly considered in the assessment.

Overall, the results across Round 1 and Round 2 indicate a strong correspondence between average IM scores and final agreed scores across panels, while also showing that the deliberation and consensus building tended to improve differentiation between applications, that is, it made it clearer which applications fit within the scheme and are high quality and fundable.

2.2.1.2 Distribution of application scores by band

The Technopolis team analysed the distribution of scores and bands to understand banding thresholds across panels and rounds.

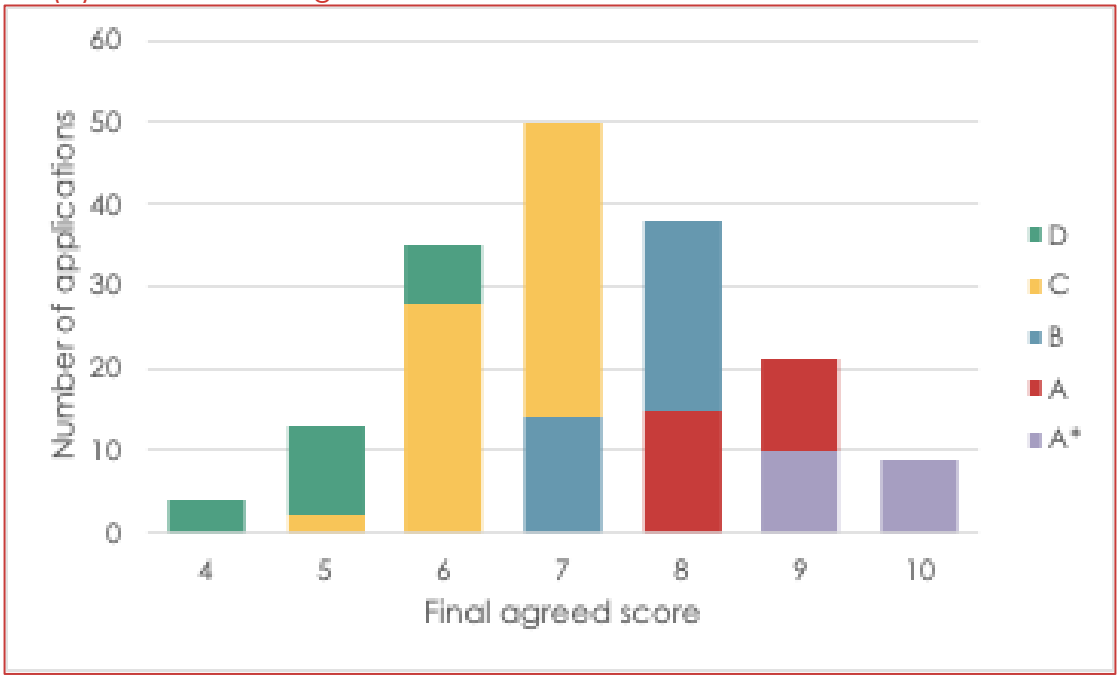
Figure 5 displays the distribution of final agreed scores by band (A* to D), where Band A* represents the highest quality applications and Band D the lowest. Round 1 outline stage data is not shown because the banding was based on two different scores. The figure shows that scores span the full scale from 2 to 10 in the outline stage, though most applications cluster between scores of 6 and 8. Full stage application scores ranged from 4 to 10.

No applications were left unbanded in Round 1, but 27.7% (n = 148) were left 'unbanded' at the Round 2 outline stage. This is due to the change in process between Round 1 and Round 2, with the introduction of a pre-panel sift where outline applications with consistently low scores from the 3 IMs were not discussed at panel.

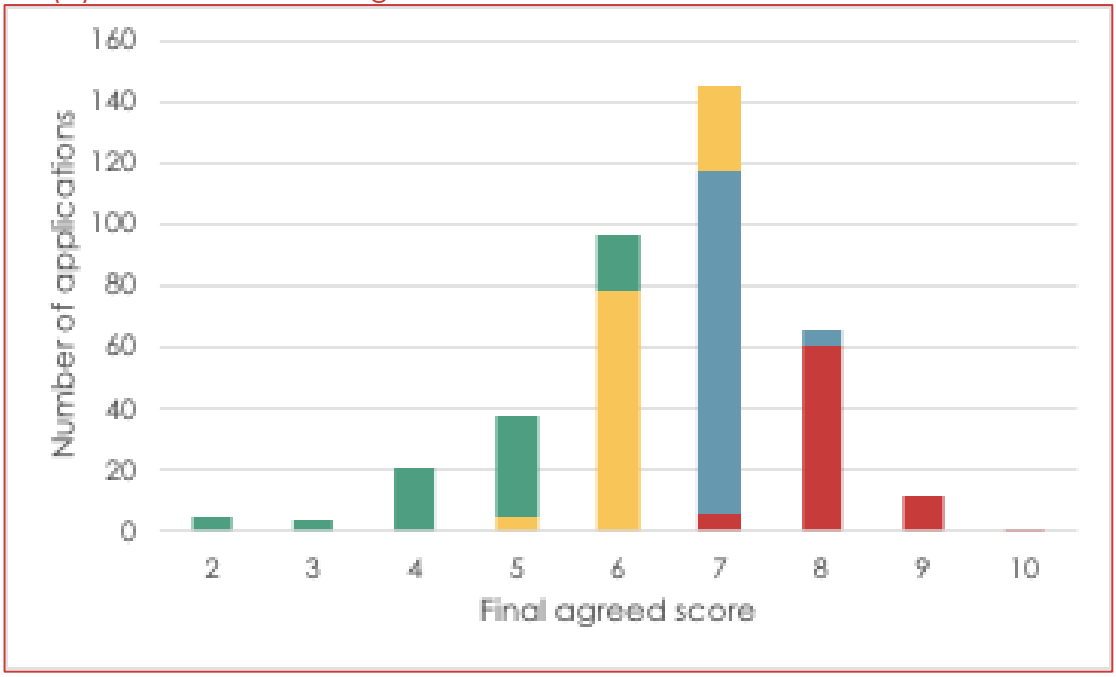
[16] Langfeldt, L. (2001) 'The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome', Social Studies of Science, 31: 820–41.

Figure 5 Distribution of final agreed scores by band

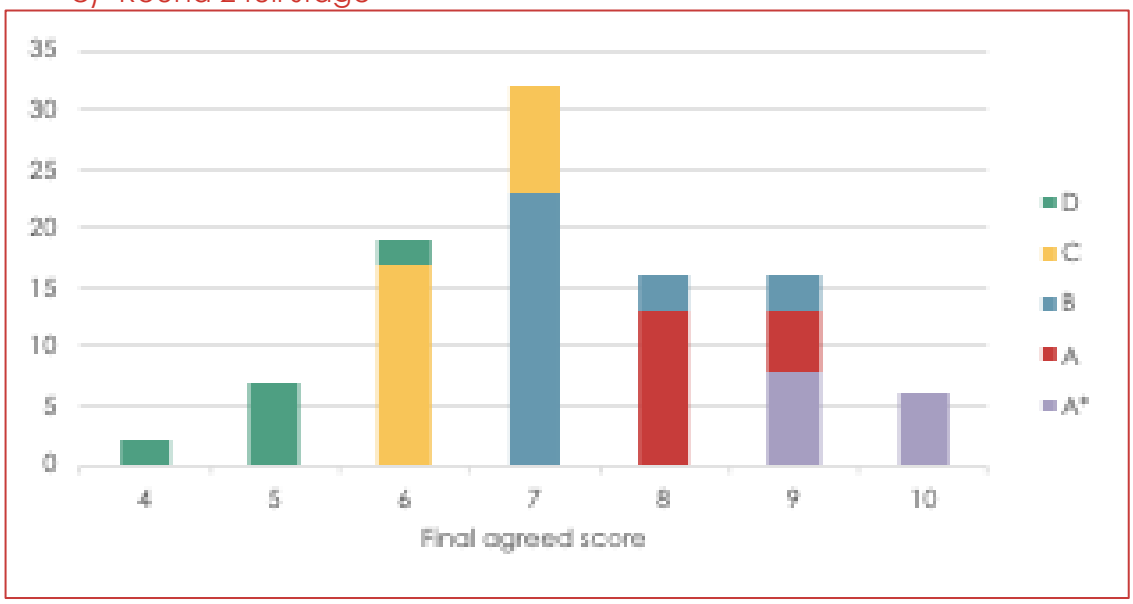
(a) Round 1 full stage



(b) Round 2 outline stage



c) Round 2 full stage



Source: Technopolis analysis of CRCRM Round 1 and 2 application data

Band B and C dominate the distribution (55-60% of the banded applications) in the full application stages, particularly around scores of 6 or 7, suggesting that a substantial number of applications were considered good or very good but not excellent, outstanding or exceptional. Band A, representing the highest-rated applications, appears more selectively applied as it should be, typically in the score range of 8 and 9. Very few applications achieved a score of 10 (exceptional) usually at the full stage assessment.

Band D shows a wider spread, appearing across the lower score ranges usually 6M or below. 27.7% of applications were left unbanded in the Round 2 outline stage with 15% assigned to Band D. Since no applications were left unbanded in Round 1 outline stage, about half of the Round 1 outline applications were in Band D. Feedback from panels to UKRI suggests that many of these applications were considered to be multidisciplinary where tools such as AI were being applied to a different discipline and integration and reciprocity were missing. As expected only a small proportion of full applications fall into Band D after the initial sifting at the outline stage, highlighting the value of having a two-stage application process.

Table 4 Proportion of applications per band

Round	Stage	n	Band A*	Band A	Band B	Band C	Band D	Unbanded
1	Outline	899	-	14.6%	13.7%	20.9%	50.8%	-
1	Full	170	11.2%	15.3%	21.8%	38.8%	12.9%	-
2	Outline	535	-	15.0%	21.7%	20.7%	15.1%	27.7%
2	Full	98	14.3%	18.4%	29.6%	26.5%	11.2%	-

Source: Technopolis analysis of CRCRM Round 1 and 2 application data. Note: Unbanded Round 2 outline applications (n = 148) were not included in this analysis.

Overall, the analysis of panel scores and bands suggests that the CRCRM application criteria and processes have allowed high quality applications to be identified and funded. The banding process enabled UKRI to tension the funding across the panels. Different panels or disciplines can be harsher or more lenient scorers, and the banding process allowed the best proposals per interface to be funded regardless of any disparity in how the scoring scale was applied by different interfaces.

2.2.2 Communications and guidance

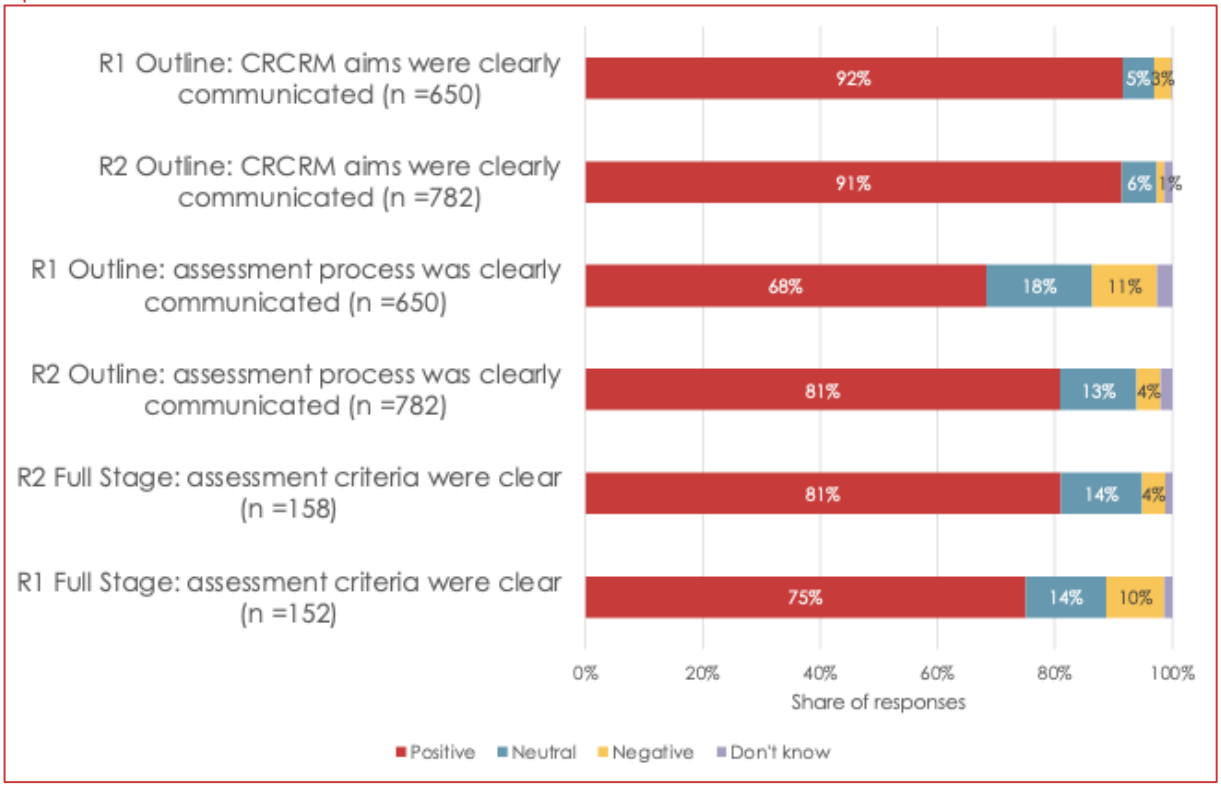
Summary findings

- Stakeholders were broadly positive about communications.
- Webinars were highly valued across both rounds, with notable improvements in the content and format of webinars for Round 2 following Round 1 feedback.
- Guidance improved significantly between rounds. Round 1 applicants were more likely to report problems (for example overlapping sections, unclear IP requirements).
- Understanding and demonstrating IDR remained a persistent challenge for some applicants across both rounds.

2.2.2.1 Communications

Stakeholders interviewed and surveyed across the four stages of the CRCRM programme generally commented positively on the communication and engagement activities. Across Round 1 and 2, in the outline, stage stakeholders felt the CRCRM aims were clearly communicated (see Figure 6). Survey data also shows that stakeholder's view on the assessment process improved between Round 1 and Round 2.

Figure 6 Stakeholder views on communication about CRCRM aims, assessment processes and assessment criteria, Round 1 and Round 2



Source: Technopolis analysis of survey data. Percentages show respondents' views on the statements above, grouped as Positive (agree/strongly agree), Neutral (neither agree nor disagree), and Negative (disagree/strongly disagree).

The combination of the online funding call and direct engagement with universities was effective in raising awareness and generating interest.

> "Overall, we're really positive about the information and the support and the webinars that were given, which were extremely useful." - Applicant

Some applicants qualified this view, highlighting practical issues: In interviews after Round 2 outline stage, a few applicants reported that searches for information on the scheme on the UKRI website often directed applicants to legacy content about Round 1 and not to updated information on Round 2. Several applicants interviewed at the Round 2 full stage also reported that they had missed the announcement of Round 1 and nearly missed the notification for Round 2 in the long list of regular UKRI announcements. One suggested that something additional could be done to highlight unusual opportunities such as CRCRM in the future.

Evidence from surveys suggest that webinars were a valued resource to support the application process (see Technical Annex A, B, C and D). Engagement with webinars was high: 79% of respondents in the Round 1

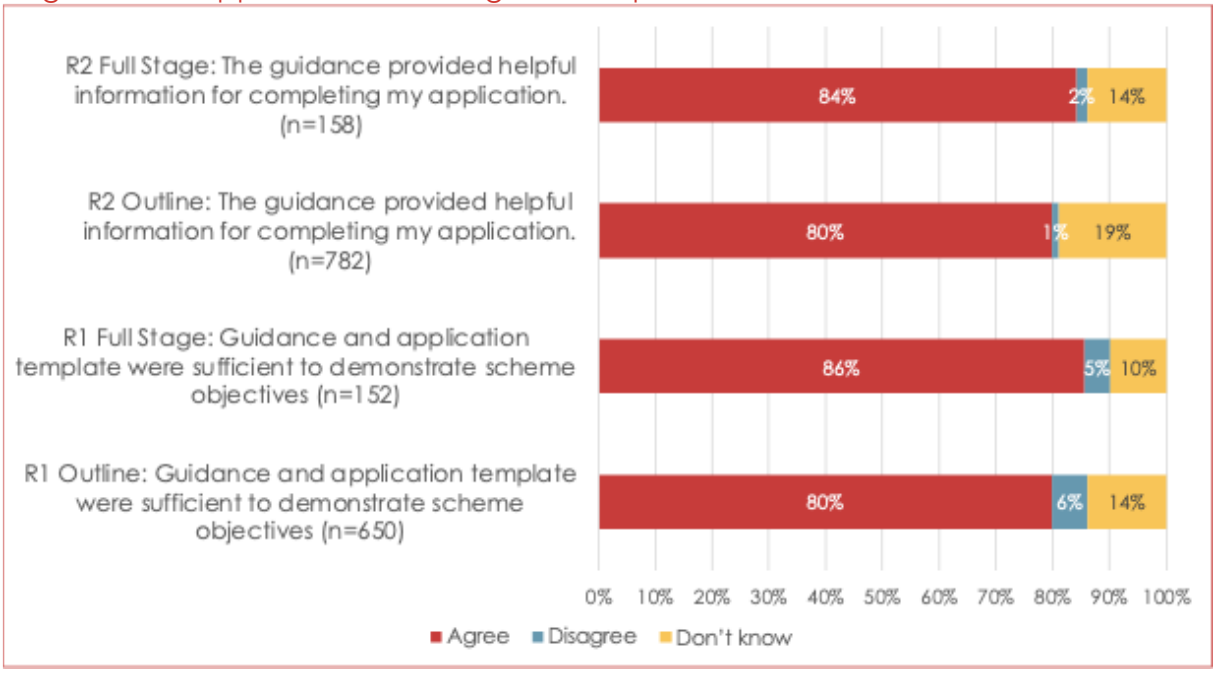
outline stage and 59% in Round 2 outline stage reporting attendance to a webinar or watching the webinars' recording. In Round 2 full stage, 79% attended a webinar or watched the recording. Among those who engaged with the webinars, perceptions were consistently positive, with 73% of respondents at the Round 1 outline stage, 92% at the Round 2 outline stage, and 95% at the Round 2 full stage reporting that they found them useful. These findings indicate that webinars played an important role in helping applicants understand scheme requirements and navigate the application process.

The views on webinars for applicants differed somewhat between cohorts. Many applicants, especially those interviewed in Round 2, found the webinars valuable in supporting understanding of IDR and in developing outline applications. Some applicants from round 1 were less enthusiastic and found that they replicated the written documentation. Survey responses confirm this view, with several applicants suggesting that written guidance was preferred over webinars. This likely reflects improvements made to the content of the webinars in Round 2 as the understanding and learning within the IRM team grew from feedback received during Round 1. Interviewees from the IRM team also noted that the information in the Round 2 webinar was improved to support researchers in understanding the different facilities available from UKRI councils. One applicant suggested that the webinar for applicants progressing to the full application stage could be shown to applicants at the outline stage, as it contained useful information to improve applications and would reportedly have been as part of a further round.

2.2.2.2 Guidance

Survey data suggest that applicants were mostly positive about the scheme's guidance and application template. As shown in Figure 7, 80% of Round 1 outline stage applicants and 86% of Round 1 full stage applicants replying to the survey agreed that the guidance and template were sufficient to demonstrate how their proposals would meet the scheme's objectives. Similarly, most Round 2 applicants found the guidance was helpful for completing their full application on The Funding Service (TFS), with 79% expressing this view in the outline stage and 84% in the full stage.

Figure 7 Applicants' view on guidance provided, Round 1 and Round 2



Source: Technopolis analysis of survey data.

In open-text survey answers across both rounds, respondents suggested that the inclusion of examples of successful projects and clearer instructions about the definition of IDR in guidance would be useful.

Evidence from interviews confirm survey findings. Applicants interviewed for the evaluation were broadly positive about the application guidance, although there were somewhat differing views. Notably, applicants to Round 1 were more likely to bring up issues than applicants from Round 2, suggesting an improvement in guidance over the two rounds.

Some applicants for Round 1 indicated that the guidance was clear, but other individual applicants identified specific challenges with regard to the application requirements for submission. This included redundancies or overlaps between sections in the outline and full-stage application templates, lack of clarity around management of intellectual property (IP), use of subcontractors and other issues. Few Round 2 applicants interviewed for the evaluation had strong views on the guidance, with some describing it as "improved" or "standard". We note that the guidance documents for applicants and IAC members were reviewed and made more user-friendly in Round 2, and practical examples from Round 1 were introduced in guidance and training materials.

Survey findings indicate that the 'Fit to Opportunity' section was received positively by most Round 2 full stage applicants, with 77% agreeing or strongly agreeing that it was helpful, while only 6% disagreed.

> "The interdisciplinary definitions were very useful and impacted the proposal's development. It helped structure the conversations between partners." - Applicant

However, interview evidence suggests that some applicants found aspects of this section challenging. The definition of IDR was the most prominent issue, brought up by applicants across all three cohorts. They suggest that interpreting the definition of IDR is inherently challenging for applicants and not an issue that can be solved easily through guidance. Relatedly, two applicants noted that the requirement to explicitly describe the "fit to scheme" was unusual compared to other schemes and one specifically highlighted challenges in completing this section:

> "I found it very challenging to write about the vision and the fit to opportunity, what the difference was, and how you're supposed to be highlighting the fit to opportunity. That was really challenging. […] Not seeing an example of a successful application made it hard to know what say." – Applicant

2.2.3 Applicants' experience of CRCRM processes

Summary findings

- Applicants were strongly motivated by the scheme's explicit IDR focus, its broader scope compared to other funding opportunities, and the legitimacy it conferred on interdisciplinary work.
- The two-stage process was widely appreciated for reducing burden, though some found the outline stage still time-consuming and Round 2's in-text referencing requirement reduced the usable word count.
- Outline stage feedback was seen as too generic to be actionable, though understood as a practical necessity given application volumes.
- Full-stage clarification questions for applicants to respond to were seen as genuinely useful by both applicants and assessors.

2.2.3.1 Motivations to apply

Applicants interviewed for the evaluation described a diverse range of motivations for applying to the CRCRM scheme.

The explicit emphasis on interdisciplinarity was attractive and encouraging to many applicants who were looking to integrate different disciplines to solve complex problems. Some also saw the scheme as stamp of approval and legitimisation of their own long-standing commitments to interdisciplinarity vis-à-vis their home institutions.

"Whilst a lot of effort and sheer hard work have gone into putting this project together, it has been a joy learning with and collaborating with new disciplines, without which this project would not be possible." - Applicant

The scheme provided a broader scope for more imaginative ideas. Several applicants stated that their projects struggled to meet the specific requirements of other funding bodies, such as philanthropic organisations or European funders. Many of these other funding opportunities require a narrow IDR focus on specific combinations of disciplines or alignment with specific priorities, whereas the CRCRM pilot's broader scope and interdisciplinary ethos enabled researchers to pursue more speculative and imaginative ideas, encouraging risk-taking and creativity. The open scope also meant that some applicants found it difficult to know what was expected.

Several interviewees suggested that the scheme had enabled them to pursue long-standing ideas that they had been unable previously to find means of support from other funders. In some cases, these were ideas that had been rejected by other funders due to poor fit with the funding criteria. Other applicants had developed ideas for research, often more unusual ideas different from what they would otherwise consider, and had been waiting for the right opportunity to fund it. For them the scheme provided a good "match" for them to bring their ideas forward, and the outline application offered a low-commitment entry point to try out unusual ideas.

"Freedom to authentically describe the idea and way of working that was needed, rather than trying to have to force-align things with a dominant council." – Applicant

Applicants also appreciated the effort to train a dedicated college of assessors. This gave researchers confidence that their applications would be evaluated fairly and with an understanding of the nuances of interdisciplinarity, which other funding opportunities had previously failed to do.

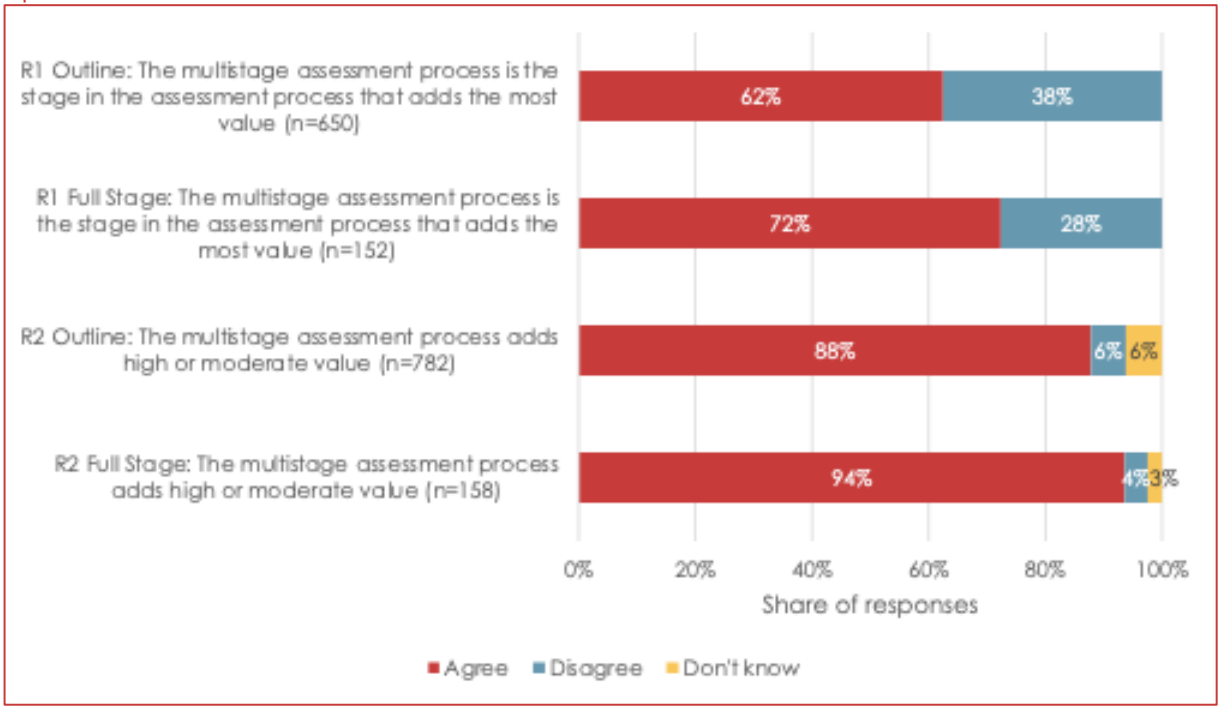
"The reviews we got on our full application were very supportive and were clearly written by people who understand the interdisciplinary ethos. It was heartening to read reviews by people who understood our bid and who were looking for reasons to support it rather than reasons to reject it" - Applicant

2.2.3.2 Two stage application process

Researchers were broadly supportive of the two-stage application, emphasising its value in reducing the overall burden of application writing. As shown in Figure 8, respondents across both rounds found the multistage

assessment process valuable. However, respondents in Round 2 were more positive than those in Round 1.

Figure 8 Stakeholder views on the extent to which the multistage assessment process adds value, Round 1 and Round 2



Source: Technopolis analysis of survey data. For the Round 1 analysis, respondents either selected the multistage assessment process as adding the most value (shown here as Agree), or not (Disagree). For Round 2, respondents stating the multistage assessment process had high or moderate value were aggregated into 'Agree', and for no or low value this was aggregated into 'Disagree'. Respondents that marked 'Not sure' have been represented as 'Don't know'.

Applicants considered the outline application to be more manageable and allowed them to submit risky ideas without a large initial time commitment. Some further saw it as a constructive constraint that forced them to focus on their ideas. While still supportive of the two-stage process, some researchers nevertheless noted that they had spent a significant amount of time to prepare the outline stage applications (in the case of Round 2 there was also the internal institutional selection as part of demand management). This seemed particularly true for applicants developing new ideas and collaborative relationships specifically for the scheme.

For IAC members too, the two-step process was a welcome feature, not least as a necessary mechanism for coping with the high volume of proposals.

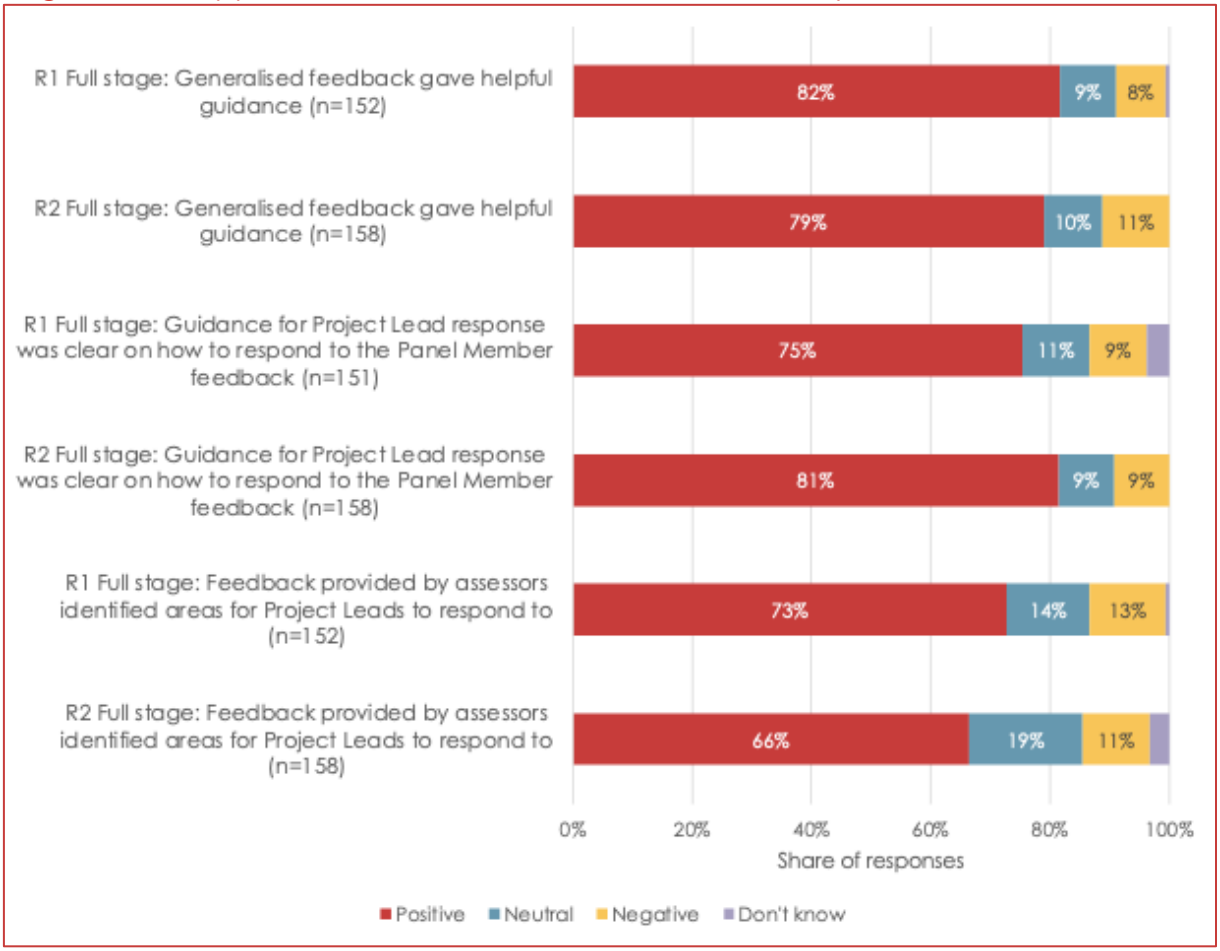
> "An outline stage is essential for this programme due to the incredibly high level of interest and the interdisciplinary nature of the programme requiring a range of panel members."
> – Panel member

Although longer, some interviewed applicants still found the Round 2 applications too short to communicate their ideas effectively. Several applicants interviewed after the Round 2 outline and final stages noted that unlike Round 1, the Round 2 application form required references to be included within the main text box. This reduced the available word count for the application, making it challenging for applicants to present their idea effectively. Some saw this as particularly challenging in the context of CRCRM where many applications included complex ideas with several different disciplinary component parts.

2.2.3.3 Feedback

Evidence from the surveys suggests that most respondents view the feedback process positively across both Round 1 and Round 2 stages. As shown in Figure 9, around three quarters or more respondents agreed that guidance on responding to panel feedback was clear, and that generalised feedback provided helpful guidance. In addition, IAC members responding to the survey also reported positive views regarding the clarification statements provided at Round 2 outline stage, which were intended to provide more specific information alongside the generalised feedback. Nearly 70% expressed a positive view about their value (see Figure 169 in the Technical Annex). Survey responses indicate that IAC members were overwhelmingly positive about the Round 2 Full stage process whereby the three assigned IMs discussed each application to develop compiled feedback and requests for further information for the Project Lead. Over 80% expressed a positive view of this approach (see Figure 103 in the Technical Annex).

Figure 9 Applicants' views on the usefulness and clarity of feedback



Source: Technopolis analysis of survey data. Percentages show respondents' views on the statements above, grouped as Positive (agree/strongly agree), Neutral (neither agree nor disagree), and Negative (disagree/strongly disagree).

In interviews, most applicants agreed that the generalised feedback after the outline stage was helpful if a little generic and lacking in actionable details, which made it difficult to improve subsequent applications. Some applicants described it as "generic", "useless", or "frustrating", and in some cases contradictory. Still, some applicants found it useful and there was a widespread understanding that it was the only feasible option for providing any feedback at all, given the large volume of applications.

At the full application stage, both applicants and IAC members commented positively on the opportunity to give and respond to feedback on full applications prior to the assessment panel meeting. Several applicants reported using the process of answering clarification questions as an opportunity to communicate additional elements that they had to leave out of the initial application.

> "It was really nice because there were a couple of things that we could clarify that maybe hadn't been articulated well. And it was really nice to be able to be asked those questions and have people be interested and then have the opportunity to say back to them, 'oh, no, we meant this'…And we actually explained 'this is why it's interdisciplinary, because we're doing it this way, because we want to be able to analyse it' So, it was a positive thing for me, stressful, very stressful, but positive thing to be able to get the feedback." – Applicant

Feedback after the assessment of the full application stage was primarily remarked upon by unsuccessful applicants, who said they did not necessarily agree with the comments but accepted it, and appreciated the effort to communicate a consensus view from the panel.

2.3 Effectiveness of the IAC approach

2.3.1 IAC recruitment and training

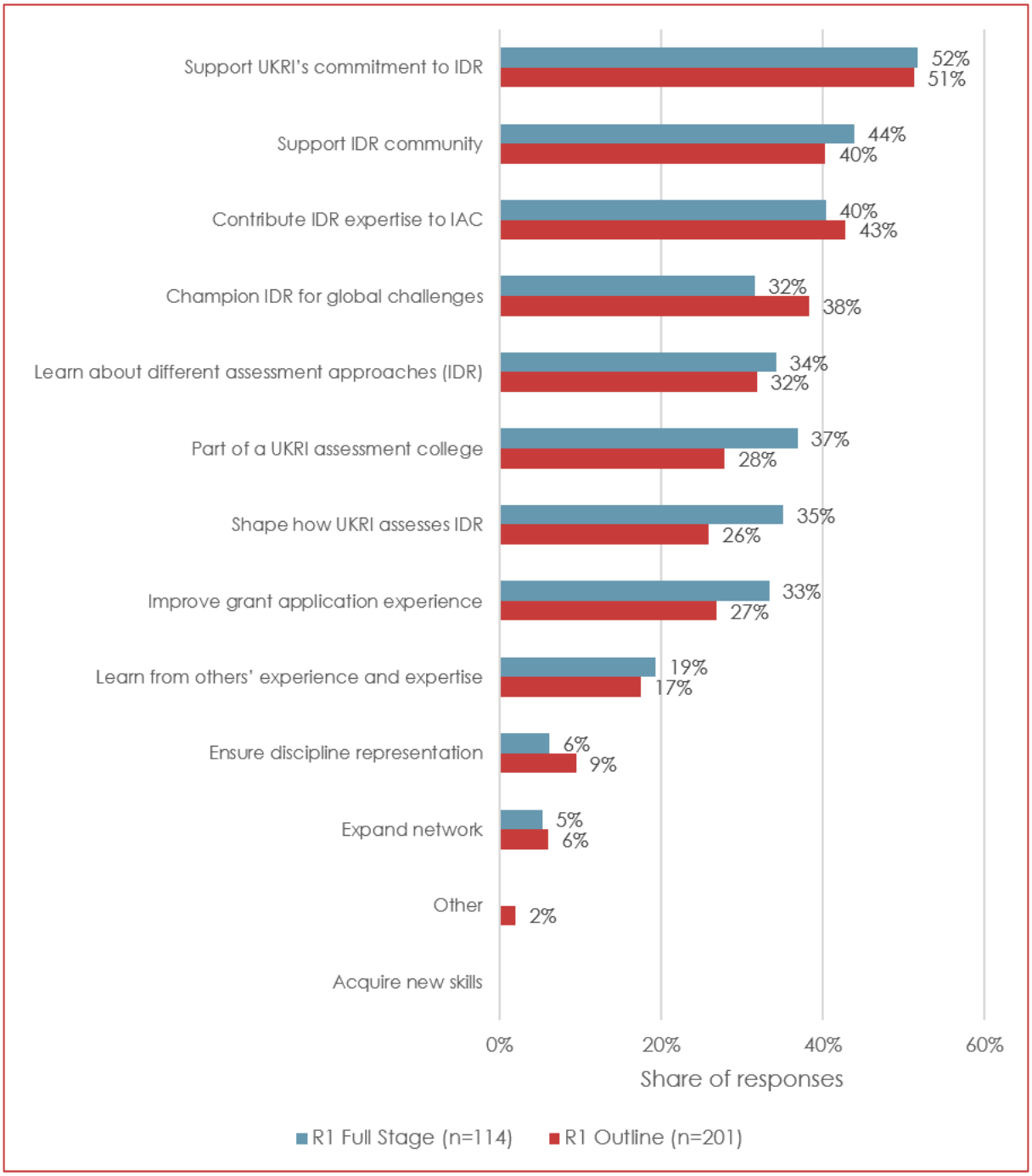
Summary findings

- Interest in IAC membership was high, attracting over 1,300 applicants for ~300 places.
- The college was broadly seen as diverse and well-balanced across UKRI council remits.
- Training was rated highly, particularly in Round 2 where real CRCRM case studies replaced the more abstract Round 1 training albeit some felt training was overly detailed for experienced researchers.
- Suggestions for expanding the IAC included snowballing via existing members and targeted outreach to government, charity, and non-profit professionals.

The evaluation team interviewed 28 members of the IAC during over three rounds, including IAC chairs, deputy chairs, panel members, and rovers (see Appendix B.2). Reflecting the IAC's overall composition, most panel members interviewed were research-active academics, though interviews also included some in research management and other professional roles.

IAC members had personal, professional, and institutional motivations for applying to join the assessment college. As shown in Figure 10 the most common motivation across both the Round 1 outline and full stages was to support UKRI's commitment to IDR (51% and 52% of IAC members respectively). Other popular motivations similarly centred around IDR, such as supporting the IDR community, and contributing IDR expertise to the IAC.

Figure 10 IAC members' motivations for applying to join the assessment college, Round 1.



Source: Technopolis analysis of survey data.

In interviews, other IAC members also expressed a desire to learn about IDR research from others, expand their professional networks, and gain experience with the UKRI funding processes. A few also cited general encouragements from their home institution to get involved in such roles as a factor.

IAC members were mostly positive about their experience of applying for the IAC. Most indicated the process was 'very straightforward', and in line with the applications for other similar roles. Several interviewees commented on the role of narrative CV, with some highlighting it as an effective tool to describe their IDR experience, whilst others suggested it was closer to business as usual. One interviewee expressed reservations about the usefulness of narrative CVs in the context of AI, as narrative CVs tailored to the opportunity can be generated automatically with minimal personal input from the applicant.

Two interviewees emphasised the inclusive nature of the process, with many colleagues from non-academic backgrounds choosing to apply (for example, research managers, representatives from industry) and panels displaying a good diversity of expertise. Yet, both interviewees and survey respondents suggested that UKRI could have taken further steps to encourage more engagement from non-academic applicants to increase the diversity of expertise in panels, for example through more targeted recruitment initiatives. It is worth noting that the UKRI team did share the details of the IAC membership call with the UKRI Research Organisation Consultation Group, ARMA, Praxis Auril, Association of Independent Research and Technology Organisations, and letters to senior university representatives.

Across the surveys conducted during both rounds and both stages, the vast majority of IAC respondents agreed or strongly that:

- Overall, the training and guidance materials provided were sufficient in preparing them for their role in the CRCRM scheme panels
- The training materials were of high quality and informative

These points were exemplified in the Round 2 full stage survey results as shown in Figure 11 and Figure 12. In the Round 1 surveys, only a small minority of IAC respondents disagreed (2% to 9%) or neither disagreed or disagreed (3% to 11%), whereas virtually all respondents agreed or strongly agreed in the Round 2 surveys.

Figure 11 Overall, the training and guidance materials provided were sufficient in preparing you for your role in the CRCRM scheme full stage panels (n = 63)

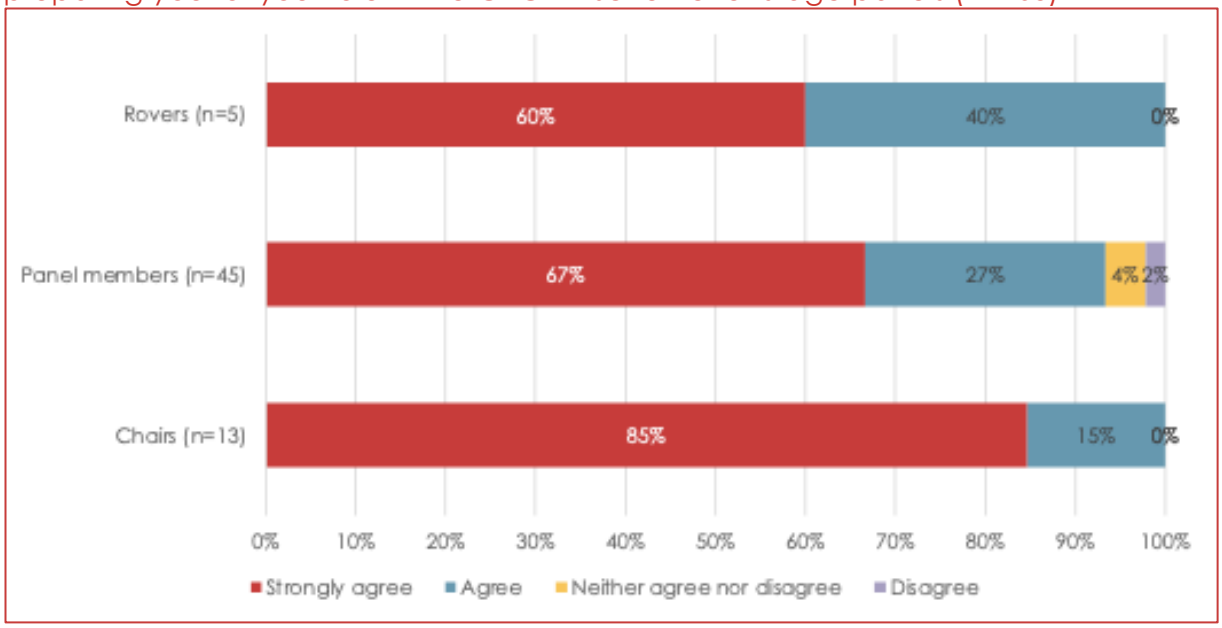
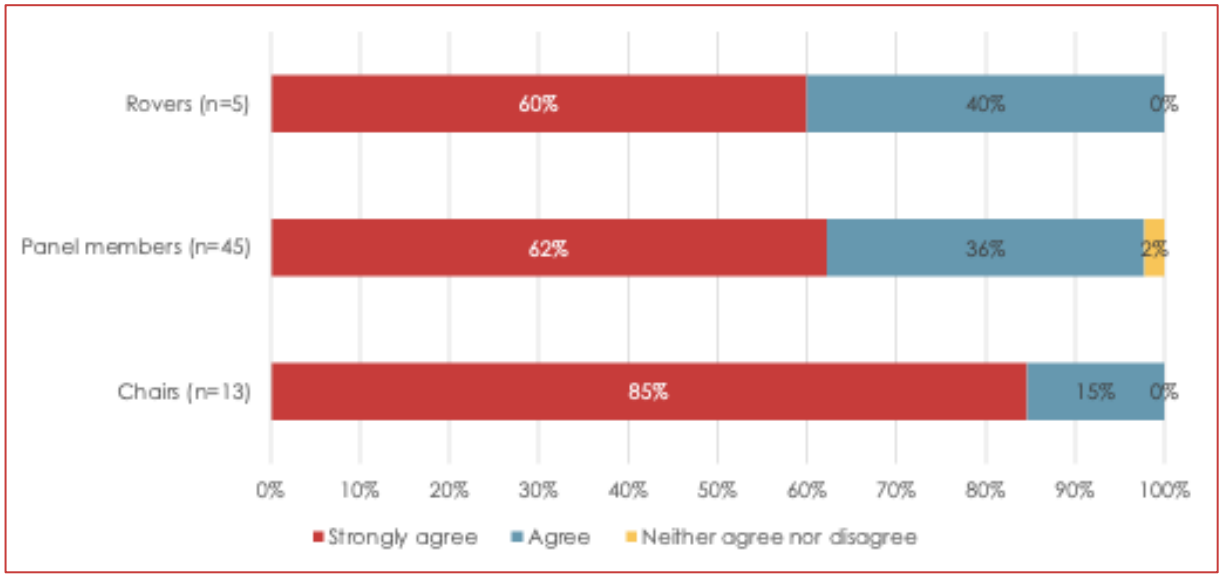


Figure 12 The training materials for the full stage were of high quality and informative



Most of the IAC interviewees praised the high quality of the training and reported a positive experience. Some valued the training as an opportunity to gain additional perspective and reflect on IDR in a way that is not usually possible. A few interviewees identified the quality and content of the training as stand-out aspects for them, stating that the training was clear and helpful.

> "Being a part of the IAC did not just provide the most rigorous, reflective and effective training on assessing funding applications I've ever received, it also provoked discussion,

> debate and a huge gain in everyone's understanding around the nature, demands and challenges of interdisciplinarity." – IAC member

Qualifying this view, some felt that the training was a bit too detailed, and/or that it had been less relevant for more experienced researchers. Several IAC members saw improvements in the training for Round 2, especially the use of examples and experience from Round 1. This was further evidenced in IAC surveys: In the Round 1, respondents requested more interaction to improve the training and by Round 2, a majority of respondents indicated that the training required no improvements. The breakout rooms and Padlet outputs aimed at sharing good practice and approaches for feedback to applicants as well as preparing oral reports at panel meetings were valued as useful to at least some extent by all Rovers, 92% of chairs and 85% of panel members in the Round 2 full application stage survey (n=63). UKRI presentations on the assessment process and trainings and presentations from Samantha Aspinall and Catherine Lyall on assessing IDR were widely considered useful: for example, 53% of IAC members found the training session facilitated by Lyall and Aspinall to be very useful in the Round 1 full stage.

Round 1 training was seen as more abstract/theoretical in comparison. For chairs and rovers, one interviewee highlighted regular meetings as a valuable part of ongoing training:

> "Not only a separate briefing for rovers before the panels, but regular meetings with the rovers before during & after panel assessments, in addition to the training that they gave to the reviewers. This was a good, iterative way of receiving and providing feedback in real time during the panel assessments." – IAC Chair

To expand the IAC, IAC members suggested adopting a snowballing approach via existing members promoting the IAC among their networks and institutions. Another suggestion was reaching out to individuals with research experience in government agencies, charities, and non-profits. IAC members interviewed noted that early career researchers may particularly benefit from participating actively in the College, gaining practical experience with funding processes and IDR at a formative stage in their career. In Round 2, UKRI did set up an opportunity for UKRI Future Leader Fellows to observe outline and full stage panels, supporting 54 fellows through this process. Promoting the IAC as an opportunity to help shape UKRI's approach to supporting IDR could appeal to those interested in influencing research policy and strategy.

2.3.2 IAC experience of CRCRM assessment processes

Summary findings

- Panel time allocations were seen as tight but adequate.
- The three-introducer (IM) structure worked well and was seen as equitable.
- Chairs were generally effective at building consensus, and the process was described as "robust but collegiate."
- In some cases, a panel member with specific topic expertise could take over discussions, though this was not systemic.

Most IAC members consulted for the evaluation agreed that the time allocation for panel discussions was appropriate. In the surveys, the majority of IAC members agreed or strongly agreed that there was enough time to build consensus in panel meetings, including 85% in the Round 1 outline stage survey, 88% in the Round 1 full stage, and 98% in the Round 2 full stage. This figure was notably lower (58%) for the Round 2 outline stage (see Figure 168 in the Technical Annex). According to interviewees, the time available was tight at times but ultimately allowed the right amount of detail to be discussed while keeping focus on the essential aspects of the applications. Qualifying this view, a few interviewees felt that time constraints hindered consensus-building in cases with divergent views (Round 2 outline stage), and one also observed that some panels were busier than others due to the number of applications.

For the full applications, the ability of IMs to communicate about the applications beforehand in discussion boards to align on clarification questions for the project lead meant that panel discussions often focused on the responses to the clarification questions according to evaluation team members who observed panels. When asked what aspect added the most value to the fair assessment of IDR in the Round 1 full stage assessment process, panel members, chairs, and rovers considered this to provide equal or greater value than the panel meeting process (see Figure 91 in Technical Annex). This positive sentiment was also reflected in the Round 2 full stage survey, with 80% of rovers and 93% of chairs agreeing or strongly agreeing that the discussion boards facilitated a fairer assessment during panel meetings.

> "*(Confidence in assessing applications)* wasn't any issue, really, because it was clear which of the other 2 IMs on the team had more disciplinary or methodological expertise. So, my confidence was built through the back and forth of the discussion. I was able to ask naive questions and receive considered, expert responses." - Panel member

Most IAC interviewees agreed that all voices were heard, with different perspectives offered by panel members contributing positively to the discussion. The process with three presenters (IMs) worked well and in interviews, several IAC members thought the inclusion of additional readers in the Round 2 full stage assessment panels helped broaden the discussion and in some cases act as substitutes for IMs who were unable to attend. This was also the observation of IRM team members.

However, it was also noted that the presence of a panel member with specific expertise with aspects of the application could sometimes discourage others from voicing their opinions, as also observed by a Technopolis evaluation team member.

Debates or tensions in the panel assessments were mostly resolved through the intervention of IAC Chairs or, in some cases, UKRI guidance on consensus building. Most interviewees agreed, across both IAC and UKRI interviewees, that IAC Chairs had the right skills to facilitate discussions towards consensus building. The process was described as "robust" but "collegiate" overall.

> "We had an excellent panel in which people were keen to reflect upon their approach to scoring and keen to have their assumptions tested. There was no one who saw it as their role to 'defend' a particular project based on a particular approach to interdisciplinarity. People collectively wanted to come to the right decision." – IAC Chair

A couple of IAC members suggested that more could be done to instigate wider participation from non-academic panel members who could tend to be quieter during panel discussions.

2.3.3 Characteristics of pool of assessors

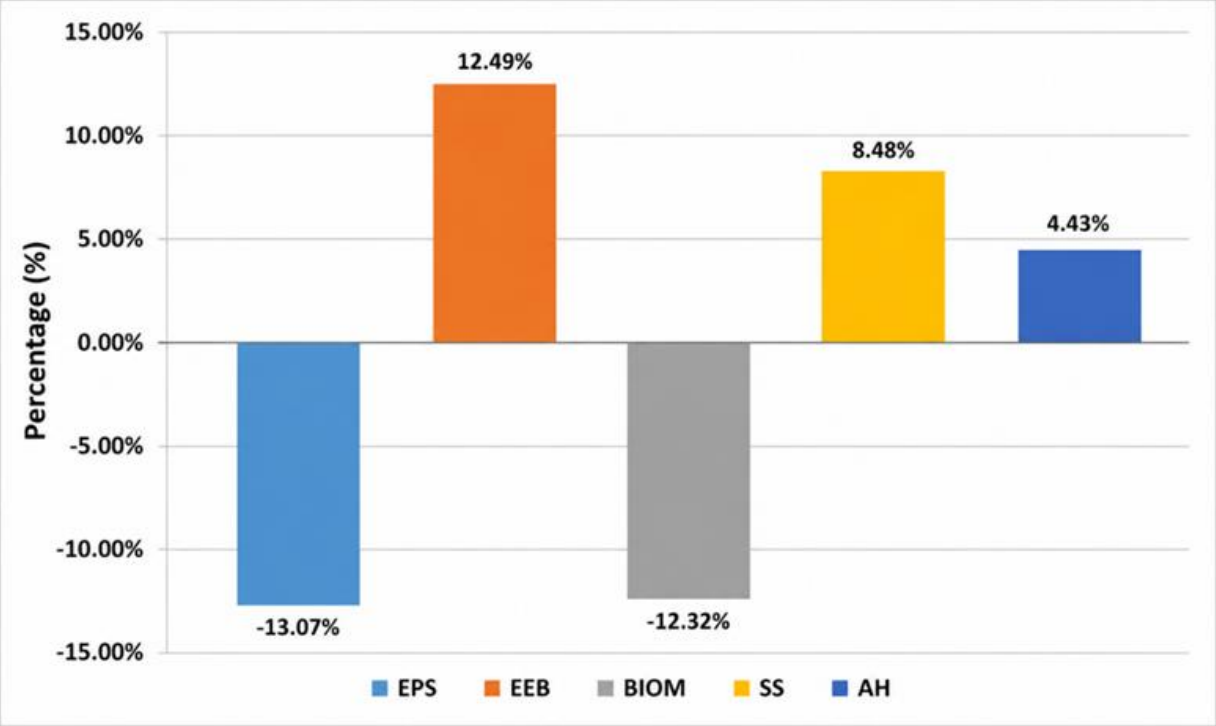
Summary findings

- The IAC membership expertise was well-fitted with the remits of applications received, based on the data from ~65% of assessors who consented to share their assessment data.

IRC analysed the IAC membership comparing the remits of applications received relative to IAC membership expertise, however, the full dataset of IAC membership was limited to assessors that consented to share their information with the evaluation teams. For Round 1, data was available for 67% of IAC members, and for Round 2 for 63%. Figure 13 shows the differences between the percentages of each remit in the populations of assessors and applications. For example, the percentage of assessors having an EPS expertise to all assessors was 13.07 points less compared to the percentage of applications categorized as EPS to total applications. For the AH remit, the

percentage of AH assessors was higher than the percentage of AH applications by 4.43 points. The analysis showed that, while the breadth of expertise was appropriate, some areas did not have the same proportion of assessors in the IAC relative to the submitted applications. Note that this should not be viewed as a deficiency. IRC analysis shows that all applications were assessed by at least one IAC member with a shared remit based on the data analysed, but it does point to an opportunity to recruit more IAC members in underrepresented topics and areas such as BIOM and EPS to ease assessor workloads in those areas. Given that application proportions across the interfaces were relatively stable across rounds, this finding does offer some potential insights to future interdisciplinary funding schemes.

Figure 13 Differences between applications' and assessors' remits across rounds



Source: IRC

Overall, panel members reported that many applications fell within or close to their area of expertise, with 47% of those responding to the survey indicating that more than half of applications were within their expertise and a further 25% reporting that around half were (see Figure 34 in the Technical Annex). Despite this, respondents were generally confident in assessing applications beyond their immediate expertise. Most panel members reported feeling confident, with 64% either fairly or completely confident, and a further 27% somewhat confident (see Figure 35 in the Technical Annex). Only a small

minority (9%) reported low confidence in assessing applications outside their immediate expertise.

The IAC was seen to have a suitable breadth of expertise to support fair assessment, with 68% of chairs and 92% of rovers agreeing or strongly agreeing (see Figure 37 in the Technical Annex). Similarly, most chairs (81%) and rovers (83%) agreed or strongly agreed that applications were appropriately allocated to panels (see Figure 38 in the Technical Annex).

2.3.4 Efficiency and resources

Summary findings

- Round 1 delivery was pressured by under-resourcing and a team structure not designed for simultaneous delivery and evaluation. Round 2 was considerably smoother.
- Demand management (institutional pre-sifting, introduced in Round 2) was broadly welcomed for reducing burden on UKRI and reviewers, though some worried it might disadvantage less well-established researchers or lead institutions to favour safer applications.

2.3.5 Overview

Interviews with UKRI IRM staff considered the issue of efficiency in the delivery of the scheme.

The scheme processes were developed iteratively during Round 1 and adapted as issues arose. This was constructive but also caused capacity challenges related to the need to devote time to addressing unforeseen issues that had arisen. For Round 2 there was more time to prepare, and UKRI processes were harmonised and designed better compared with Round 1. Most internal processes were ready prior to Round 2 being launched.

UKRI staff interviewed reported that the IRM programme team was somewhat under-resourced for Round 1, leading to pressure on staff capacity; more staff resources at a strategic level would have ensured a smoother development of the Round 1 processes without overburdening the team. This issue was particularly prominent in the first six months, as the full team had not been recruited, and was exacerbated by the fact that the team structure was based on programme delivery and so lacked the resource to develop and evaluate the programme at the same time. These pressures notwithstanding, one interviewee also noted that staff were well supported and benefited from an inclusive culture in the team during Round 1.

The implementation of demand management, experience of delivering Round 1 and more streamlined processes enabled greater efficiency in

Round 2. The IRM team members also reported that planning and resourcing were improved for Round 2, with more time allocated for delivering panel meetings over four days rather than two. This was reflected in the surveys with one response praising the "support and preparation of all the UKRI staff," and describing it as "absolutely invaluable."

2.3.6 Institutional support

Most of the applicants reported receiving some form of institutional support. 61% of Round 1 applicants surveyed agreed that there was support in their host organisation to develop their interdisciplinary research ideas.

The nature and extent of the support reportedly varied significantly between institutions, from standard grant preparation support to more tailored support, especially in the context of internal selection related to demand management measures in Round 2.

Some interviewees experienced increased institutional buy-in and several applicants described how internal selection triggered additional support from their home institution to prepare the application to UKRI, for example in the form of advice and offers of internal peer review of applications, as well as more technical support.

2.3.7 Demand management

Demand management, whereby institutional caps were put in place and research institutions conducted initial sifting of applications before submission to the CRCRM scheme, was introduced in Round 2 because of the high volume of applications (n = 979) received in Round 1. Most of applicants and IAC members interviewed saw this as a practical and necessary step which helped to reduce burden on UKRI and reviewers as well as support efficient use of institutional resources. Based on the recommendation in the Careers Research and Advisory Centre (CRAC) report to establish good practices to enhance equality, diversity and inclusion (EDI), institutions submitting applications were asked to provide data on their submitted internal applications and an Inclusive Selection Statement, outlining their internal sift process. Feedback on this process from the institution research management teams was that collection of sensitive EDI data, which was often held by HR rather than research management teams, added a significant burden to the research organisations. Collecting data on external co-leads, where multiple organisations were collaborating on an application, was particularly difficult and many were not able to do so. To support the research management offices, the IRM team co-developed a workshop with the Association of Research Managers and Administrators (ARMA) in October 2024 with central

UKRI support, and presented the rationale for collecting the data and provided a forum for questions and feedback.

The internal sift process gave IAC members and successful applicants from Round 1 an opportunity to provide advice to applicants at their own institutions.

> "For a non-research-intensive organisation like my own, an internal sift process is very useful as it is not just a decision-making process but a way of helping to focus development. For time poor academics, it means that they engage with the process earlier and so their ideas can benefit from increased support from a range of experts. For this call, this meant being able to link applicants up with those who had experience from Round 1 to work with them on developing their interdisciplinarity further". - Rover

Of the 146 IAC members who responded to the Round 2 outline survey (see Figure 160 in the Technical Annex), 48 (33%) reported being involved in an institutional sift process, 54 (37%) that their research organisation had run an internal sift but that they were not involved, 12 (8%) that their organisation had not run an internal sift, and 32 (22%) were unsure whether an internal sift had taken place. For those involved in the sift process, most were positive about the process (74%).

> "The sift process at my university was an opportunity to ensure that applicants understood the core principles and criteria of CRCRM funding. Those principles and criteria include identifying and clearly demonstrating how reciprocal benefit to all participating disciplines will be achieved."
> – Panel member

The introduction of demand management also had an effect on the nature of the application process. Several applicants described formal internal selection processes set up by the institutions. Interviewees offered different views on this change: some were concerned that institutions were insufficiently able to judge IDR and would therefore tend to select fewer risky and exciting applications to go forward or that it would disadvantage less well-established researchers. UKRI analysis of data from the institutional selection statements and data on interested vs submitted applicants suggests that demand management did not have a major impact on the portfolio of applications or introduced any biases in relation to applicants' protected characteristics.

The review of Research Organisation Inclusive Selection Statements found variable practices. Stronger submissions demonstrated clear efforts to ensure

openness and transparency through wide, neutral-language promotion across established networks, alongside proactive engagement activities such as webinars, workshops, and tailored 1:1 support to help applicants shape interdisciplinary ideas. Good practice also included consistent use of CRCRM assessment criteria, structured peer review processes involving interdisciplinary expertise, and deliberate application of EDI principles (e.g. panel briefing, diverse membership, and supportive applicant feedback). Notably, several research organisations maintained proportionate and developmental processes even when under subscription caps, ensuring fairness and quality assurance. Weaker statements highlighted gaps where institutions relied on passive dissemination, did not actively promote opportunities, or bypassed formal selection and review processes when applicant numbers were low. In some cases, EDI considerations and transparency of decision-making were insufficiently evidenced and processes did use assessors with IDR expertise or assess applications against the specific IDR assessment criteria for the call.

2.4 Successes and challenges

Summary findings

Successes

- The IAC created a dedicated community with genuine expertise in IDR assessment, and it is widely seen as one of the scheme's most significant achievements.
- Assessment processes were judged to be fair and increasingly effective, with Round 2 improvements including clearer criteria, better training, and smoother panel operations.
- The applicant response stage at full application assessment combining feedback and questions from three assessors was an innovative and well-regarded feature that helped panels make more informed decisions.
- Partial randomisation was broadly accepted as a practical necessity, though some applicants found it instinctively unfair.

Remaining Challenges

- Articulating reciprocal benefits of IDR was the hardest dimension for applicants and the weakest element in assessed applications across both rounds.
- "Pure" AHSS combinations were underrepresented in funded projects (1 vs 21 pure STEM projects), likely due to the larger cognitive distance between disciplines within single AHSS councils compared to STEM councils,

potentially creating difficulties for AHSS researchers to articulate interdisciplinarity.

- Some scoring challenges remained, particularly in differentiating between "excellent" (score 8) and "outstanding" (score 9) applications.
- Assessors with both STEM and AHSS backgrounds tended to give lower scores, suggesting reviewers with a greater of breadth of IDR experience apply a higher bar (a useful insight for future scheme design).

Two key challenges are seen to negatively affect funding for IDR:

1. IDR applications do not always fit the remit of one specific funder and/or its departments which are usually organised on a disciplinary basis
2. It can be difficult to find peer reviewers with appropriate expertise and knowledge of IDR to fairly assess IDR applications and monodisciplinary perspectives of reviewers adversely affect the funding of IDR

The CRCRM pilot addressed the first challenge by specifically requiring applications to span the remits of at least two UKRI councils but allowed applications not to be constrained by any remit boundaries. The IAC approach effectively addressed the second challenge as described previously.

> *"This call gave us an opportunity to develop ideas* beyond what would ever be funded by single research councils". – Applicant

2.4.1.1 Assessment processes were more efficient and effective in Round 2

Most interviewed IAC members and applicants found the assessment criteria clear, although often with some tension in practice between interdisciplinarity and quality. The two-score approach used in Round 1's outline stage attracted mixed reactions among IAC members: Some found that it reduced complexity and ensured that due consideration was given to IDR, but others felt that it had precluded discussion about the relationship between the two criteria. In Round 2, there was greater emphasis on the quality, fit, and co-creation aspects of applications, but some IAC members still felt that one aspect was sometimes given more weight over the other.

Survey evidence broadly supports these findings. Through open-text questions in the survey, IAC members noted that Round 2 improved transparency and enabled discussion of differing viewpoints.

Around half of the rovers and chairs/deputy chairs agreed that a panel meeting with all members present was more effective at identifying strong IDR proposals and enabled a more robust assessment process compared to the process at the Round 1 outline stage (Figure 172 and Figure 173 in Technical

Annex C). Most remaining respondents were neutral, while only a small minority expressed disagreement. Respondents were particularly positive about the outline stage of the assessment process. 94% of panel members, and all chairs/deputy chairs agreed that the outline stage helped identify and prioritise high-potential applications that met the scheme's objectives and should progress to the full stage. Respondents also agreed that the outline stage reduced burdens on both applicants and UKRI when managing a high volume of applications.

Several successful applicants commented that, in fact, they were not confident that they understood the criteria, for example, the "fit to scheme" in Round 2, but still found them to be fair. Some IAC members as well as applicants also commented that the specific criteria were less important and that high quality applications tended to win out regardless.

IAC members observed an improvement in Round 2 with respect to the clarity of the criteria and the fit of the applications to the criteria. For example, one interviewed IAC member felt that UKRI had improved their guidance on what constitutes an A* scored application. However, a few IAC members raised concerns (through open-text survey questions) that, while Round 2 improved operational efficiency, it did not fully resolve questions about how best to evaluate and differentiate interdisciplinary. They noted persisting issues in distinguishing between high quality applications, inconsistencies in score interpretation, and concerns that a single score could obscure weaknesses in interdisciplinarity.

IAC members also noted that, whereas in Round 1 they received several emails related to the scheme's processes which for some caused confusion rather than clarity, in Round 2 communication was streamlined, contributing to a more efficient process overall.

Overall, stakeholders saw the assessment processes as fair, robust, and more effective in Round 2. The factors underpinning this view included more clarity on and better understanding of scoring criteria, improvements to training, and improved information and guidance for both applicants and the IAC. The Round 1 experience contributed to this improved clarity and understanding.

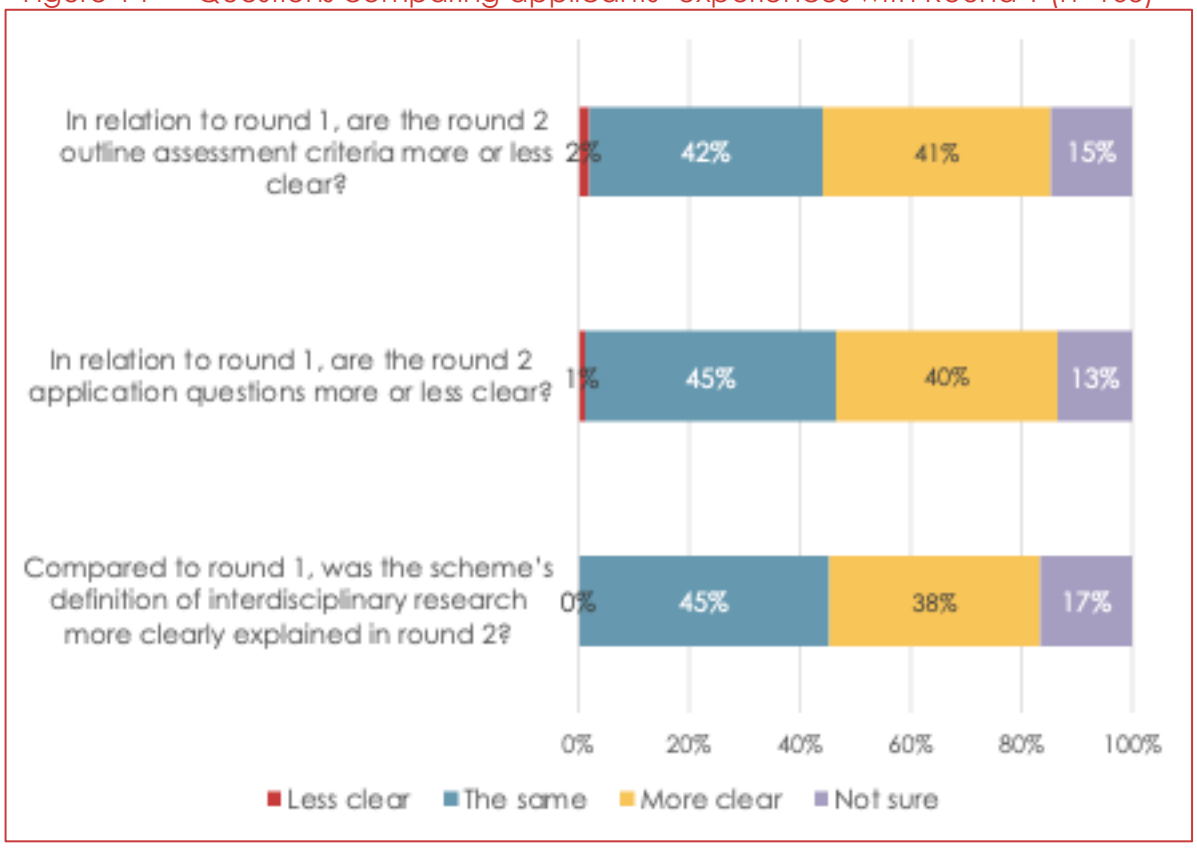
> "I've been a member of the IAC since its inception. I've seen us growing stronger with each panel round. The conversations have progressed to a more integrated level of understanding, with members from difference disciplinary backgrounds engaging with one another in a stronger understating of interdisciplinary. I feel sure that the training, which has obviously been responsive to our previous feedback, has contributed to

that progression. Thank you for providing this support, which I am confident has led to better decisions." IAC member

"The scheme has been a terrific success. This is new territory, and all kinds of issues were faced by UKRI and the process design. I'm not sure it could have been handled better or gone any better." IAC Chair

This progress between Round 1 and Round 2 was also highlighted in the survey responses (Figure 14). The majority of applicants (>80%) found the scheme's definition of interdisciplinary research the application questions, and the assessment criteria to be either "more clear" than, or "the same", as Round 1.

Figure 14 Questions comparing applicants' experiences with Round 1 (n=163)



Source: CRCRM Round 2 outline stage survey

2.4.2 A community dedicated to IDR assessment

Another key success, recognised in Round 1 and reinforced in Round 2 interviews, was the creation of a community trained in IDR assessment. Several interviewees observed that a broad pool of expertise in the IAC strengthened the assessment process, and that this expertise was appropriate for assessing the IDR applications received in the pilot.

> "This is an exciting scheme that recognises the value of high risk, interdisciplinary research and combines assessment of interdisciplinary working as well as the research itself. To this end it's been invaluable being on the panel to understand the UKRI perspective on what good interdisciplinary research looks like at all sorts of levels, and feed that back into both my own, and my institution's approaches." – Panel chair

2.4.3 IDR Assessment processes allow for all voices to be heard

The assessment panel process was effectively designed to ensure that all voices could be heard.

While Technopolis and IRC evaluators could not observe every panel in the Round 2 outline and full stages, the ones we were able to attend had very similar structures that we suspect were replicated across panels. In the panels we observed in both stages, each application was introduced by a lead assessor (IM) who summarised the application and reflected on its strengths and weaknesses and then shared their final score. Then the second and third IMs were invited to do likewise, often choosing to build on or challenge things the lead assessor had already said and sometimes to raise additional strengths or concerns. This process accorded each of the three IMs more or less equal time to speak without regard to their career stage, discipline, experience, etc.

Once the initial scores were presented the chair facilitated a conversation between assessors in an effort to find consensus on a final score. Where there was little disagreement, these conversations were relatively short. In cases with larger differences between scores, the process could take longer. IRC tracked these conversations to try to understand whether, once these conversations were more freeform, assessor characteristics affected how often they spoke up or whether some were reluctant to share their thoughts. They found that the frequency that individual assessors spoke was more reflective of the degree of divergence in scores and how much interaction was required to reconcile them than of "fairness" or imbalances due to seniority/confidence etc.

> "Getting the chair and co-chair together early to work through the process was an excellent idea". – Panel chair

Where there were disagreements over the final score, panelists had to compromise. The main difference between the panels we observed was how active the chairs were in managing these discussions. In some they would collate what they thought the main points of disagreement were and asked specific panelists specifically about issues they had raised while other chairs

were more open to letting those discussions happen more organically among the panelists. The former approach was perhaps more effective at ensuring a summary of the views of all panelists; it also potentially gave less time for individual panelists to think out loud about whether they still held those views after hearing others' perspectives. The latter approach enabled panelists to surface the issues they felt were still sticking points after everyone had had a chance to speak and to add depth to their arguments. This tended to give more time to those panelists that had assigned outlier scores. While each approach has its advantages and disadvantages, we do not believe that either reduced the ability of panelists to have their voices heard. Indeed, the consensus requirement meant that each panelist had to take the others' views into account and all of the discussions we witnessed showed panelists very willing to speak up and explore each other's perspectives and constructively arrive at a final score.

While all panelists were ultimately willing to shift their scores to some degree, the main disagreements came down to how far individuals needed to come up or down in order to align with others. However, there did not appear to be any clear pattern in willingness or unwillingness to compromise on scores. People without direct disciplinary expertise relevant to an application were slightly more likely to defer to other IMs but, in most cases, the divergence was about interpretations of interdisciplinarity than discipline/subject specific issues.

> "It seemed clear to me in the panels I observed that the very exceptional ground-breaking and transformative IDR research applications were properly identified" – Rover

Overall, survey responses across both rounds indicate that discussions to build consensus in the panel meetings were generally well balanced and fair (around 80 to 90% across both rounds selecting 'strongly agree' or 'agree', see Technical Annex A, B, C and D). Further, in Round 1 full stage, over 60% indicated they felt able to contribute to discussion where they were not an IM (n = 90), and in Round 2 full stage, almost all respondents (92%, n = 63) agreed or strongly agreed that the IAC will ensure that interdisciplinary research is fairly assessed across disciplinary boundaries/council remits. According to interviewees, in some panels, compromise was facilitated by the trust built over time between panel members.

Survey responses were overwhelmingly positive regarding the breadth of IDR expertise and experience represented on the panels. Across all stages, a large majority of respondents agreed that panels had a suitable breadth of expertise and experience to provide fair scores and comments, with

agreement ranging from around 80% to almost 100% (see Technical Annex A, B, C and D).

Panelists rarely mentioned that they lacked confidence to assess the application or reservations that one of the disciplines was not sufficiently covered. Notably, though, in all the panels IRC observed, the STEM panelists seemed confident about assessing applications with AHSS components. This could be because they also have experience in interdisciplinary working in those areas (that was not necessarily evident from their declared expertise in the applications). We suspect that there may be cases where STEM assessors assumed that they fully understood AHSS contributions where AHSS assessors would not likely feel the same about their grasp of the STEM components. Because of the composition of the panels IRC observed, there is no direct evidence to support this view, but this possibility is worth mentioning and potentially exploring further.

In all cases, there was a spectrum of experience levels with the specific topics of the applications across assessors. Typically, assessors represented a relevant discipline, but they would only rarely have direct experience with the specific content of the application. This is not a particular weakness, either from an alignment perspective (perfect alignment would be impractical) and it did not seem to affect the degree to which panelists were able to participate in the discussion (most would acknowledge and speak from their individual perspectives). When panelists had expertise specifically relevant to an application, they did not necessarily dominate the conversation. In these cases, expertise was usually only mentioned to highlight novelty (rather than overruling other less expert or experienced assessors). This observation from IRC contradicts that of a Technopolis observer who in a couple of panels saw that the consensus building or wider discussion of an application was sometimes influenced by the panel member with the most expertise in the topic area.

2.4.4 Stakeholder views on partial randomisation

Interviewees across stakeholder groups (IAC, UKRI and applicants) expressed different views on the use of partial randomisation. The majority of interviewees tended towards a positive view with some seeing this as the right thing to remove biases, whereas others saw it as more of a practical necessity when the quality of applications cannot be meaningfully ranked. Others, particularly IAC members, felt that the need for randomisation could be removed through improvements to the assessment process, for example, through better calibration of the scoring (that is, reducing the number of high scores). Some applicants could see the practical rationale for using randomisation but instinctively felt the practice to be unfair. This is confirmed

in survey responses (open-text), in which stakeholders broadly accepted the rationale for partial randomisation as a mechanism to reduce bias in highly competitive funding decisions. Yet, a few applicants noted that randomisation was a demotivating factor to apply to the scheme, as it reduces the value of the expert assessment of the IAC and undermines the merit of a high-quality application. One suggestion was to increase transparency about how partial randomisation is used, the evidence supporting its benefits and how it operates in practice.

In addition, one interviewee suggested that the partial randomisation should be stratified across panels, to ensure applications are randomised in proportion to the numbers within each panel. Another theme discussed by a few successful applicants was the re-submission process for high-scoring applications which were not funded as a result of the partial randomisation process. This issue had already been addressed for Round 2, where Band A applications that were unsuccessful through randomisation in Round 1 were invited to resubmit to Round 2 under the UKRI invited resubmission policy. Applicants in this situation were made aware that they would be subject to selection by their organisation as part of the demand management process, that their resubmission should address any feedback received and that their application would be assessed through both stages of the assessment process and treated as a separate submission, with no guarantee of being invited through to the full stage or awarded funding. There was a marked shift in IAC members' survey responses from Round 1 full stage to Round 2 full stage. In Round 1, Chairs had a mixed view on partial randomisation, with an equal proportion of chairs (23%) feeling extremely satisfied, satisfied, neutral, or dissatisfied. However, in Round 2, 70% of chairs felt the partial randomisation process added value and helped to counteract potential bias (see Figure 268 in the Technical Annex). It should be noted that the partial randomisation process had not been publicised at the time of the Round 1 surveys, so it wasn't clear at the time that applications still had to meet a high-quality threshold to be included in the randomisation process.

2.4.5 Challenge in assessing fit to scheme

A key challenge identified in outline stage panel observations was the assessment of fit to scheme. Specifically, panels spent a substantial amount of time judging whether an application could possibly be funded through other UKRI council-led responsive mode funding schemes and whether it adequately fitted the scheme's definition of interdisciplinarity. This chimes with some interviewees and survey respondents reporting some uncertainty about how to best approach the assessment of IDR. The feedback received on training, particularly the call for greater emphasis on practical examples, suggests that this aspect could benefit from further clarity even though it is

evident that there were improvements in the information and resources provided for Round 2, and that both applicants and IAC members had a better understanding of the scheme in that round. This was evident in the Round 2 full application panel meetings which were smoother, more efficient and involved focused discussions, albeit these meetings covered fewer applications.

A connected issue observed in the Round 2 full stage panel meetings was some lack of clarity among panels about some of the scoring definitions and differentiating between excellent (score 8) and outstanding (score 9) applications. While guidance had been provided to panels on what to consider as weaknesses, panel observations and panel feedback indicated that some lack of clarity still remained.

2.4.6 Technical challenges

Some IAC members faced technical challenges in conducting their work. These included access problems to the extranet, not understanding how to log on, or being unable to transfer their notes from the 'notes screen' to the discussion board as required. Furthermore, at the round one full stage, UKRI transitioned to The Funding Service (TFS), which confused IAC members who now had to familiarise themselves with a new system and work across two systems including the TFS and extranet. These technical challenges were compounded as some IAC members did not know who to contact for support. UKRI offered weekly drop-in sessions for both rounds and stages throughout the assessing period to support IAC members on this aspect.

2.4.7 Remaining challenges

IRC analysis highlighted some challenges to assessing IDR that surfaced as part of the evaluation. None of these are serious deficiencies of the assessment processes followed in the CRCRM pilot, but they do reveal some considerations for future schemes seeking to support IDR.

IRC team members observed an overall increase in the quality of interdisciplinarity across rounds and stages as measured by the confidence intervals of mean Signals of Interdisciplinary Focus in Text (SIFT) scores, a measure of interdisciplinarity (see Appendix B.5), at each stage. Applicants demonstrated integration of disciplines and co-creation in their applications, however, it was noted that applications continued to struggle to effectively articulate reciprocal benefits. Reciprocal benefits were described to applicants in the Round 2 call webinar as the disciplines being changed or transformed in some way by working together. Applicants needed to demonstrate that the contributing disciplines will gain something significant from the collaboration e.g. that the discipline's knowledge base has been

changed or transformed by working with other disciplines therefore showing that that discipline is not just being used as a tool for a different area of research. It was made clear that changes or gain did not need to be equal across the disciplines.

In reading applications and comparing these to assessor scores, comments, and panel discussions, the IRC evaluators think that assessors (and panels) were mostly likely to give applications the benefit of the doubt on this dimension of interdisciplinarity, especially when other dimensions (for example, integration of ideas and teams) were judged to be sound. For example, in line with instructions to assessors, the trend was (for the sample of applications that we reviewed) that if reciprocal benefits could be reasonably expected for *at least two disciplines* involved then that was sufficient to score relatively highly on that dimension of interdisciplinarity. It was considerably more difficult to demonstrate the potential for reciprocal benefits for every discipline involved (much less that those benefits would be equally felt).

The lesson here is that convincingly articulating that research will generate reciprocal benefits, is the highest bar for IDR, and the most difficult aspect to assess. Future schemes should put some thought into how important this dimension is in assessment and devote additional effort to demonstrating what thresholds are and are not acceptable (for example, if strong reciprocity for some, but not all disciplines, is sufficient) - to both applicants and assessors. Ideally, the application pool from the CRCRM pilot would be made available to those administering future schemes so that they can test whether their planned assessment protocols are realistic.

In the context of differences in the number of "pure" AHSS collaborations and "pure STEM" combinations and the wide variety of disciplines covered by the AHSS councils, possibly it is more difficult for AHSS disciplines to articulate interdisciplinarity because the cognitive distance between them tends to be higher – these disciplines lack, and must therefore develop, a common language to describe dimensions of interdisciplinarity – more specifically problems and reciprocal benefits. In contrast, STEM disciplines may not need to work as hard to bridge those gaps. While these questions were not explored in this evaluation, this is an area that is ripe for analysis using measures of disciplinary distance alongside indicators of lexical complexity in panel applications.

2.5 Assessing IDR in applications – key learnings

2.5.1 Assessors that had IDR backgrounds combining STEM and non-STEM gave lower scores

In their application to the IAC, assessors were asked to specify their councils and remits. For some, the councils selected were exclusively STEM, for others they were exclusively AHSS. Some assessors indicated experience in both STEM and AHSS councils. An interesting observation from IRC was that, in both rounds, assessors whose selected councils spanned both STEM and non-STEM (i.e., AHRC or ESRC plus a STEM council) were tougher graders. More specifically, assessors with both STEM and non-STEM background gave lower scores in both rounds (overall score in all of Round 2 and Round 1 full applications and interdisciplinarity score in Round 1 outline stage) compared to assessors belonging only to STEM or non-STEM councils. This suggests that assessors with interdisciplinary experience that spans STEM and non-STEM areas were more reluctant to give higher scores, probably because they were not being easily persuaded by the claims articulated within projects concerning integration of disciplines, compared to assessors with interdisciplinary experience that did not cross the boundaries of STEM and non-STEM.

Another finding related to assessors' interdisciplinary experience was that the higher the number of councils identified by UKRI for each assessor, the lower their scores. This suggests that assessors who have prior experience in interdisciplinary research may have higher standards for interdisciplinarity and, therefore, are stricter and more sceptical in their scoring. This finding was detected in both rounds, although the score in Round 2 refers to overall score while the score in Round 1 refers to the interdisciplinarity score.

Analysis of Round 1 scores showed that the assessors who (1) reviewed more applications or (2) assigned scores across proposals within a narrower range (that is, clustered in a narrower part of the scoring range/narrower deviation) tended to give higher scores but this was not detected in Round 2 (overall score in Round 2 and interdisciplinarity score in Round 1).

Finally, in both rounds IRC did not find any relation between assessors' career stage or role and their given scores.

2.5.2 A two-stage process incorporating deliberation is very valuable

The two-stage process with an outline stage was valued because it reduced the burden for writing, processing and assessing applications.

Outline stage panel meetings as in Round 2 were preferred compared to a moderation meeting with Chairs, Deputy Chairs and Rovers as in the Round 1

outline stage. Deliberation between experts made for a more robust, fair and efficient process which reduced the 'double jeopardy' of IDR applications being seen through multiple monodisciplinary lenses. It was noted in the IAC surveys that the Round 2 outline stage panel meetings were more effective at identifying strong IDR applications compared to the moderation meeting in Round 1 – 53% of chairs and deputy chairs (n=15) and 38% of rovers (n=13) agreed or strongly agreed with this view.

The introduction of readers in the Round 2 full stage further supported in-depth deliberation between experts and more robust assessments, especially if introducing panel members were absent in the meeting.

2.5.3 Combined feedback and obtaining clarification from project leads adds value to assessment

Using discussion boards to create a combined set of feedback mitigated the risk of contradictory reviews from different disciplinary perspectives in the project lead response stage. In interviews, applicants and IAC members both agreed that the opportunity for panel members to ask clarification questions of applicants was very useful. It improved assessment as panel members could make more informed decisions on applications. The Round 2 process involved improved UKRI guidance for IMs on how to structure their feedback and questions following prior coordination and discussion between the three IMs. This made for better structured feedback and clarification questions and a smoother process. This was appreciated by IAC members.

3 CRCRM portfolio

The findings in this chapter show that the CRCRM pilot was successful in funding high quality, transformative, and well-integrated research that is likely to be highly interdisciplinary.

3.1 Quality of applications

Summary findings

- Panel members found the quality of outline proposals to be variable, including very strong as well as weaker proposals
- The assessment process was seen to be effective in identifying high-quality applications
- Round 2 applications were judged to be of higher quality and better adapted to the scheme than Round 1 overall, although some panel members also felt Round 2 applications were more risk-averse.

In interviews, IAC members were asked about the quality of applications at the outline and full application stages. For Round 1, they stated that the quality of applications varied at the outline stage. However, interviewees emphasised that the processes allowed the identification of high-quality applications to take through to the full application phase. The full stage applications that followed were of much higher quality, including those that were ultimately unsuccessful.

> "There were some very strong proposals in our panel that I believe will deliver world-leading research. Encouraging more of the community to consider IDR by demonstrating that these higher risk ideas can be funded, if realistically planned, should make a very strong contribution to the UK's international competitiveness in research" – IAC member.

A few interviewees mentioned that, while a large portion of the very low-quality applications had been weeded out in the outline stage in Round 1, a few full stage applications of low quality were detected. This may reflect a difference in quality between the outline and full stage applications or a limitation of the Round 1 outline stage assessment process which involved two separate scores and a moderation panel rather than a panel meeting.

For Round 2, IAC members interviewed for the evaluation generally found the applications to be of a higher quality and better adapted to the scheme's aims compared to those in Round 1. Some IAC members, however, also felt that the applications were more risk-averse and less innovative compared to

Round 1. These individuals speculated that the introduction of demand management in Round 2, or Round 1 applicants and IAC members sharing their experiences with colleagues at their home institutions, may have led to researchers and institutions putting forward research applications that they perceived to be 'safer bets' with a higher chance of being funded.

Survey findings broadly support these observations. When asked to compare the quality of applications between Round 1 and Round 2 full-stage, IAC respondents generally perceived improvements in the interdisciplinary of applications in Round 2. Many IAC rovers (75%) and almost half of IAC panel members (45%) also considered Round 2 applications to have a greater potential to deliver ground-breaking outcomes whereas IAC chairs (69%) felt these aspects were similar across rounds. Novelty and ambition were considered comparable across both rounds (see Figures 247 and 248 in the Technical Annex).

3.2 Characteristics of collaborations being supported

Summary findings

- Social sciences disciplines featured far more widely across the portfolio than the panel interface labels suggest. They appear to some degree in applications across nearly all interfaces in both rounds, indicating the scheme may be supporting disciplinary combinations that go beyond its formal categorisation.
- IRC's SIFT analysis indicates that funded applications were more interdisciplinary than unsuccessful ones, and significantly more interdisciplinary than a random sample of UKRI-funded research. The gap widens in Round 2, suggesting the improved assessment processes were better at identifying true interdisciplinarity over time and/or applicants were better at articulating interdisciplinarity.
- Full-stage applications tended to be more regionally and organisationally diverse than the outlines from the corresponding round, suggesting the process was actively selecting for broader collaboration.
- On protected characteristics, there was weak evidence of slightly higher success rates for applicants declaring White ethnicities at the Round 2 outline stage, but the overall evidence of bias was limited and inconclusive: no reliable bias was found for age, disability, or sex.

3.2.1 Co-occurrence of different disciplines

The network diagrams in Figure 15 and Figure 16 represent the disciplinary combinations in the outline and full applications versus the awards for Round 1 and Round 2. The size (and colour) of each disciplinary node represents the

total number of applications that incorporated that discipline, while the thickness of the connecting line represents the number of applications that combined the two connected disciplines.

These diagrams show that social sciences feature more prominently among the applications than is implied by analysis that focuses only on the panel interfaces. In fact, it appears that social sciences are included to some extent in many of the CRCRM applications regardless of the interface in both rounds.

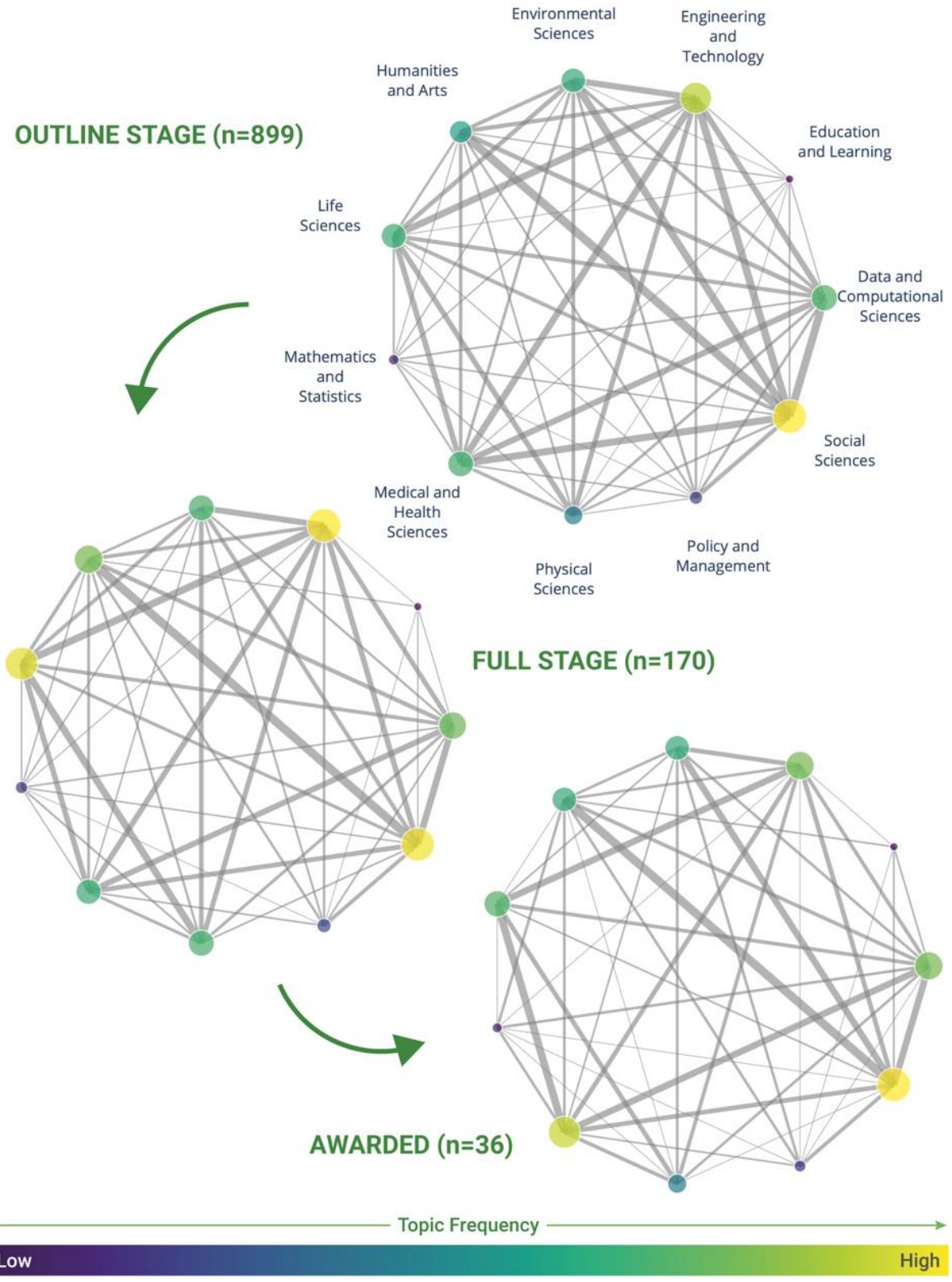
> "The response of the research community into integrating social and environmental science... was an exciting development." - Panel chair

In Round 1, medical and health sciences are more predominant in awarded applications than in the outline and full stages and for Round 2, as is the case with environmental sciences.

Panel-wise network maps of outline and full stage applications show the expected results, with the two main disciplinary areas at the panel's interface being the most represented (see Technical Annex H). This suggests that the remit tags and assignment to panels were very accurate. Social sciences also appear to some or a small extent in the panels where social sciences is not one of the two main disciplinary areas, which further evidences the point that social sciences are included in many of the applications and awards regardless of the interface.

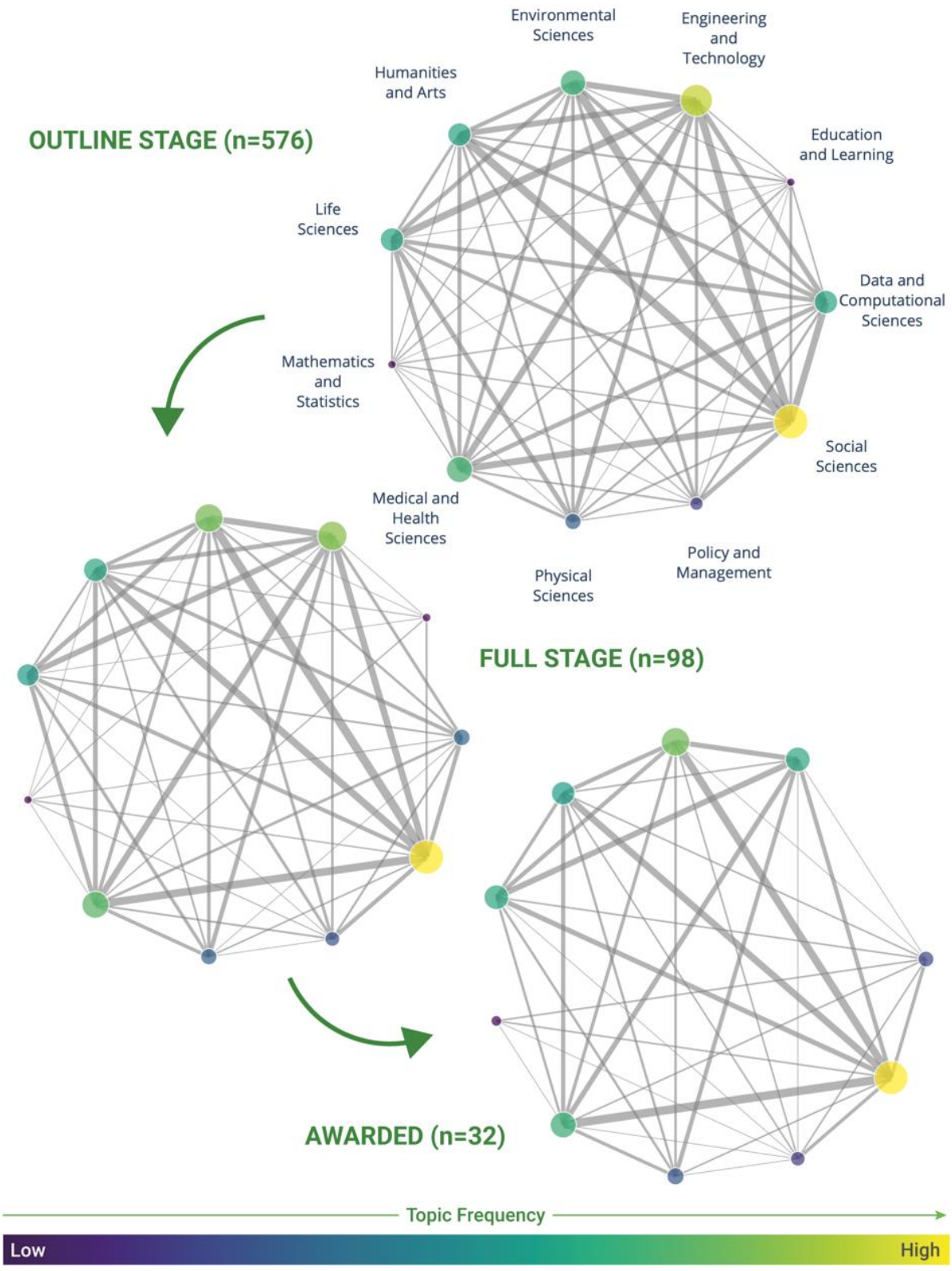
Network maps for awarded applications were not created per panel due to the small number of data points.

Figure 15 Network maps showing disciplinary combinations present in outline, full and awarded applications in CRCRM Round 1



Source: Technopolis. Infographic by Célia Couchot.

Figure 16 Network maps showing disciplinary combinations present in outline, full and awarded applications in CRCRM Round 2



Source: Technopolis. Infographic by Célia Couchot.

3.2.2 Signals of Interdisciplinary Focus in Text (SIFT) analysis

The co-occurrence and combination of different disciplines do not necessarily indicate interdisciplinarity, as applications may be multidisciplinary rather than truly integrative. The IRC developed a methodology to analyse interdisciplinarity as a measure of how well integrated disciplines, ideas, and teams are, as well as the characterisation of reciprocal benefits in application text – Signals of Interdisciplinary Focus in Text (SIFT). The methodology consisted of analysing application text to identify and score the strength of statements related to interdisciplinarity. For example, pertaining to the integration of disciplines, ideas, teams, ways of working, reciprocal benefits, or challenges of interdisciplinary research.

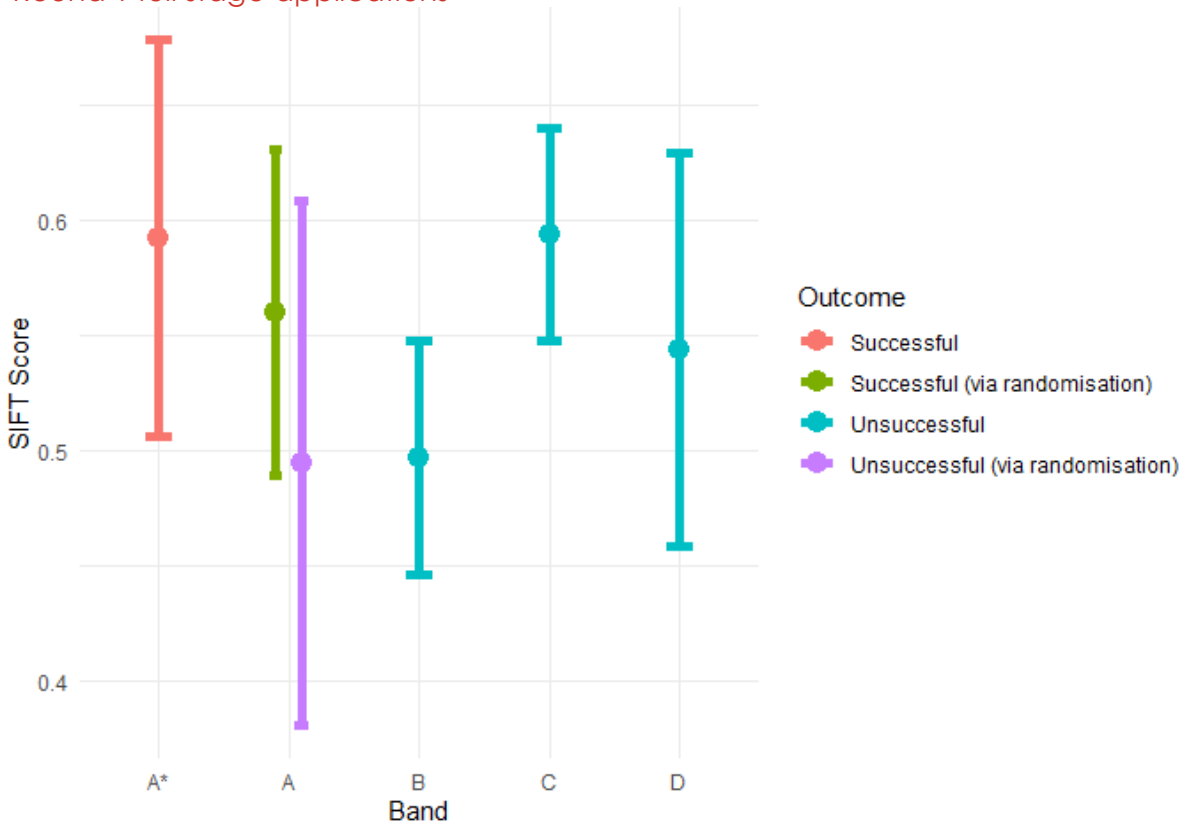
Over the two rounds, interdisciplinarity was assessed by panels in several ways due to the changes in processes implemented to improve the robustness of the assessments. In Round 1, interdisciplinarity was scored separately by assessors at the outline stage. It was noted by the moderation Panel Chairs and Rovers that the score was not always justified by the comments and there was inconsistency in the understanding of what is considered good IDR I by the assessors at this early stage of the scheme. For Round 1 full stage and Round 2, IRC used the SIFT methodology to determine a measure of interdisciplinarity. In Round 2 full stage, the panel was also asked to provide an interdisciplinary panel rating of each application to support the evaluation. These different scores offer slightly different interpretations and together provide a more nuanced perspective on the interdisciplinarity of successful and unsuccessful applications.

It was found that those applications that scored highly on interdisciplinarity, were, on balance, most likely to advance to the next round. While this might suggest that everything that advanced was highly interdisciplinary it is noteworthy that this was the round with the highest correlation between interdisciplinarity score and overall team diversity (organisational and disciplinary). This, coupled with assessor comments that often focused on team composition, suggests that in this earliest round assessors may have been relying, consciously or not, on team diversity as a proxy of interdisciplinarity and possibly did not weight factors such as integration, co-creation, or reciprocal benefits as highly. The result is that the projects that advanced may have been *diverse* but the research approach might not be as interdisciplinary as expected.

In Round 1 full stage, IRC pioneered the SIFT score with slightly more nuanced results. This analysis shows that funded applications generally had a higher SIFT score than unfunded applications (see Figure 17, below). Of the funded applications, A* had a higher range of SIFT scores than other bands. Successful A band applications also tended to have higher scores (and a

higher mean) than the unfunded applications. This result was somewhat fortuitous given that successful A band applications were selected by randomisation. However, it should be noted that there were 18 Band A successful applications and 9 Band A unsuccessful applications and confidence intervals become wider as sample sizes become smaller, and they become narrower as sample sizes increase.

Figure 17 Mean SIFT score by band and outcome with 95% confidence intervals, Round 1 full stage applications



Source: IRC

Band B (unfunded) applications had the lowest ranges of all the bands. Band C (unsuccessful) applications generally had stronger SIFT scores than Band B, while Band D also had a higher range of SIFTs than B. IRC suspect that these discrepancies in lower bands reflect applications being sufficiently interdisciplinary but being graded lower because of deficiencies in other areas (for example, novelty, proposed methods, team experience, value for money, etc.).

In Round 2 outline, IRC used some indicators to help test this hypothesis through the clarification statements sent to outline applicants as part of the feedback, chosen from pre-written set of statements, and selected by deputy

Chairs during the panel meetings. The feedback statements relevant to each of the assessment criteria including the Fit to Scheme relating to the interdisciplinarity of the application help to confirm these patterns (see Appendix B.3). In this round, 96 applications were characterised as having "clear fit to scheme" but nearly 40% of those (38) did not advance to the full stage. In other words, all of these applications were deemed highly interdisciplinary (and, therefore, not assigned any further critical feedback in this area) but only 58 advanced to the full stage. Of the applications that were deemed to have "some level of fit to scheme" (read: with interdisciplinary potential), only a small percentage (~11%) progressed. None of the applications that were assigned "low level of fit to scheme" advanced to the final round. This overall pattern demonstrates that Round 2 outline effectively weeded out the applications that did not clear the interdisciplinary hurdle but also that many applications that did, with high or some fit, also did not advance.

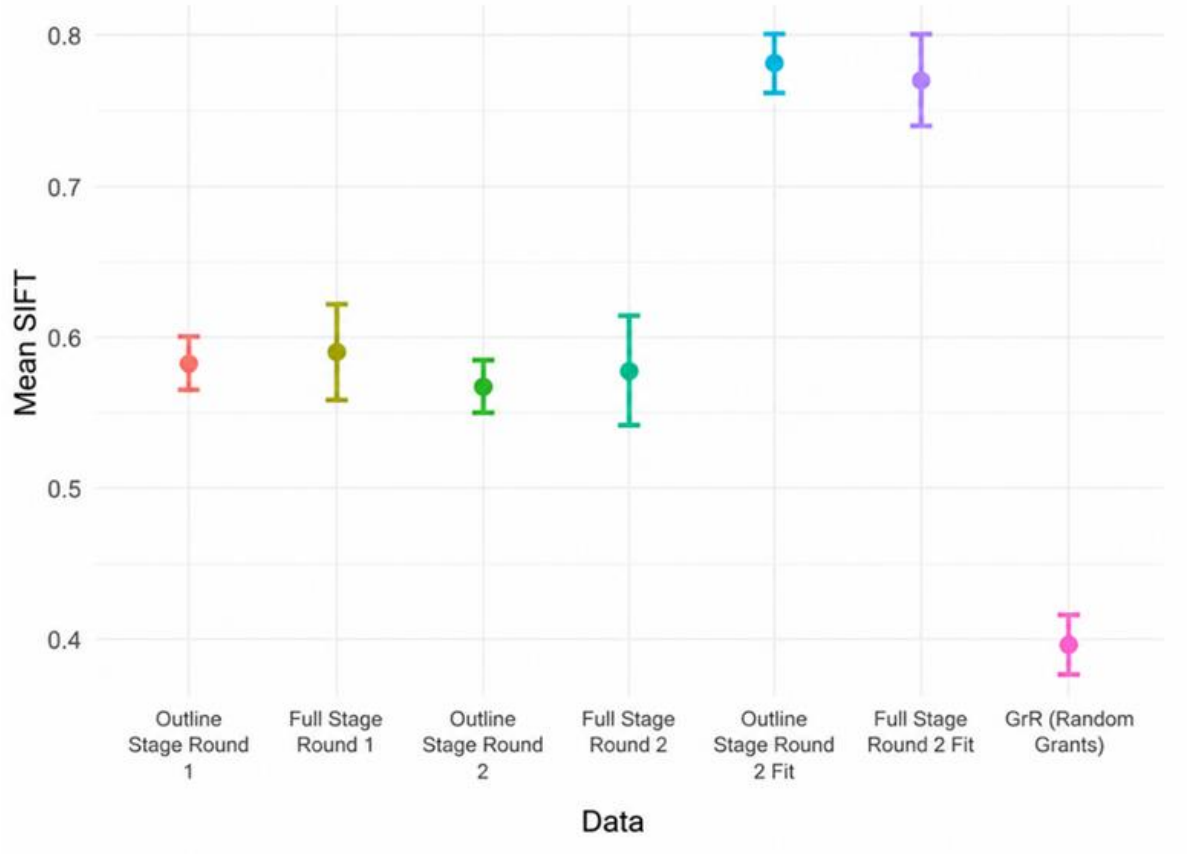
Looking closely at the additional feedback confirms our initial assessment. Of the 38 applications that were rated highly on interdisciplinarity but were unsuccessful, 37 of them received feedback that indicated minor or significant concerns about "the effectiveness and appropriateness of the outlined project approach in achieving objectives". The one project that was judged to have an approach effective and appropriate to objectives but that was also unsuccessful had feedback indicating concerns about team capability to deliver as well as requiring further clarity about how it could impact world-leading research. These results confirm that unsuccessful highly interdisciplinarity projects were largely held back due to concerns on overall research quality. Otherwise, a combination of concerns was also enough to prevent applications from advancing.

In Round 2 full stage, the feedback process to applicants was in the form of a written statement to the applicant and no separate quantitative measure of interdisciplinarity was done. Therefore, to support the evaluation and analysis of interdisciplinarity in applications, panels were asked to assign "Low", "Medium", or "High" scores to the following categories: "Integration of Disciplines", "Reciprocal Benefits", "Co-creation/co-design", and "Interdisciplinarity overall". This was done by the panel Deputy Chair after the application had been discussed and scored and was not sent to the applicant. IRC noticed that 72% of applications that scored "High" in "Integration of Disciplines" were successful, while the equivalent percentage was 66% for "Reciprocal Benefits", 56% for "Co-creation/co-design", and 67% for "Interdisciplinarity overall". It is also noteworthy that we found five unsuccessful applications that scored "High" in all four categories but were considered a lower priority for funding in the banding process when

considering the other assessment criteria. There were three successful applications that didn't score "High" in any of the four categories but were recommended for funding in the banding process due to overall excellence.

The full stage in Round 2 of the scheme had much higher SIFT scores than in previous stages/rounds (see Figure 18). This is largely due to improvements in training and guidance to applicants, which provided much stronger cues about the three components of high-quality interdisciplinary applications and how to demonstrate this in an application. It is also due to updated instructions on applications where the "fit to scheme" section became a dedicated space to make arguments about the interdisciplinarity of proposed projects. When using this section alone for the SIFT score the resulting score is much higher than when using the summary sections.

Figure 18 Mean SIFT score for CRCRM grants by round and stage using the application summary sections and the fit to opportunity section (Fit) for Round 2 applications compared with summary text from previously UKRI funded research available on Gateway to Research (GtR)



Source: IRC

Figure 18 also shows the SIFT scores from each round contrasted with a random sample (n = 576, chosen to match the number of summary texts available in the outline stage of Round 2) of summary text from previously

UKRI funded research available on Gateway to Research (GtR). The significant differences in mean scores and confidence intervals demonstrates that projects submitted to CRCRM were significantly more interdisciplinary than this random sample. This is especially striking from the "Fit to scheme" (Fit) text provided in Round 2.

On the whole, these results indicate that the process was working as intended. Research that was given high ratings for interdisciplinarity *and* assessed to be high quality advanced while those projects judged to be sufficiently interdisciplinary but not considered high enough quality research for other reasons did not. What this suggests is that there is a tranche of research that may be quite interdisciplinary but that is not being funded because of other deficiencies in the application and that, with proper feedback and guidance in research design, might have great potential in future interdisciplinary schemes.

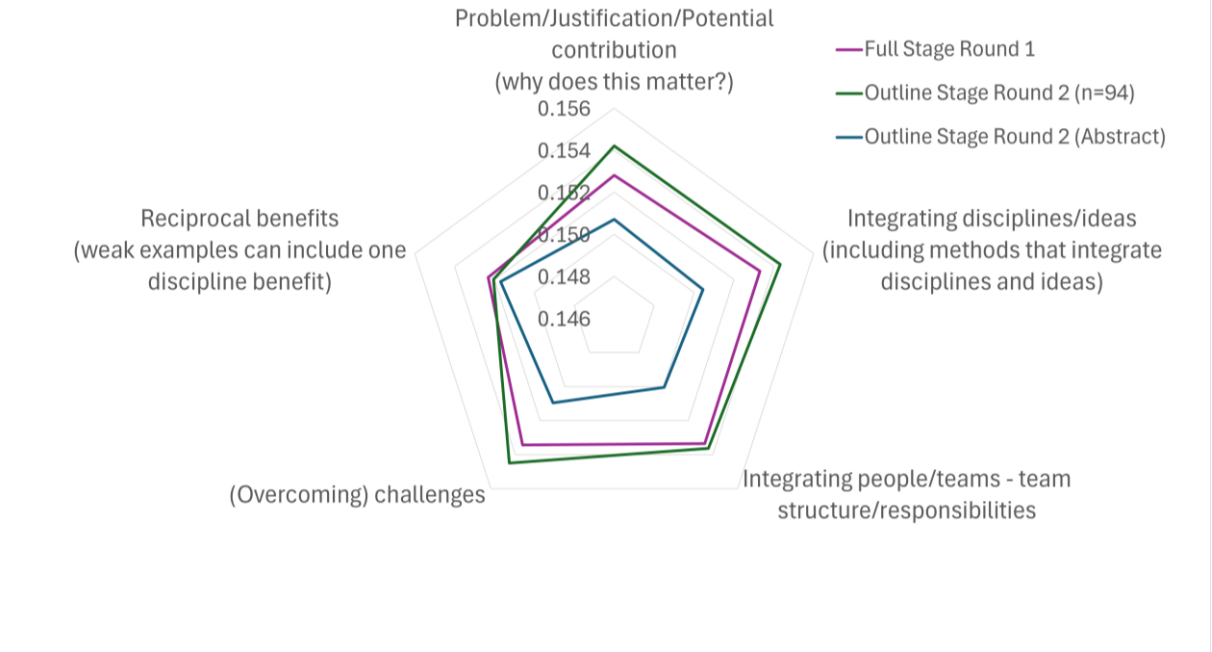
3.2.2.1 Exploring dimensions of interdisciplinarity

IRC did a deeper analysis of the nature of interdisciplinarity in Round 1 full in order to understand how different aspects of interdisciplinarity were affecting the overall SIFT score. We were interested to learn whether higher scoring applications were achieving those results because they were very good at articulating one or a small number of aspects of interdisciplinarity or whether they were effectively addressing them evenly. And, if there was a disparity between the frequency and scores of different aspects of interdisciplinarity, IRC wanted to know which ones were better articulated.

In this analysis, IRC categorised sentences that were selected and scored by the SIFT methodology as dealing with several aspects of interdisciplinarity in applications, notably: 1) the justification for an interdisciplinary approach in problem statements or characterisation of project outputs or contributions; (2) how disciplines and/or ideas would be integrated; 3) how people and teams would be integrated into interdisciplinary ways of working; 4) specific anticipated challenges of interdisciplinary research and how they can be overcome; and 5) articulations of reciprocal benefits.

IRC found that applications tended to contain more statements related to problem statements and the integration of disciplines/ideas and teams than discussions of challenges or reciprocal benefits. Furthermore, statements related to the first four factors tended to have high mean scores (with challenge statements scoring the highest), statements of reciprocal benefits tended to have lower mean SIFT scores.

Figure 19 Mean SIFT scores for dimensions of interdisciplinarity in Round 1 Full stage and Round 2



Source: IRC

Furthermore, we found that applications that had high SIFT scores overall generally scored higher across all dimensions of interdisciplinary sentences, disproving the possibility that high scores overall were obtained from excellence on only a few dimensions.

In Round 2 full application stage, the assessment panel process introduced feedback for dimensions of interdisciplinarity to support the evaluation. This data largely reinforces the trends in our findings. In this stage, panels gave applications higher scores overall for integration of disciplines and co-creation and design (roughly similar our previous team integration dimension). However, the panel feedback for reciprocal benefits and interdisciplinarity overall categories had lower scores, indicating that the panel more frequently found weaknesses on these dimensions. There was a strong relationship between the interdisciplinarity overall score – which is, as named, an overall assessment of interdisciplinarity intended to capture the overarching assessment of the three main dimensions (discipline integration, reciprocal benefits, and co-creation) – and reciprocal benefits. This indicates that applications that scored on the lower end of the overall dimension typically did so either because of a perception of extreme weakness on reciprocity or because of feedback about weaker reciprocity paired with weakness on at least one other dimension.

Overall, these different analyses show that it is possible to parse elements of interdisciplinarity (first, using SIFT and then using panel feedback scores) to

determine the different characteristics of interdisciplinarity. To summarise, there is strong evidence here that applicants are making efforts to articulate how their research approach will break down silos (hence stronger scores on dimensions of integration). That we see generally strong SIFTs – or low frequencies of low panel feedback – among higher banded projects indicates that the assessment process rewards these efforts (and penalises those projects that do not as effectively make cases for interdisciplinary approaches). The overall lower scores on reciprocal benefits demonstrates that this is the most difficult area to articulate and often the hurdle at which otherwise high-quality applications stumble.

Finally, while these findings indicate that this scheme achieved its goal of attracting and funding projects that are likely to serve IDR goals (and were not conceived in silos), it is important to note that a final determination of how successful this endeavour was cannot be made until projects are completed and their actual ways of working, integrating, and generating benefits are assessed.

3.2.3 Diversity of organisations, career stages and protected characteristics

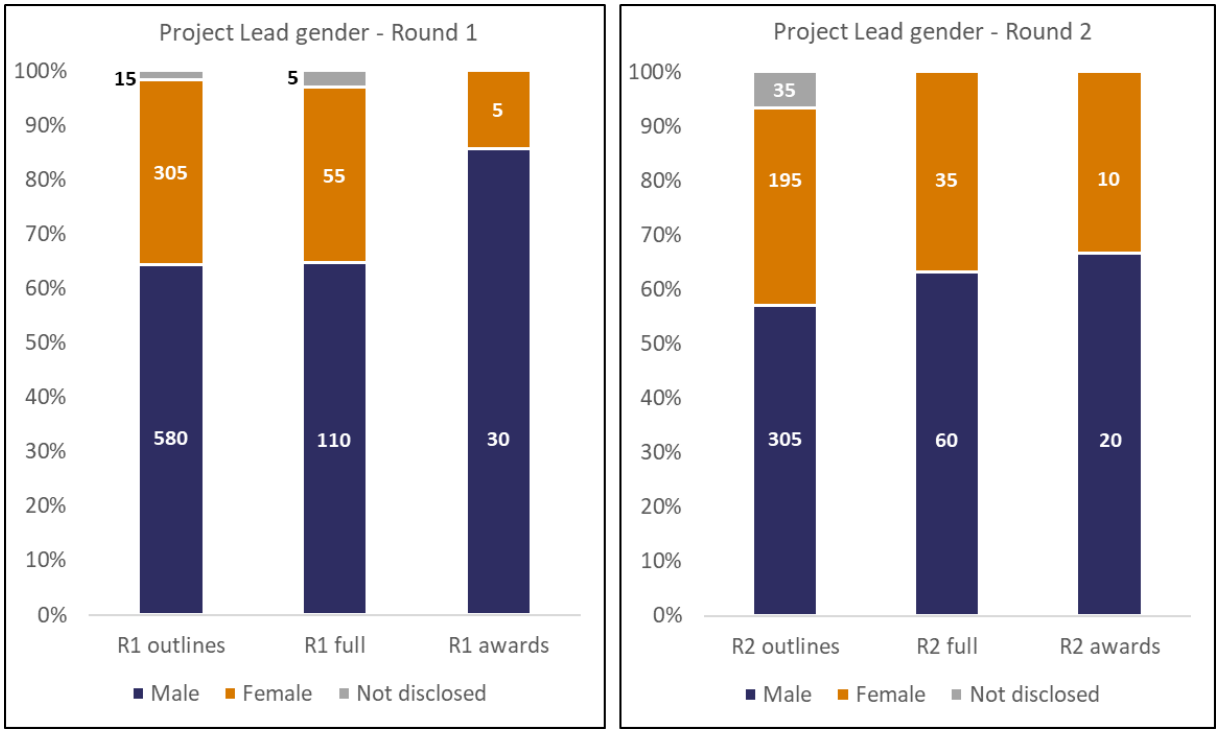
UKRI conducted analysis associated with protected characteristics for age, disability status, ethnicity and sex. In addition, they also analysed the portfolio of applications based on research organisation, region and team make up.

Diversity data is suppressed for categories with numbers between 1 and 4. These values are shown as <5 in diversity data presented in Tables 10-13 in Appendix B.2. Counts of 5 or more are rounded to the nearest multiple of 5. Therefore, counts do not always add up as expected when comparing to known application numbers for each round and stage. It should be noted that the applications for Round 1 were submitted to Je-S, and Round 2 applications were submitted to the UKRI Funding Service which has impacts for how applicant data was collected. For Project Lead gender data were collected under the 'Gender' section on Je-S, but on TFS are collected under the 'Sex' section, where there is a separate question on 'Gender assigned at birth'. Different ethnicity categories were used on Je-S and the UKRI Funding Service, therefore, there are differences in the ethnicity categories between Rounds.

Across both rounds, male project leads make up the majority at every stage, with the proportion generally increasing from outline to award stage. In Round 1, female project leads account for around one-third at outline and full stage applications but are a smaller proportion by the award stage. In Round 2, although the proportion of female project leads still decreases from outline to award stage, the drop is less pronounced, resulting in a more balanced progression compared to Round 1 (see Figure 20).

Figure 20 Project Leads gender at each assessment stage and awards for Round 1 (a) and Round 2 (b).

(a) (b)

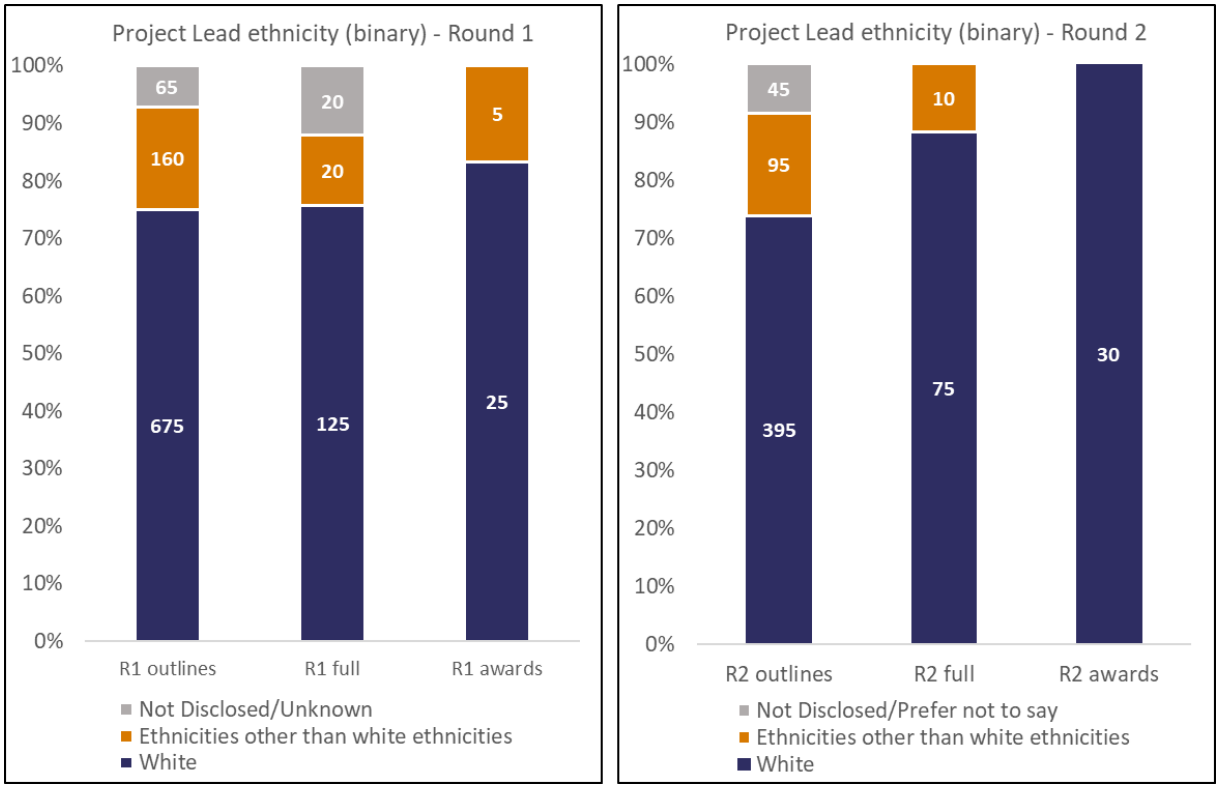


Source: UKRI

Across both rounds, White project leads represent the majority at every stage, with their proportion generally increasing from outline to award stage. In Round 1, project leads from ethnicities other than White make up a notable minority at outline and full stages but decline by the award stage, alongside a reduction in "not disclosed/unknown" responses. In Round 2, a similar pattern is observed, with representation of project leads from ethnicities other than White decreasing from outline to award, resulting in an almost entirely White cohort at the award stage (see Figure 21).

Figure 21 Project Leads Ethnicity (binary) at each assessment stage and awards for Round 1 (a) and Round 2 (b)

(a) (b)



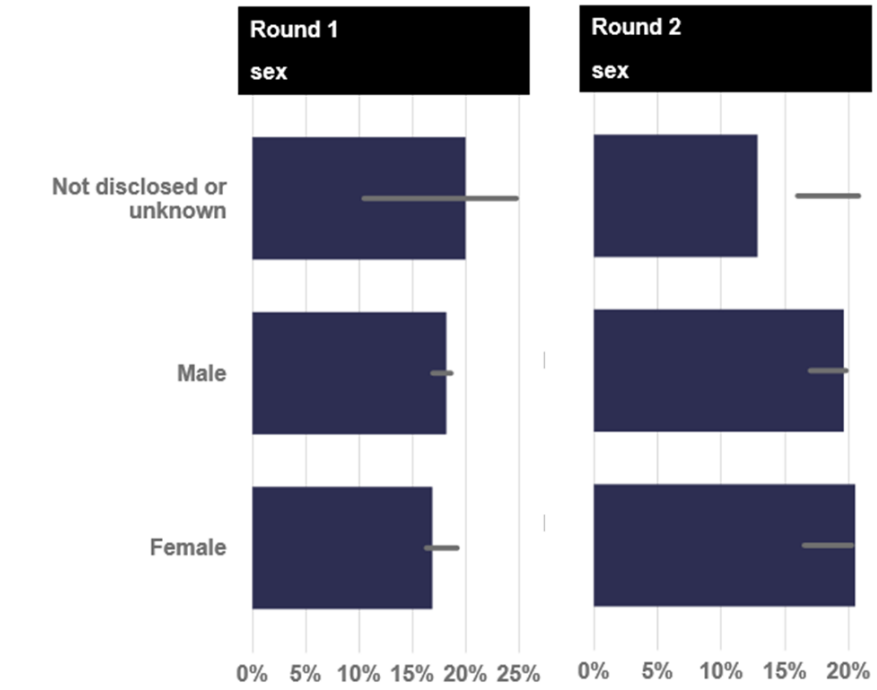
Source: UKRI

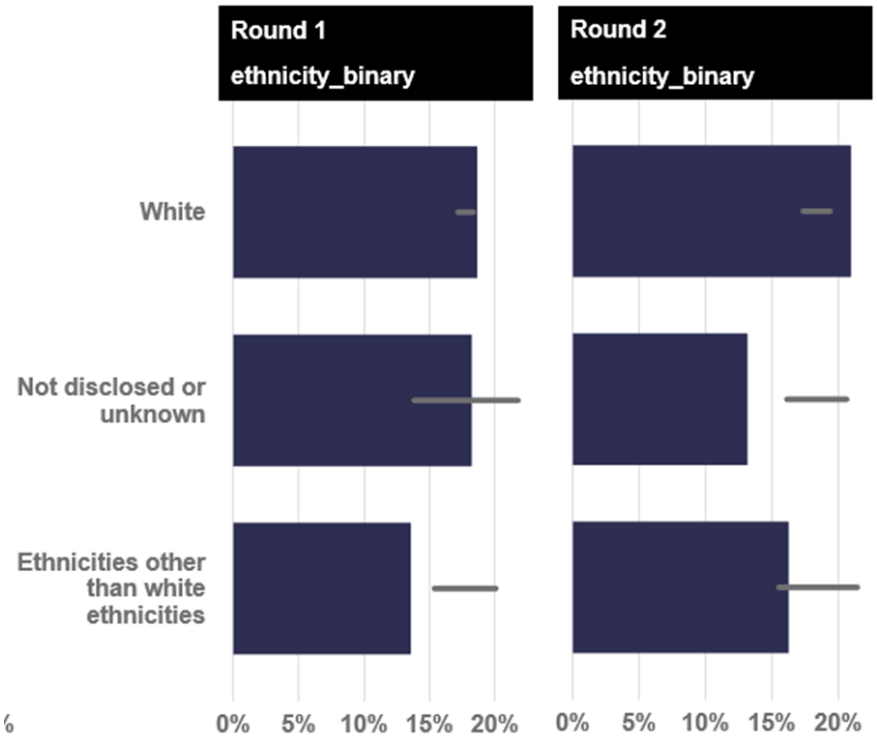
Data on Project Lead disability and age ranges can be found in Appendix B.2.

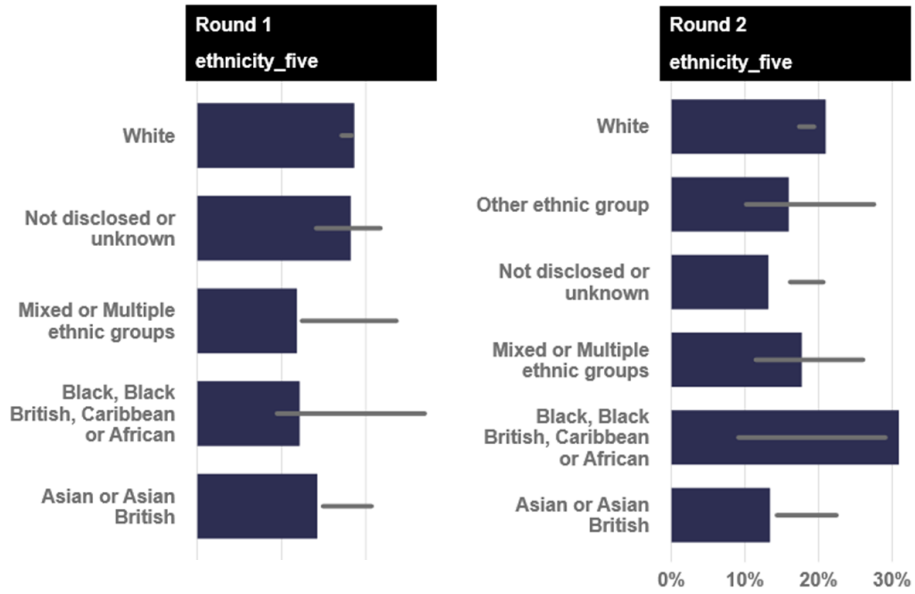
Analysing the success rates for these characteristics, there was some suggestion of bias related to ethnicity, but none of the others. In Round 2 outline stage, applicants declaring White ethnicities had slightly higher success rates. The data is harder to interpret at the full application stage as there are fewer applicants, meaning that the expected ranges become wider and the observed values more variable. There is some evidence suggesting that applicants declaring White ethnicities had lower success rates in the full application stages, but this is not definitive. Overall, the evidence of bias relating to ethnicity is quite weak and for other characteristics, there is no reliable evidence of bias (see Figure 22).

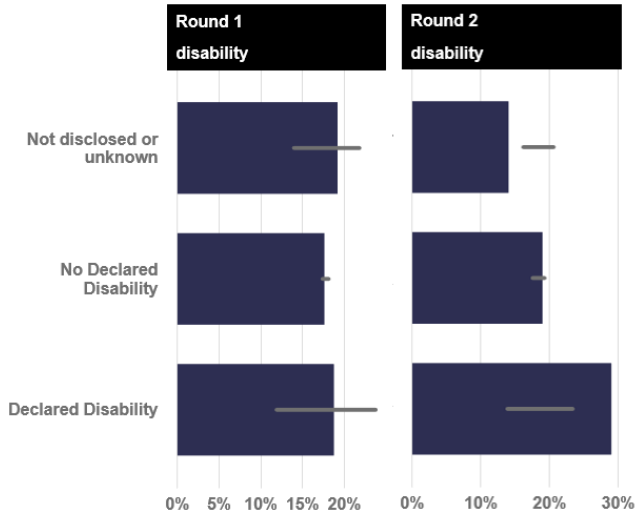
Figure 22 Success rates associated with protected characteristics at outline (a) and full (b) stages for Rounds 1 and 2

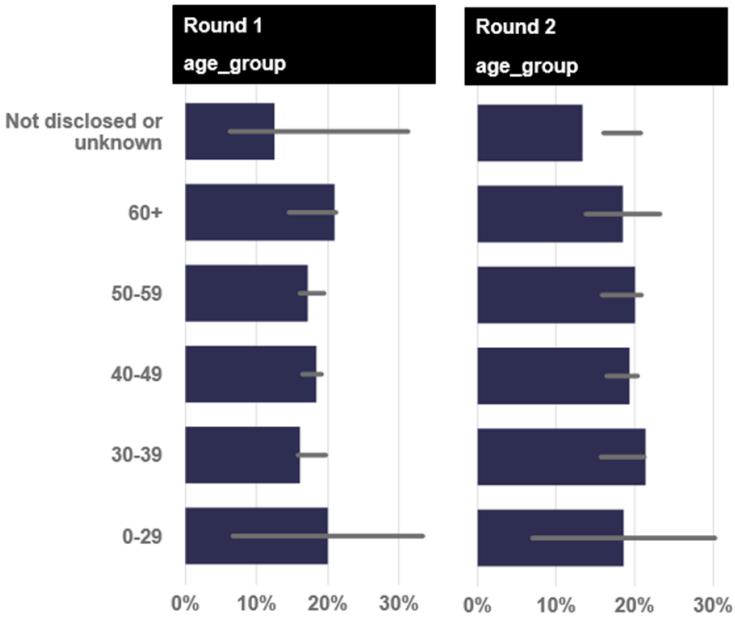
(a)



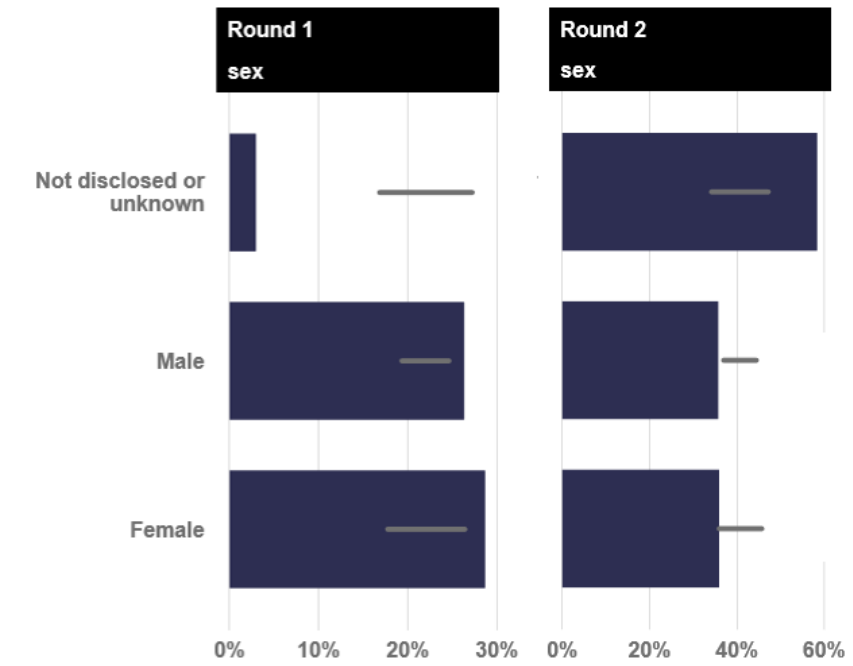


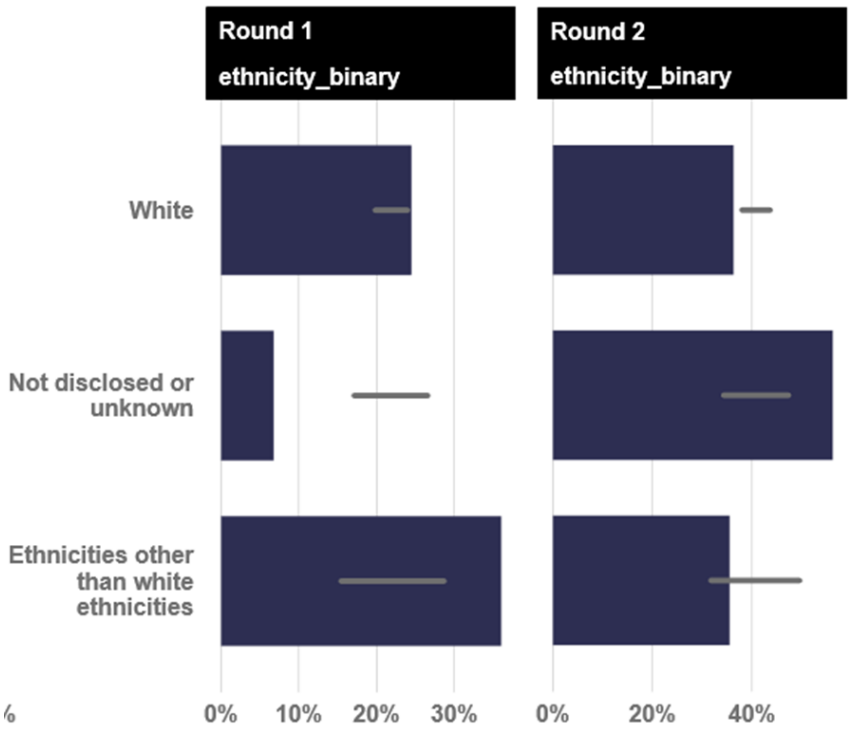

Round 1
ethnicity_five
White
Not disclosed or unknown
Mixed or Multiple ethnic groups
Black, Black British, Caribbean or African
Asian or Asian British
Round 2
ethnicity_five
White
Other ethnic group
Not disclosed or unknown
Mixed or Multiple ethnic groups
Black, Black British, Caribbean or African
Asian or Asian British
0%
10%
20%
30%

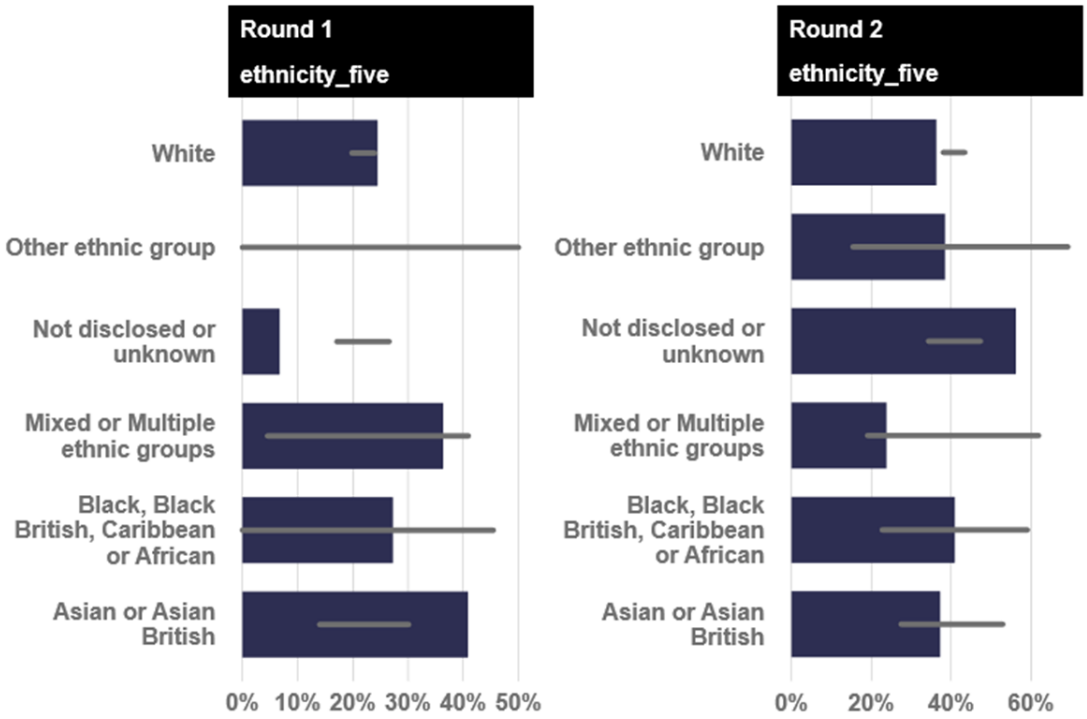

Round 1
disability
Round 2
disability
Not disclosed or unknown
No Declared Disability
Declared Disability
0%
5%
10%
15%
20%
0%
10%
20%
30%

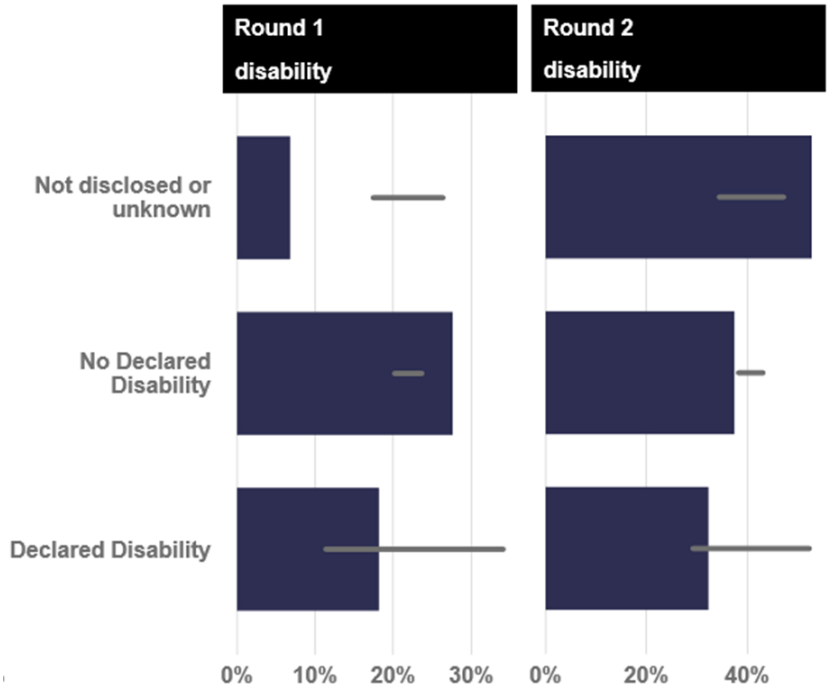

Round 1
age_group
Round 2
age_group
Not disclosed or unknown
60+
50-59
40-49
30-39
0-29
0%
10%
20%
30%
0%
10%
20%
30%

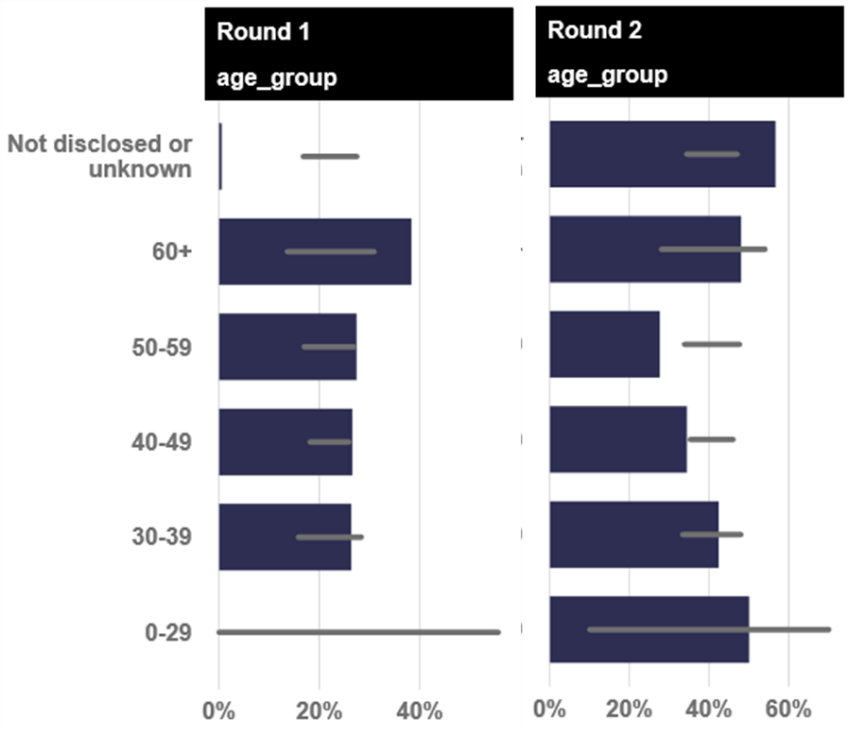
(b)


Round 1
sex
Round 2
sex
Not disclosed or unknown
Male
Female
0%
10%
20%
30%
0%
20%
40%
60%


Round 1
ethnicity_binary
Round 2
ethnicity_binary
White
Not disclosed or unknown
Ethnicities other than white ethnicities
%
0%
10%
20%
30%
0%
20%
40%


Round 1
ethnicity_five
White
Other ethnic group
Not disclosed or unknown
Mixed or Multiple ethnic groups
Black, Black British, Caribbean or African
Asian or Asian British
0%
10%
20%
30%
40%
50%
Round 2
ethnicity_five
White
Other ethnic group
Not disclosed or unknown
Mixed or Multiple ethnic groups
Black, Black British, Caribbean or African
Asian or Asian British
0%
20%
40%
60%


Round 1
disability
Not disclosed or unknown
No Declared Disability
Declared Disability
0%
10%
20%
30%
Round 2
disability
0%
20%
40%



Source: UKRI

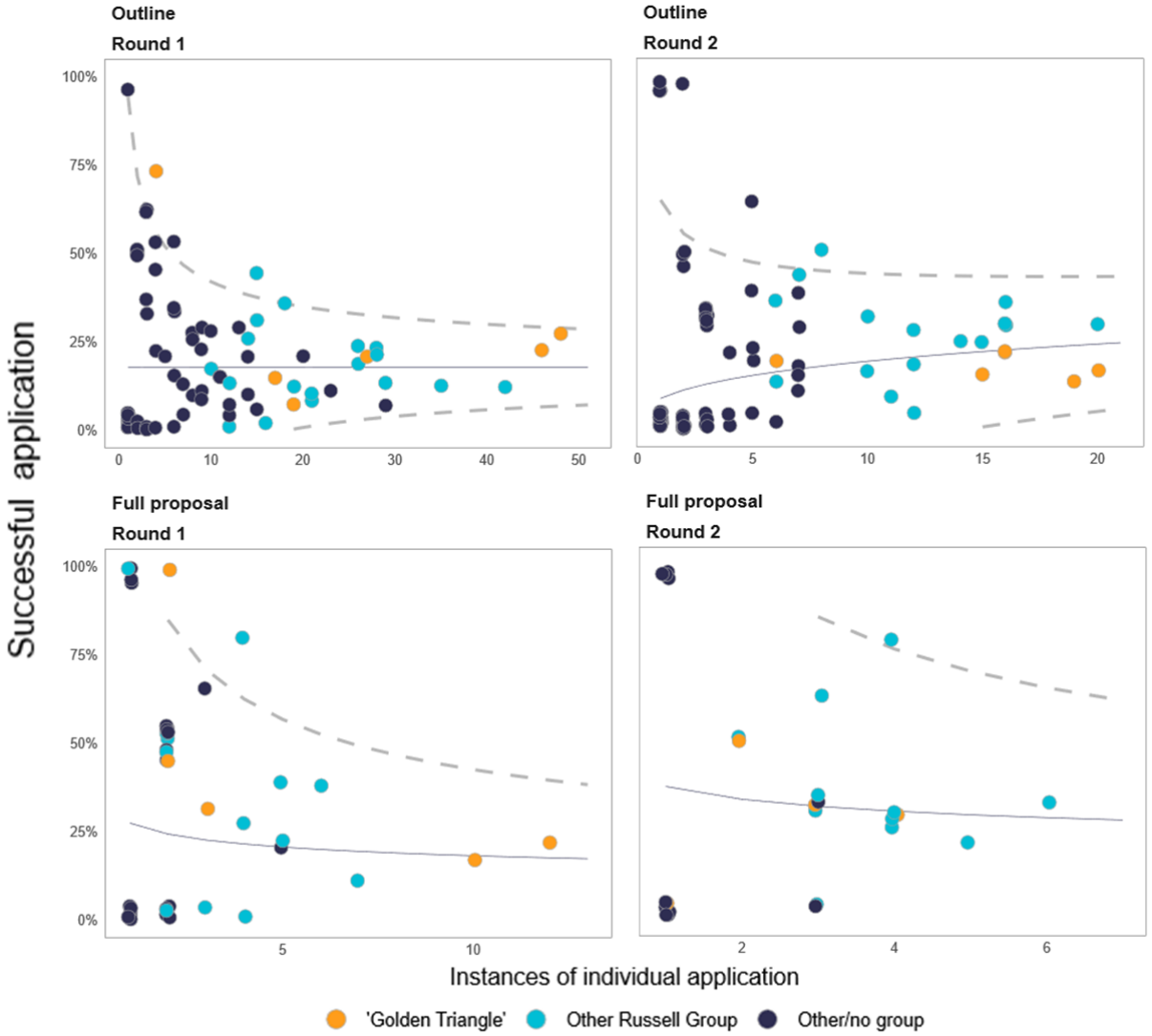
UKRI also analysed the diversity of organisations and teams.

Funnel plots (Figure 23a) based on organisation, with each circle representing one named organisation, show that for project leads there is evidence of a clear success-volume relationship based on organisation only for round 2 outline stage. However, this is not a strong relationship. Allowing for fluke results, there is no evidence of any organisation having an unexpectedly high or low success rate when project leads' organisations are considered.

For all applicants (Figure 23b), success-volume relationships are not visible. There is a bit more scatter, with some organisations having quite high invitation rates for their applicants at full stage (less so with outlines).

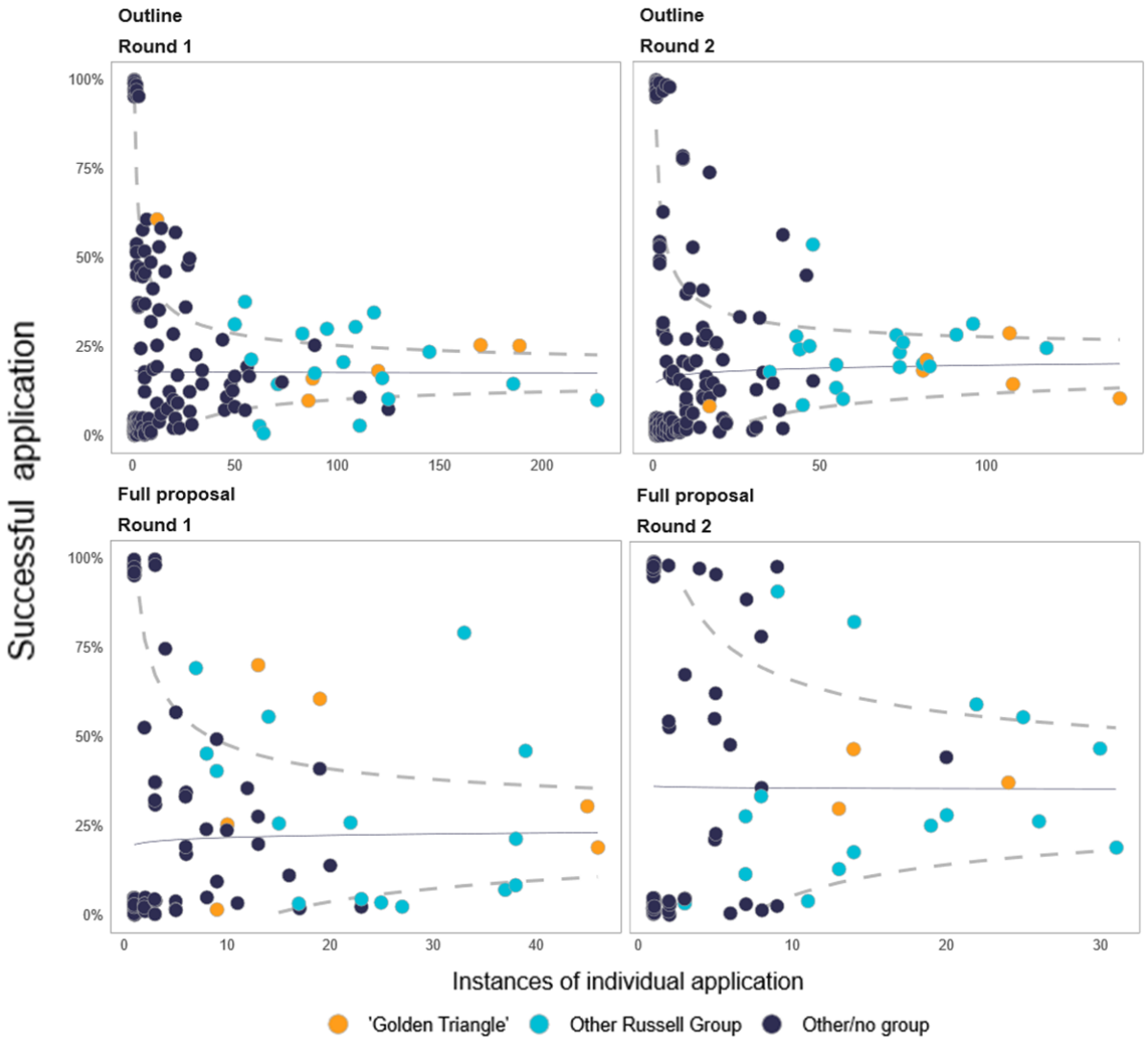
Figure 23 Invitation and success rates by organisation for project leads (a) and all applicants (b)

(a)



Source: UKRI

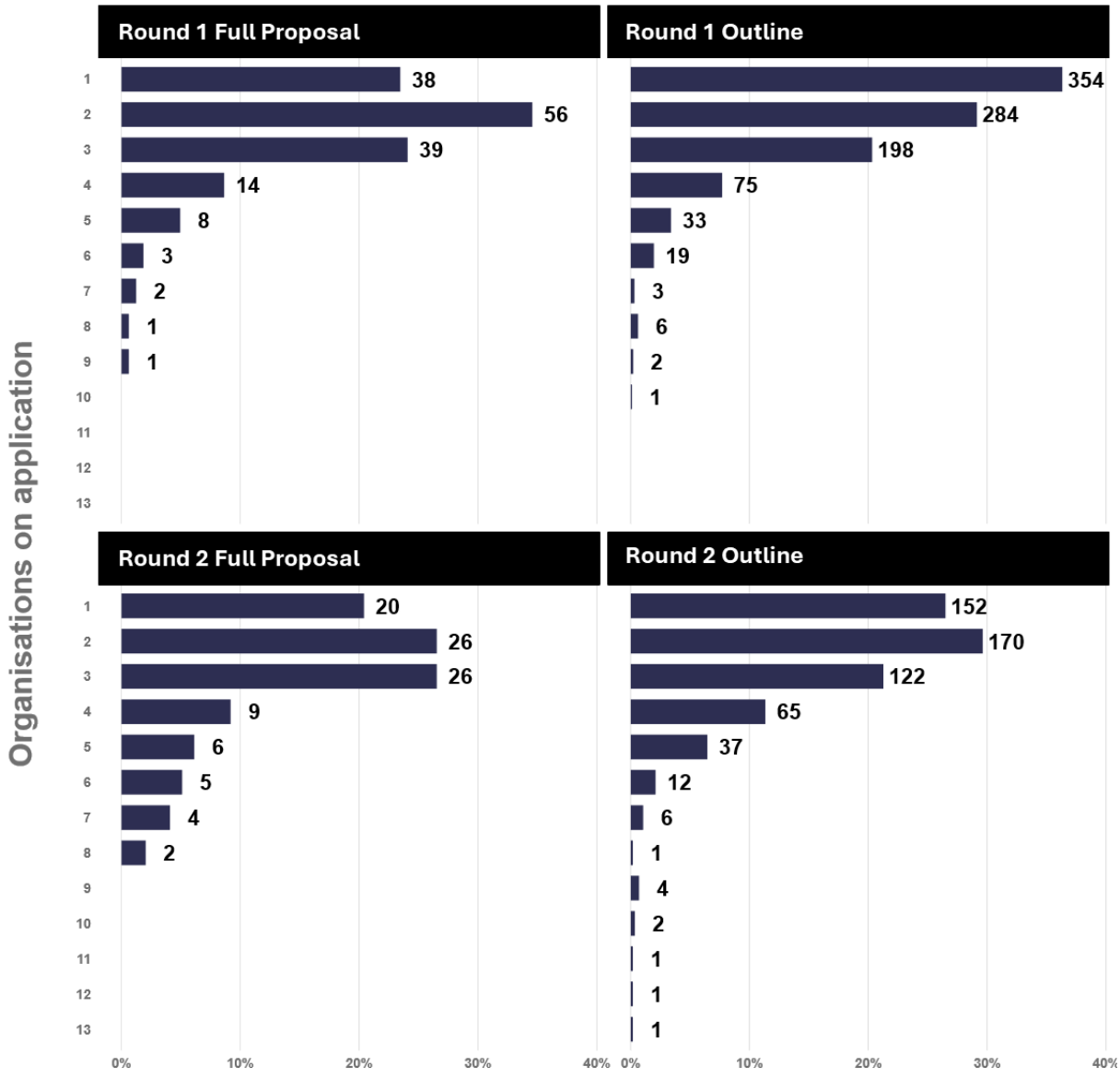
(b)



Source: UKRI

Analysis of the number of organisations on applications (ranging from 1 to 13) shows evidence suggesting full applications were more organisationally diverse than were their corresponding outlines (Figure 24).

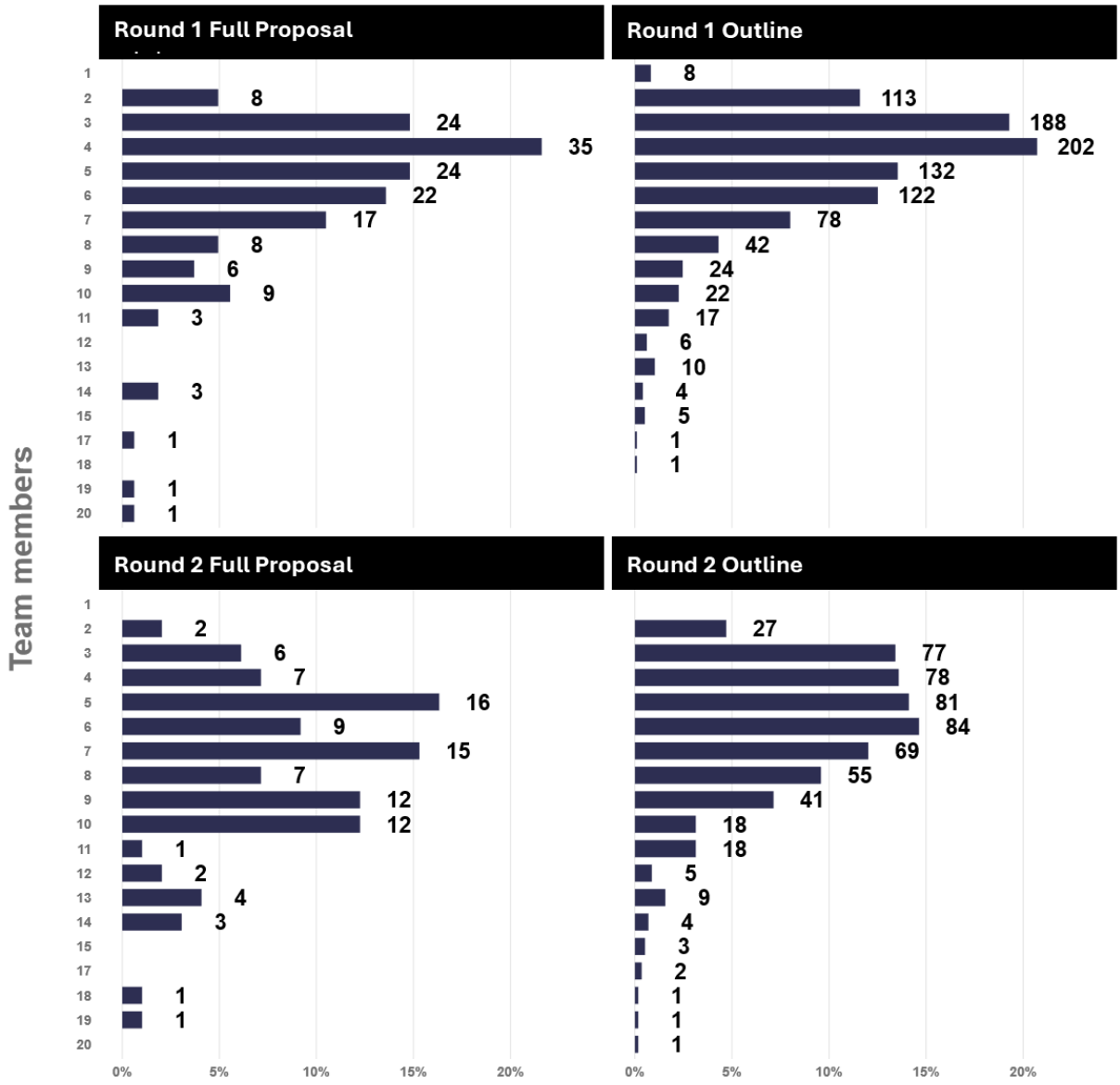
Figure 24 Number of organisations on applications



Source: UKRI

Full applications tended to have more team members than the corresponding outlines, especially in Round 2 (Figure 25).

Figure 25 Number of team members on applications



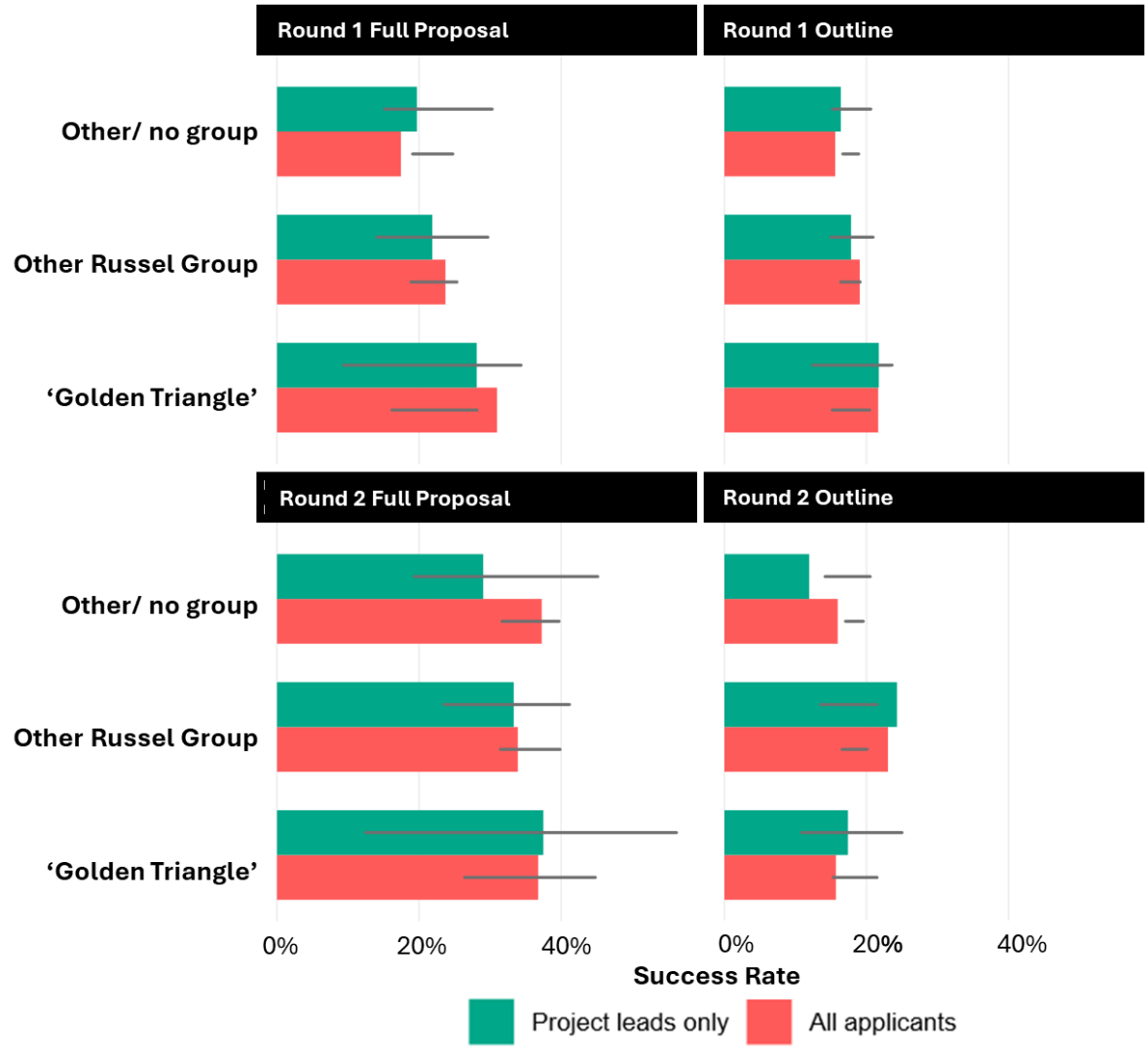
Source: UKRI

In Round 1 outline stage, IRC found that the diversity of organisations and number of collaborators involved in an application had a positive impact on interdisciplinarity scores. However, this pattern did not hold, and results were not statistically significant in Round 1 full stage and Round 2 outline and full stages. This supports, but does not conclusively prove, the possibility that assessors in the first round were using diversity as a proxy for interdisciplinarity. That this correlation disappeared in Round 2 suggests that relationship was either a fluke or that further training for assessors and changed scoring methods eliminated this behaviour.

When looking at research organisation 'mission groups', in round 1 applicants from organisations from 'Other/no group' had low success rates in both stages, and those from the Golden Triangle had high rates. In round 2, the 'Other/no group' had low success rates at outline stage while those from the broader Russell group had high success rates. Similar results were seen for project leads only.

Figure 26 Percentage of individuals associated with mission groups (a) and success rates (b) for project leads and all applicants by round and stage (a)

(b)



Source: UKRI

Further figures are available in Appendix C.2. These figures show that applications representing a single region were most common except in round 2 full application stage where applications covering two regions were most common. In general, full application stages tended to feature applications that were more regionally and organisationally diverse than their corresponding outlines, suggesting that the process was selecting in favour of this kind of diversity.

4 Impact on portfolio

4.1 The CRCRM pilot addresses a gap in the UK's IDR funding landscape

Summary findings

- Demand was strong across both rounds, and almost all interviewees confirmed the scheme fills a genuine gap that existing UKRI mechanisms do not cover.
- Without CRCRM, most applicants said they would have had to significantly rework their applications for other funders, wait for another opportunity, or simply not pursue the research, suggesting the scheme is additional rather than substituting existing funding routes.
- Feedback from UKRI councils indicates that CRCRM is funding IDR that would likely not be funded through their usual responsive-mode funding opportunities.

The strong demand for this scheme across the two rounds indicates that the CRCRM pilot addressed a gap in the UK's IDR funding landscape. In Round 1, 979 outline applications were received, of which 175 were invited to submit full applications and 36 awards were made. Strong demand was also observed in Round 2, where demand management was introduced. 576 outline applications were received, of which 100 were invited to submit full applications and 32 awards were made. UKRI collected data from the research organisations submitting applications on the number of applications coming forward in the internal sifts for Round 2. The data showed that 1060 applications were submitted, demonstrating that interest in the scheme remained high.

Almost all of the applicants and IAC members interviewed for the evaluation stated that CRCRM fills a gap in the IDR funding landscape. This is also consistent with findings from the applicants' surveys where CRCRM is seen to provide funding opportunities for IDR (90% agree or strongly agree) and breakthrough and disruptive ideas (84% agree or strongly agree) that would otherwise not be funded through established responsive-mode processes (n=1742 across both rounds and stages).

> "My research interests have historically fallen on the margins of council-specific calls and have perhaps been less competitive because of this. On face value, the CRCRM call is a perfect fit for the type of work I do." – Applicant

Many interviewees and survey respondents felt that it would be challenging, or not possible at all, to obtain funding from other UKRI schemes for their project. Without CRCRM, applicants reflected in both surveys and interviews

that they would have to rework their application to adapt it for the remit of other funders, such as Wellcome Trust, Leverhulme Trust, NIHR and the European Research Council.

When asked "in the absence of the CRCRM pilot scheme what would you have done with your interdisciplinary research application?", most applicants (n=782) said they would have repurposed their submission for another UKRI responsive mode funding scheme (32%), for submission to an alternative funder (29%), or for submission to a UKRI strategic funding scheme (9%). A significant number (23%) said they would simply wait for an alternative opportunity. However, many survey respondents (26 applicants in Round 1 and 25 applicants in Round 2) expressed concerns that such repurposing for another UKRI responsive-mode scheme would necessitate significant alterations to fit the specific disciplinary focus of these councils, potentially weakening the interdisciplinary nature of the research. Several respondents expressed concerns about the limitations of existing funding routes for interdisciplinary work, noting that proposals might be "rejected for going beyond the mainstream" or "falling betwixt and between" different councils. This sentiment was echoed by those who had previously experienced difficulties with interdisciplinary proposals, such as one participant who noted that "past experience of interdisciplinary proposals with these funders is not positive," highlighting a lack of appropriately skilled reviewers. Overall, the responses highlight a strong reliance on the CRCRM pilot scheme for supporting IDR, with many researchers expressing uncertainty about alternative funding options. The scheme's absence would likely lead to a fragmented approach to funding, potentially undermining the coherence and impact of interdisciplinary projects.

IRC analysis shows that, while some applications that were assessed as likely to be highly interdisciplinary were not advanced to the next stage or funded in the full stage, these decisions were based on concerns about (predominantly methodological) quality. Those interested in advancing more high-quality interdisciplinary projects may consider identifying those applications with potential and providing workshops or training to assist scholars unused to putting together interdisciplinary applications to design more effective projects.

Comments from individual UKRI councils on the portfolio of projects funded in Round 1 and 2 of CRCRM indicate that quite a few of the successful applications that span their council remit would not have been funded through their usual responsive-mode routes. There was thematic alignment between these applications and their council remit, and the interdisciplinary remit overlap with other councils was largely per what they usually observe. However, some councils also suggested that they would have found it hard to

identify the right assessors or ensure fair and robust assessment of the applications that were eventually successful.

Several funded projects illustrate forms of interdisciplinarity that would be unlikely to emerge through single-council responsive mode schemes. For example, 'Navigating the Neural Frontier: Embedding Ethics and Epistemology in Neurotech', 'NeuroCognitive Shield', and 'Economic Early-Warning Signals of Accelerated Climate Change and Earth-System Tipping-Points' all address social or economic challenges through deeply integrated, cross-disciplinary approaches. Similarly, 'MoonRISE' was highlighted as particularly novel in bringing together STFC and ESRC perspectives in a way not typically observed in standard funding routes.

Other projects demonstrate interdisciplinarity as a source of novelty in itself. 'Wreck Assess' was widely regarded as an example of "true" interdisciplinarity and notably distinct from typical NERC applications; while its core disciplinary elements may appear less novel in isolation, particularly those that fall within the EPSRC portfolio, the strength lies in the integration of approaches, which would likely limit its competitiveness within single-council schemes.

The portfolio also includes projects that combine unconventional disciplinary perspectives, such as those integrating Indigenous and historical knowledge systems with environmental science, or those linking manuscript analysis with AI and plant or medical research. These projects demonstrate how arts and humanities methods can broaden the evidence base and enhance impact in addressing complex societal challenges, even beyond current AHRC strategic priorities.

Finally, the AI-related funded projects are notably interdisciplinary, incorporating diverse methodologies that would be difficult to assess within standard panel structures. Examples include 'BeefTwin', which applies AI to create a digital twin of beef farming systems; 'JED-AIs: Justice, Energy, Demand Flexibility and AI for Sustainability', which blends technical and social science approaches; and 'Historical horticulture: a quest for future plant propagation solutions', which develops AI-powered tools trained on historical sources. Collectively, these examples highlight the scheme's success in attracting AI-focused proposals that extend beyond conventional disciplinary boundaries typical of investigator-led routes.

4.2 Benchmarking to other UKRI schemes

Summary findings

- Compared to benchmark programmes (GCRF, CCRA, Physics of Life), CRCRM applications showed statistically significantly higher levels of interdisciplinarity and novelty, with Round 2 awards showing roughly twice the odds of being classified as high novelty.
- Importantly, the selection process appeared to amplify this effect. Awarded CRCRM projects were even more interdisciplinary than CRCRM applicants as a whole, suggesting the assessment process successfully identified and prioritised the most genuinely interdisciplinary applications.

To identify the added value from CRCRM, we compared CRCRM applications and awards to projects awarded under a group of benchmark programmes. To ensure we were able to provide a robust comparison against different configurations of IDR and assessment processes, we chose three schemes to cover a broad set of disciplinary mixes and both strategic and responsive mode schemes as benchmarks. These were:

1. A strategic collaborative research scheme: that is, Physics of Life (SPF), which was a top-down IDR funding scheme that invites applications for interdisciplinary projects specifically combining the biological/biomedical and physics, n = 25

2. A strategic challenge-led scheme: that is, Global Challenges Research Fund (GCRF), which was a top-down IDR funding scheme aimed at funding interdisciplinary collaborations to address key global challenges (not necessarily specifying the disciplinary mix), n = 1018

3. A responsive mode scheme: that is, projects funded under the cross-council remit agreement, which is a policy for funding bottom-up, applicant-led projects covering remits from multiple councils, n = 51

We compared the novelty and interdisciplinarity of applications and awards for the CRCRM pilot with projects funded by the benchmark programmes. For this we computed interdisciplinarity and novelty indicators using advanced natural language processing tools.

We computed the following **interdisciplinarity indicators**:

- **Herfindahl-Hirschman Index (HHI)**: HHI is a standard metric use to measure concentration (typically market concentration). A lower HHI reflects a more evenly distributed focus across multiple fields, reflecting higher interdisciplinarity.

- **Entropy**: Entropy measures the degree of diversity or "spread" in a distribution and thus, the entropy value is higher when it draws on a balanced range of distinct fields.[17]

It should be noted that the HHI and Entropy indicators are, like most other available metrics to measure interdisciplinarity, not able to differentiate between multidisciplinarity and interdisciplinarity. Consequently, the conclusions we draw below combine the findings from multiple indicators (including novelty, and the SIFT score presented in Section 3.2.2).

The **novelty indicator**[18] stems from a machine learning model that learns from the novelty patterns in scientific articles that led to Nobel prizes (papers renowned for their high levels of novelty). The model's features comprise solely the text of an article and the text of the closest prior art (that is, prior existent publications), measuring the originality of each scientific contribution. In an econometric framework controlling for citations, journal venues, publication year and topics, the model remains a robust predictor of novelty, effectively disentangling it from quality and other scientific attributes.

Statistical comparisons of the three indicators (Table 5) show a statistically significant difference between the means of CRCRM applications and benchmark projects. These findings ($p < 0.001$ for HHI and entropy across both rounds; $p = 0.009$ and 0.072 for novelty in Round 1 and 2 respectively) provide **early evidence that CRCRM applications present higher interdisciplinarity and novelty levels than other relevant projects from distinct funding mechanisms within UKRI**. The results from deploying the "Semantic fixed-effects" model[19] confirm the initial aggregate differences, showing that CRCRM applications have statistically significant higher levels of novelty and interdisciplinarity than benchmark projects after controlling for within-group semantic proximity. While, like most currently used interdisciplinarity metrics, HHI and entropy do not distinguish between multi- and inter-disciplinarity, the results clearly show more diversity in the CRCRM applications compared to benchmark projects.

Analysis showed that a CRCRM Round 1 application has ~1.55 times higher odds of being classified among the higher-novelty group (according to the Nobel_Novelty criterion) relative to comparable benchmark projects within

[17] Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of informetrics*, *5*(1), 87-100.

[18] Machado, D. (2024). "The meaning of novelty: predicting novelty with text embeddings". Presented at WOEP FULL4, TURIN, 2024; LIS-Bibliometrics, Brighton, 2024; and REvaluation Conference, Vienna, 2024; and upcoming at the ATLC25 and Global TechMining Conferences, Atlanta, 2025.

[19] Abbonato, D., Bianchini, S., Gargiulo, F., & Venturini, T. (2024). Interdisciplinary research in artificial intelligence: Lessons from COVID-19. *Quantitative Science Studies*, *5*(4), 922-935.

the same semantic cluster.[20] For Round 2 applications, the odds ratio is higher and is associated with ~2.32 times higher odds of being classified among the higher-novelty group. These results require cautious interpretation, as they capture variation only among projects and applications that are identified as similar to allow fair comparison between projects and applications in similar fields. That is, we are comparing subsets of the dataset rather than the entire dataset. Moreover, note that the benchmark data has a strong predominance of GCRF projects and the CRCRM applications include more unsuccessful outline applications comparatively.

The results on interdisciplinarity and novelty are interesting since we would have expected that funded benchmark projects which have been selected on the basis of quality would at least be more novel and possibly more interdisciplinary compared to CRCRM applications. However, we do not know to what extent the benchmark projects were assessed and selected on the basis of interdisciplinarity and how this was interpreted in the assessment process, so it is difficult to comment on this aspect.

Table 5 Statistical (t-test) comparisons of CRCRM application indicators against benchmark projects

(a) Round 1

	CRCRM applications	Benchmark projects	Mean Difference	p-value
HHI	0.343	0.390	-0.048	0.000
entropy	1.391	1.293	0.098	0.000
nobel_novelty	0.026	0.011	0.015	0.009

(a) Round 2

	CRCRM applications	Benchmark projects	Mean Difference	p-value
HHI	0.312	0.390	-0.078	0.000
entropy	1.435	1.293	0.142	0.000
nobel_novelty	0.045	0.028	0.017	0.072

Source: Technopolis-group based on CRCRM application data, Gateway to research and Semantic Scholar

[20] The odds ratio of approximately 1.55 for the CRCRM dummy in the Nobel_Novelty model is derived by exponentiating the estimated log-odds coefficient (0.438). The coefficient's p-value ($p<0.05$) confirms that the effect—that is, CRCRM applications having higher odds of classification among the higher-novelty group—is statistically significant. This result means we can be confident that this odds ratio is different from 1.0 (no effect) at the 5% significance level.

The aggregate indicators for CRCRM awards and benchmark programmes are presented below (see Table 6). Considering the small sample size for some programmes, it was not possible to do analyses of statistical significance. However, it is particularly worth noting that CCRA projects have lower diversity of fields (and thus lower interdisciplinarity) compared to CRCRM awards, indicating that the CRCRM scheme has clear added value and supports IDR that isn't funded through usual responsive-mode mechanisms.

Table 6 Aggregate means of key indicators per funding programme

Indicator	Awarded CRCRM (Round 1)	Awarded CRCRM (Round 2)	GCRF	CCRA	Physics of Life
HHI	0.309	0.292	0.390	0.435	0.300
entropy	1.462	1.507	1.297	1.144	1.451
N (sample size)	35	32	1018	51	25

Source: Technopolis-group based on CRCRM application data, Gateway to research and Semantic Scholar. Note: Our prior art construction process dropped one application in Round 1, leaving us with 35 successful CRCRM applications.

Successful CRCRM projects present lower concentration levels of fields measured through HHI and higher field diversity measured through the entropy indicator, representing greater interdisciplinarity. These differences are statistically significant, with p-values of 0.02 and 0.045 for Round 1 awards and of 0.007 and 0.015 for Round 2 awards, thus evidencing that successful CRCRM applications present higher levels of interdisciplinarity than benchmark projects.

Finally, the semantic fixed-effects analysis of CRCRM awards vs all benchmark projects (Table 9 in Appendix C.1) shows that for both HHI and the entropy indicator, the strength of the relationship is more accentuated for successful CRCRM applications than in the previous analysis considering all CRCRM applications, suggesting a potential enabling mechanism in the selection process promoting interdisciplinarity. That is, the explicit focus on selecting for interdisciplinarity in terms of integration, reciprocal benefits and co-creation/co-design in the CRCRM assessment process may explain the difference between CRCRM awards and benchmark projects.

4.3 Cross-disciplinary research challenges being tackled

Summary findings

- Environmental challenges and health and wellbeing dominated as the primary problem areas across both rounds and at all application stages.

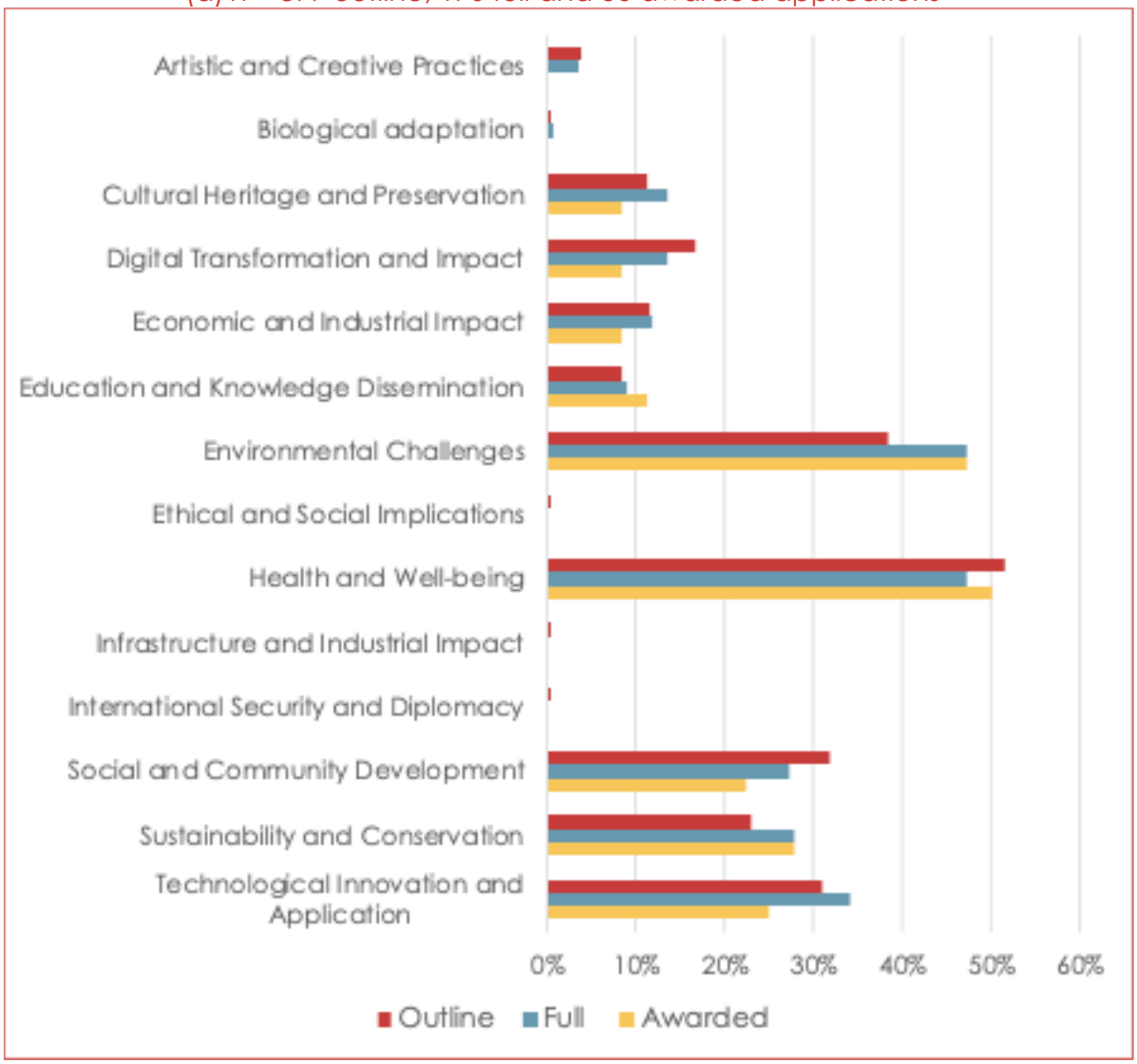
- There was a notable methodological shift between rounds: technical/biomedical methods dominated Round 1, while participatory approaches and co-design were more prominent in Round 2, possibly reflecting a better-informed applicant pool.

4.3.1 Key problems addressed

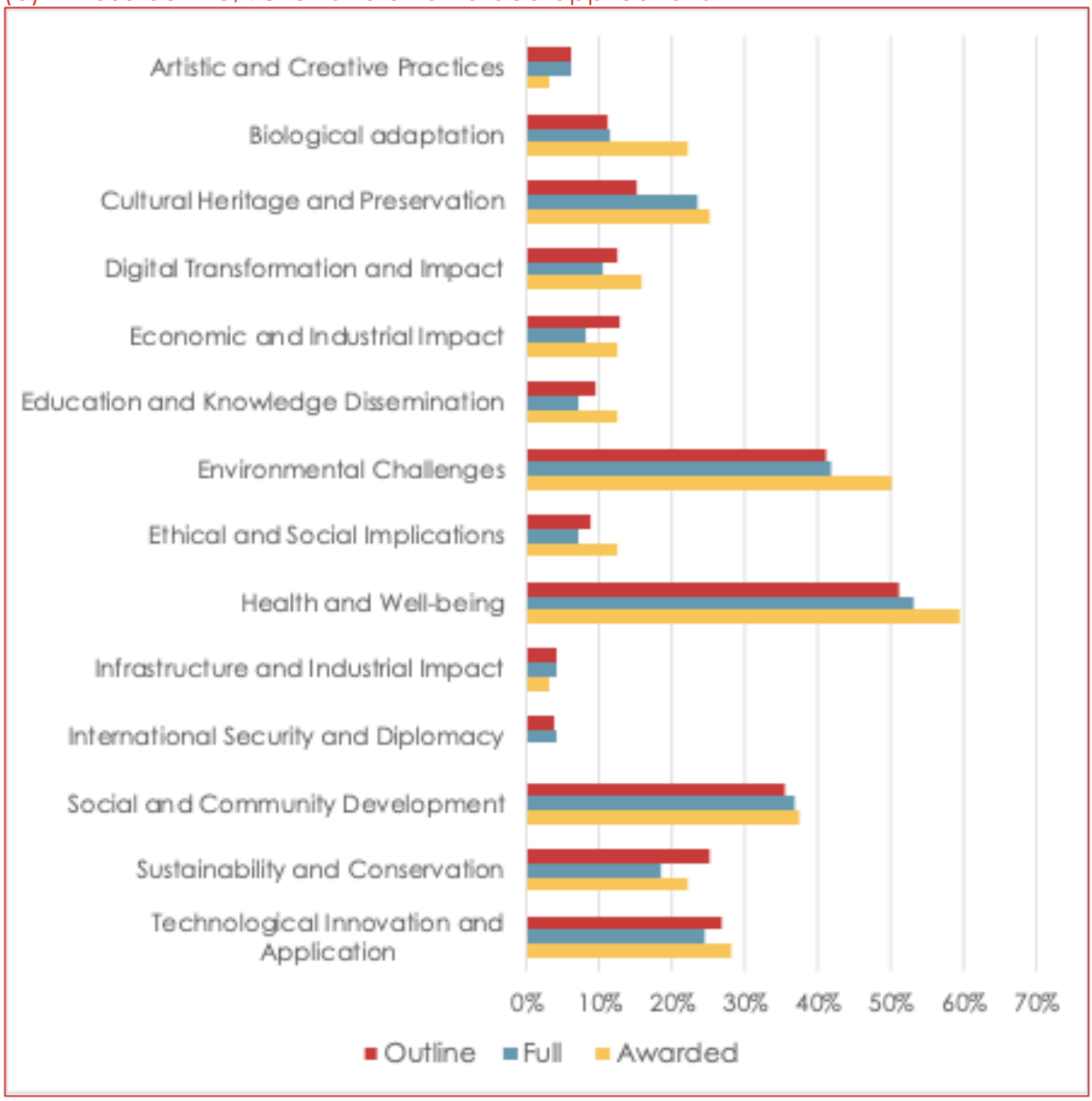
Environmental challenges and health and well-being were the most common cross-disciplinary research problems targeted in CRCRM applications (Figure 27). Societal challenges in these areas emerge as the focus in outline and full applications as well as the awards. The pattern is broadly similar between Round 1 and Round 2.

Figure 27 Bar charts showing the key problem categories addressed in outline, full and awarded applications in Round 1 (a) and Round 2 (b)

(a) n = 899 outline, 170 full and 36 awarded applications



(b) n = 535 outline, 98 full and 32 awarded applications



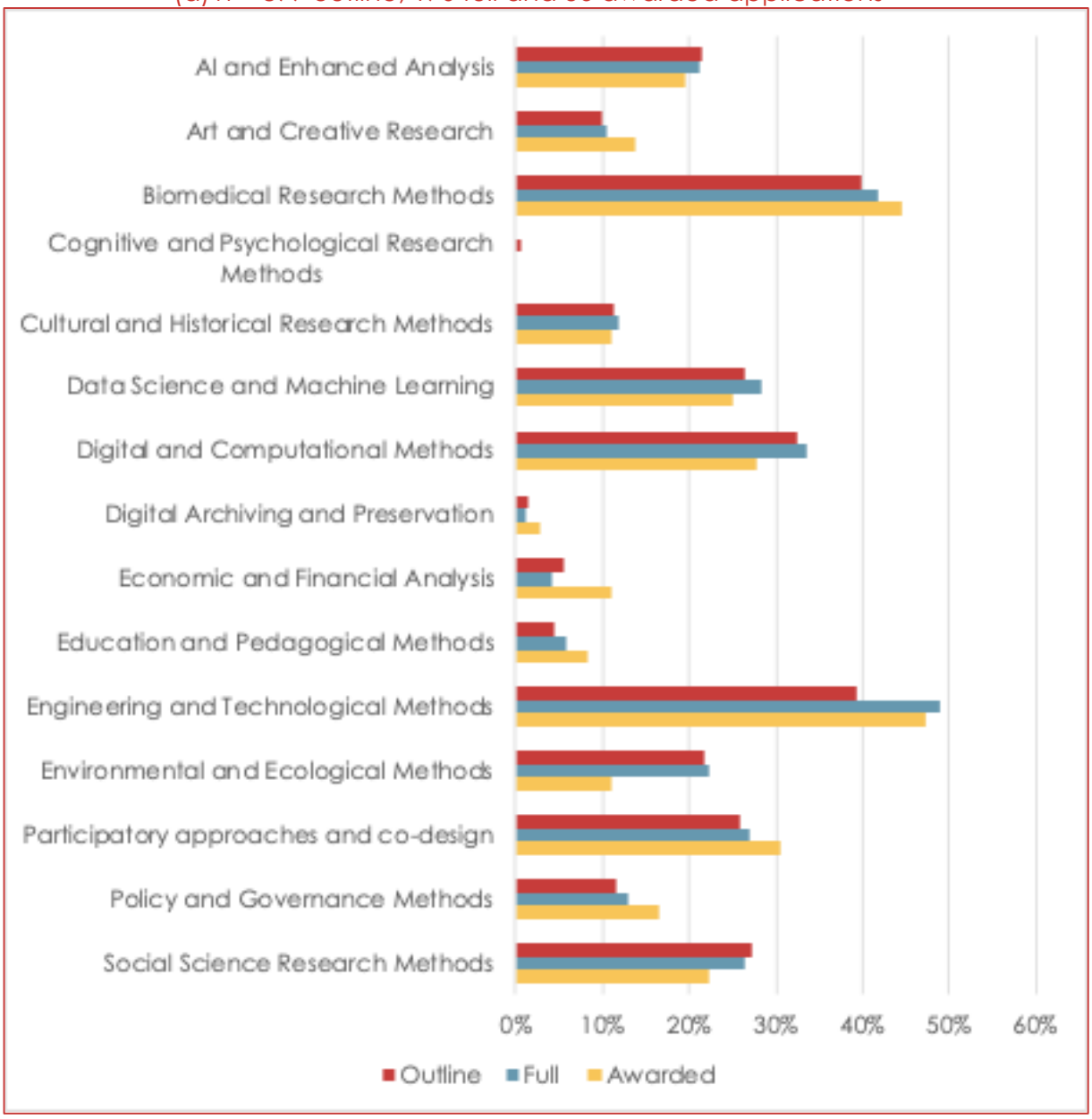
Source: Technopolis

4.3.2 Main types of methods used

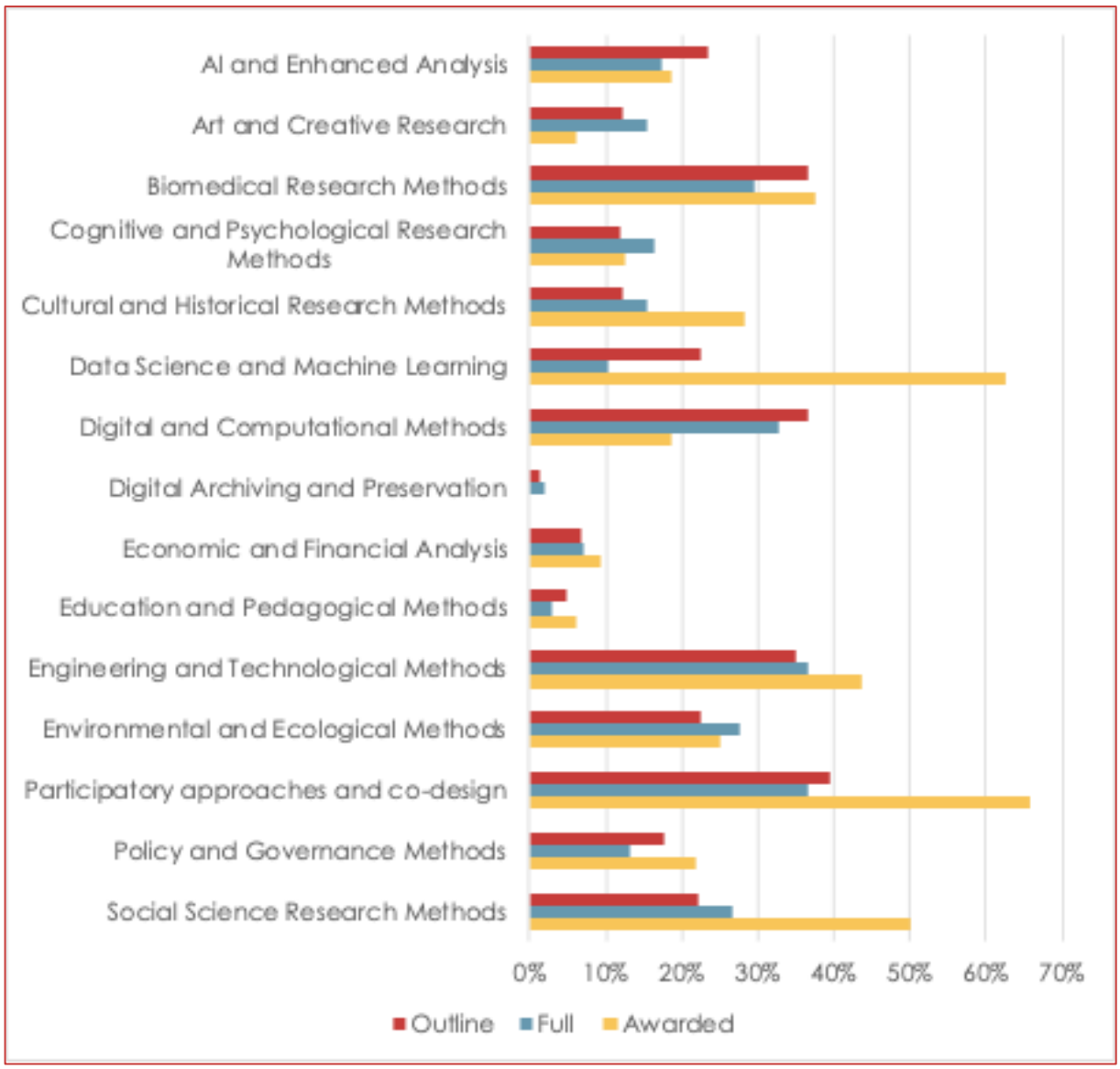
Technical methods, for example biomedical and engineering and technology research methods, dominated the applications and awards in Round 1. However, participatory approaches and co-design featured prominently among methods covered in Round 2 applications alongside the technical methods such as biomedical, engineering and technological and digital and computational research methods (Figure 28). Interestingly, applications involving data science and machine learning, and social science research methods seem to come through more prominently among awards in Round 2 alongside participatory methods compared to the outline and full stage applications.

Figure 28 Bar charts showing the key methods used in outline, full and awarded applications in Round 1 (a) and Round 2 (b)

(a) n = 899 outline, 170 full and 36 awarded applications



(b) n = 535 outline, 98 full and 32 awarded applications



Source: Technopolis

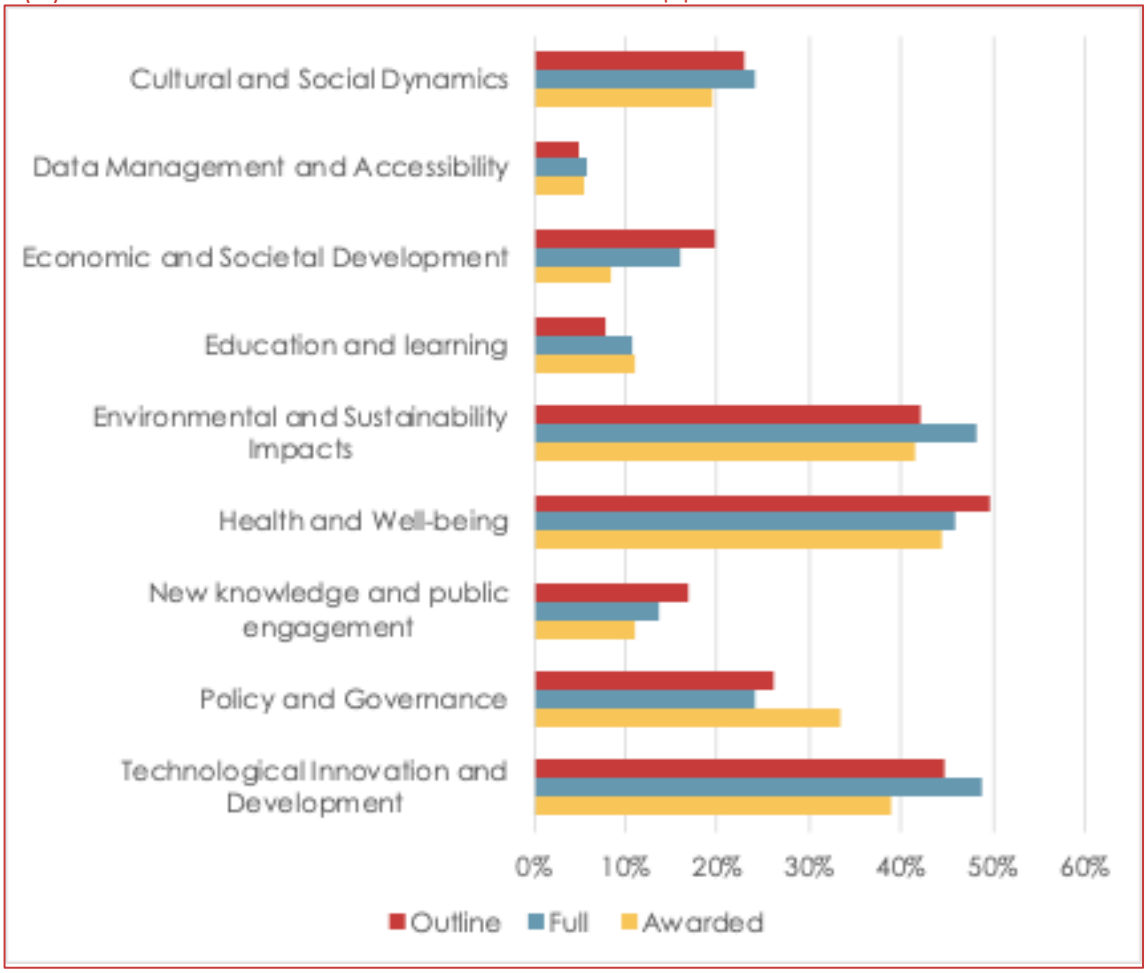
4.3.3 Main impact areas targeted

In line with the main problem areas addressed, the key impact areas targeted in the applications were health and wellbeing and environmental and sustainability impacts (Figure 29). Technological impacts are also a main target in Round 1 and Round 2 applications. The proportion of targeted impact areas remains broadly similar across the outline, full and awarded applications. This is the case for most impact areas across both rounds with the exception of new knowledge and public engagement impact which appears to be enriched in the Round 2 awards compared to the outline and full applications. Across the individual interfaces (see Technical Annex G),

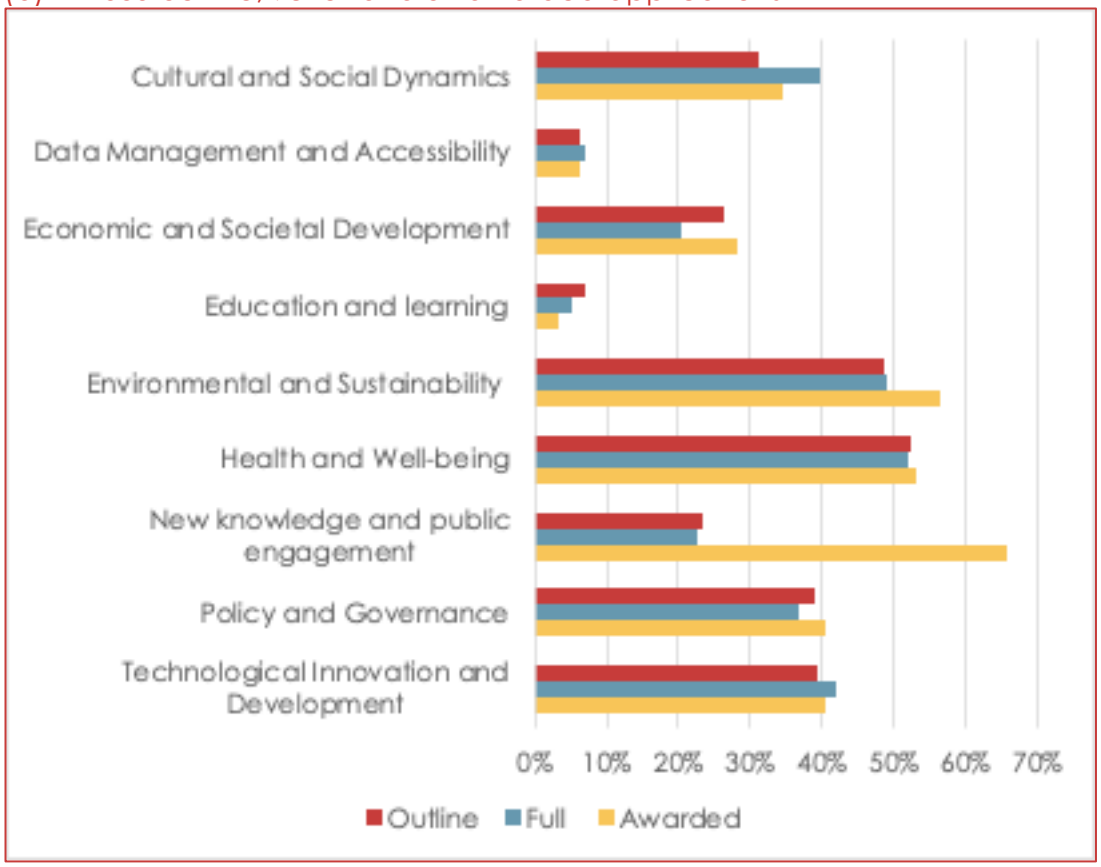
applications were frequently targeting impacts that would be most commonly associated with the main disciplines represented in the remit area. For instance, applications in BIOM interfaces targeted health and well-being impacts, while EEB interfaces targeted environmental impacts.

Figure 29 Bar charts showing the main impact areas targeted in outline, full and awarded applications in Round 1 (a) and Round 2 (b)

(a) n = 899 outline, 170 full and 36 awarded applications



(b) n = 535 outline, 98 full and 32 awarded applications



Source: Technopolis

4.4 Enablers and barriers

The CRCRM pilot scheme met its objectives. The key enablers and barriers that affected CRCRM's progress towards meeting its objectives are described in this section.

Summary findings

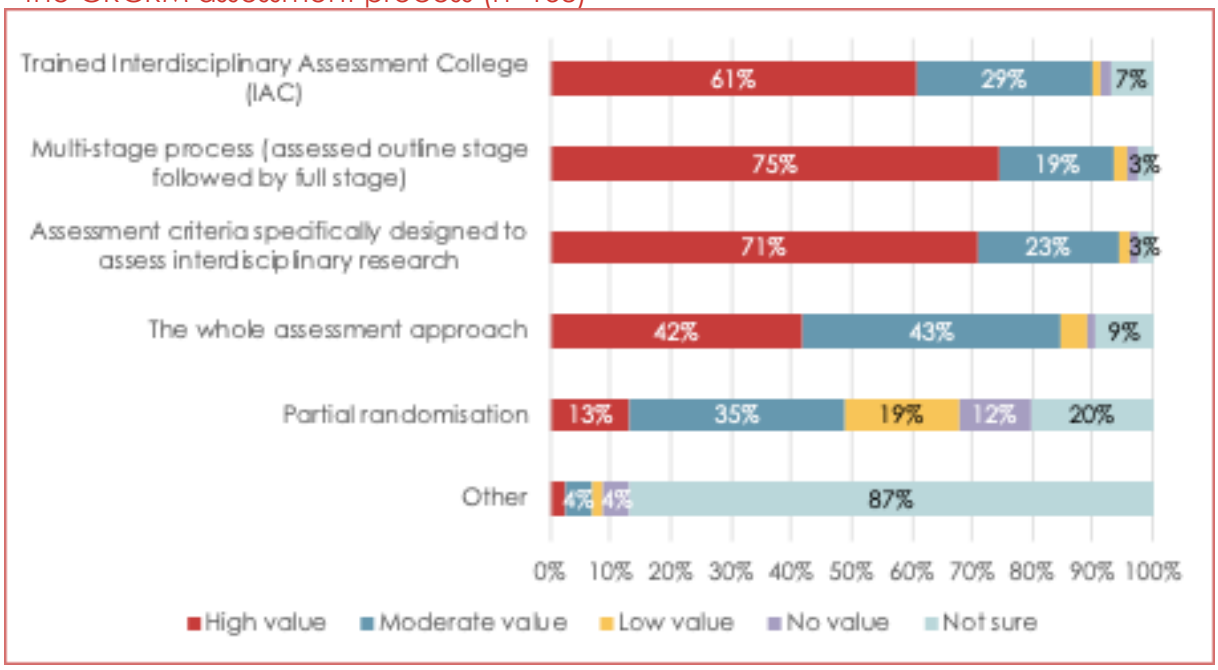
- Stakeholders identified the trained IAC, the multi-stage application process, and assessment criteria specifically designed for IDR as the elements adding most value.
- The scheme's explicit focus on interdisciplinarity was seen as genuinely 'liberating', allowing researchers to pursue speculative, high-risk ideas across multiple fields in a way that other funding mechanisms do not permit.
- Deliberation is also a key enabler – the strongly collaborative assessment process, including panel discussions and combined assessor feedback to lead applicants bringing together different interdisciplinary expertise and perspectives, made for fairer and more robust assessment.

- The two-year funding period was widely seen as too short to build meaningful interdisciplinary collaborations, particularly for novel or high-risk projects, with many suggesting even a no-cost extension to three years would make a material difference.
- While the scheme was praised for beginning to shift research culture towards valuing interdisciplinarity, participants noted that sustaining this requires broader institutional changes (that are beyond the control of this type of scheme).

4.4.1 Enablers

Stakeholders saw the trained IAC, multi-stage process and assessment criteria designed specifically for IDR as adding the most value to the assessment process. This view is consistent with views of Round 1 and 2 applicants across the surveys. In both rounds, the multi-stage process was most frequently identified as the most valuable aspect of the assessment process. However, Round 1 applicants ranked the IAC ahead of the assessment criteria whereas Round 2 applicants placed greater value on the assessment criteria.

Figure 30 Round 2 full stage applicants' views on the aspects that add value to the CRCRM assessment process (n=158)



Source: Technopolis

The scheme's explicit emphasis on interdisciplinarity was widely regarded as both a strength and a novel approach to promoting collaboration. Successful applicants emphasised the scheme's ability to facilitate collaborations across multiple fields, including Earth sciences, social sciences, cognitive neuroscience, and the arts. This interdisciplinarity enabled researchers to

experiment with novel methodologies, such as using theatrical techniques to address scientific biases and disagreements. Furthermore, the scheme's emphasis on reciprocal learning and speculative, high-risk projects gave participants the freedom to pursue innovative and transformative ideas.

> "I have never worked in such an egalitarian way between disciplines when drafting the application, in previous applications one discipline has always been dominant." - Applicant

Evidence from panel scores across Round 1 and 2 as well as feedback from IAC that panel meetings allow for fairer and more robust assessment indicates that deliberation is a key enabler. Panel meetings that bring together diverse interdisciplinary expertise and perspectives for assessing IDR applications produce more robust and fairer assessments compared to moderation panels with a smaller number of individuals that only discuss divergent scores without input from those doing the scoring. Discussion and consensus building in panel meetings allows greater differentiation between applications based on assessment criteria as observed from the greater range and higher standard deviation for consensus scores arrived at after deliberation compared to averages of individual IM scores. Moreover, deliberation reduces the risk of 'death by a thousand cuts' as observed when individual reviewers score IDR applications in isolation from the perspective of their own discipline or prior experience.

4.4.2 Barriers

Many applicants interviewed at different stages found the two-year funding period insufficient for producing meaningful collaborations, particularly in novel and high-risk interdisciplinary research areas. The consensus is that longer-term funding is critical for sustaining the initial efforts, ensuring project development and research capacity building in new areas. Some suggested that even a no-cost extension to allow projects to be carried out over three years instead of two would be helpful to better implement projects while navigating complex collaborations.

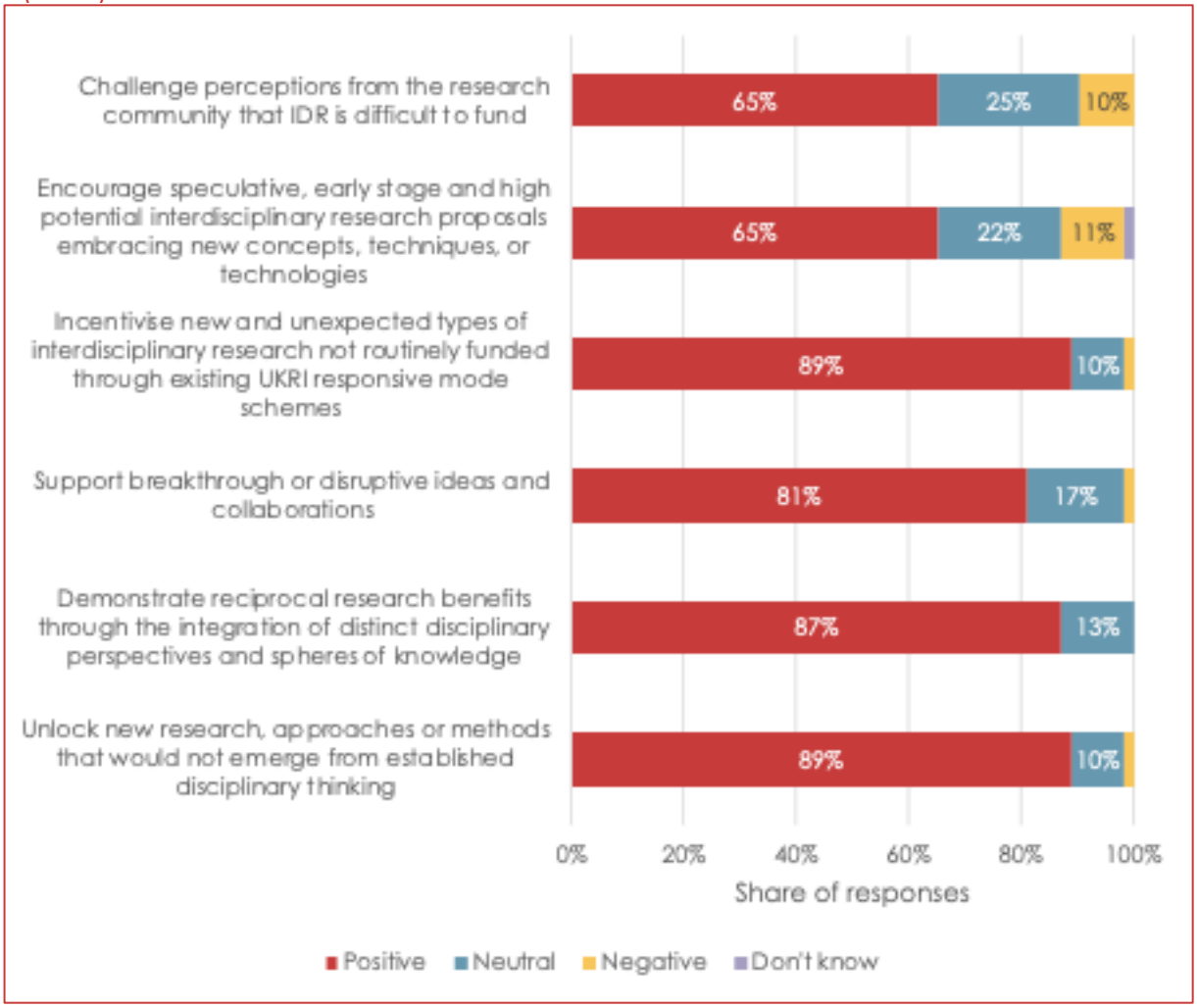
4.5 Impact on the research community

Survey evidence suggests that the CRCRM is helping to shape IDR practice and its assessment processes can support innovative and potentially transformative research approaches. The scheme was praised in the surveys and interviews for attempting to change the culture of research collaboration by prioritising interdisciplinarity over outcomes in particular challenge areas as seen in other UKRI IDR funding opportunities. The use of a trained college of assessors, as well as a focus on collaboration as a key component of grant applications, were deemed promising innovations to support the change in

research culture. However, participants recognised that achieving this cultural shift requires institutional changes, such as changes to university assessments and incentives.

> "Many universities want interdisciplinary research to thrive, but this does mean breaking down disciplinary siloes, which govern how institutions are organised… I think this call and UKRI are also helping nudge similar approaches across the sector." - Rover

Figure 31 Stakeholder perceptions about the scheme's assessment process (n=63).



Source: Technopolis

IAC members were overwhelmingly positive about the scheme's assessment process (see Figure 31). They believed it would support innovative IDR by enabling novel approaches, methods, and collaborations that might not emerge through traditional disciplinary routes. They also felt it would help

address barriers to funding IDR by encouraging speculative, high-potential research and demonstrating the value of integrating diverse perspectives.

> "I was struck by the more risky applications which really combined unusual ideas. I think we need to be careful to encourage risk and not become risk averse." – Panel Chair

Several interviewees felt the CRCRM scheme is already influencing perceptions of IDR by raising the profile of IDR projects and helping to recognise the benefits of interdisciplinary collaborations. The pilot has helped expose researchers, both as IAC members and as applicants, to a different definition and approach to funding and conducting IDR.

Over half of Round 2 full stage applicants reported that CRCRM changed their approaches to IDR (n = 158, 55%), suggesting that the scheme is helping to shape the way researchers design and conduct IDR.

> " I think having this opportunity itself is a good way to promote interdisciplinary research as it gives a focus for groups coming together to discuss what could be done. This is something that might not happen otherwise as everyone is busy. Even if bids fail at the first stage links have been made and future work might develop." - Applicant

Other interviewees believe it is too early to understand the effects of the CRCRM scheme on the wider practice of IDR in the UK.

5 Conclusions and Recommendations

The overall evaluation findings suggest that the CRCRM pilot scheme met all its objectives. Analysis of the applications and the portfolio of funded projects and opinions of applicants, assessors and UKRI councils indicate that the scheme funded novel IDR ideas that would not have been funded through other UKRI responsive-mode funding opportunities, including breakthrough and disruptive ideas as well as projects incorporating new concepts, approaches and methodologies. The pilot also successfully tested and refined IDR assessment processes and highlighted the demand within the research community for a dedicated responsive-mode IDR funding opportunity at UKRI, outside the council structures.

5.1 Conclusions

1. **CRCRM provided a unique opportunity to support research that previously had not been supported by UKRI funding mechanisms**

   The CRCRM pilot scheme represents a strategic effort by UKRI to foster IDR by providing a dedicated funding mechanism that addresses the unique challenges faced by IDR applications: (a) fitting within the remit of individual UKRI councils and (b) assessment through the perspective of individual disciplines.

   There is a wide consensus that the CRCRM scheme addressed a gap in the UK funding landscape and supported research ideas and a research community that previously had not always been supported well by UKRI. The scheme design allowed applicants to authentically develop their applications without having to write it to align with the remit of a particular research council or strategic funding opportunity. Indeed, comments from UKRI councils indicate that several of the successful CRCRM applications would not have been funded through the usual responsive-mode routes. Thus, the scheme, with its innovative processes, such as the trained IAC and panel structures that enable constructive discussions and consensus building is recognised as a 'test bed' for developing best practice in funding responsive-mode IDR.

   The large number of applications received highlights that there is a high demand for research that does not fit the UKRI disciplinary council framework.

   > "It was an application we've been trying to do for about 5 years. We'd been writing and planning, but we just never got around to submitting it because it just didn't fit anywhere. I think our application is what the scheme was

for; something that just couldn't work in another responsive mode." – Applicant

The evaluation findings also suggest that CRCRM has led to increased confidence among IAC members in assessing IDR and also among applicants to propose highly novel IDR ideas to UKRI for assessment.

2. **CRCRM and its processes are high-quality, efficient and effective in the assessment of IDR applications**

   Most stakeholders found the CRCRM assessment processes to be high quality and efficient in surveys and interviews. In particular, stakeholders found the processes implemented in Round 2, which built on learnings from Round 1, to be more standardised, effective and efficient compared to the previous round.

   The pilot effectively addressed the 'double jeopardy' of research applications falling in between the gaps of distinct research councils' remits, which is a particular risk for interdisciplinary applicant led applications. It did this through testing and developing different approaches to assess IDR applications in response to feedback gathered over the course of the pilot, including:

- convening and training an expert college of interdisciplinary assessors
- testing different assessment criteria and scoring systems
- iteratively building a bespoke and clear definition of IDR
- developing clear guidance and webinars for applicants to develop strong interdisciplinary applications
- developing innovative approaches to the assessment of interdisciplinary applications to manage divergent opinions through deliberation at all stages of the assessment process

The key approaches and innovations that added value to the CRCRM assessment process are summarised below.

The **IAC added value** to the assessment processes by providing suitable breadth of expertise in panels with training specifically focused on assessing IDR applications. After the two CRCRM rounds, the IAC represents an experienced resource for IDR assessment in the UK. An interesting finding from IRC's work is that IAC members whose expertise spans both a STEM and AHSS research council tend to be tougher graders, possibly because they are experienced in IDR and hence more stringent in assessing levels of integration.

Deliberation was embedded throughout assessment in terms of panel discussions and providing integrated applicant feedback. Evidence from

panel scores across Round 1 and 2 as well as feedback from IAC that panel meetings allow for fairer and more robust assessment. Panel meetings helped to bring together different interdisciplinary expertise and perspectives for assessing applications. Discussion and consensus building in panel meetings allowed greater differentiation between applications based on assessment criteria as observed from the greater range and higher standard deviation for consensus scores arrived at after deliberation compared to averages of individual IM scores.

Single council responsive mode schemes use external written review which can disadvantage IDR applications, where multiple separate assessments, may not fully capture the application as an integrated whole, especially if a reviewer does not have the expertise that covers the breadth of the proposed research and their review is through the lens of a single discipline. IDR benefits from discussion and triangulation of perspectives from different disciplinary and interdisciplinary expertise and experiences.

Deliberation was also a part of **integrated applicant feedback**. The use of discussion boards to assemble the perspectives of all reviewers to come up with a clear set of clarification questions for applicants, was seen as an innovative process to mitigate conflicting critiques and led to improving applicant responses. This more structured (and streamlined) system for providing feedback helped to simplify some of the operational aspects and burden of the application assessment.

A clearly articulated IDR definition i.e. requirement for integration, co-creation and reciprocal benefits across the fields alongside clear guidance on how to operationalise this definition in applications (for applicants) and in assessment (for panel members) was provided alongside training for IAC members and chairs. This was seen to be a very valuable aspect of CRCRM.

> " The strongest aspects for me are the close attention to honing college members' understanding of IDR...and having a clear definition of IDR is integral to the scheme" – Panel chair

A clearly articulated IDR definition has enabled the CRCRM scheme to play a key role in advancing understanding of IDR across UKRI and the wider research community. Through guidance, webinars and engagement activities, the scheme has helped to embed a clear, shared definition of IDR including how to design, articulate and assess high-quality interdisciplinary approaches.

The iterative approach to the development of the scheme by the IRM team enabled **learning from Round 1 processes** to lead to improving the clarity and

understanding of the IDR definition, CRCRM scope, assessment criteria and processes among applicants and IAC members in Round 2. It also contributed to improvement not only in the guidance and training of IAC members and applicants, but also the efficiency of delivery and the robustness of the assessment processes with improved understanding and confidence among IAC members in how to provide feedback to applicants, score and band applications. Examples of approaches that improved the effectiveness of the assessment processes include, improved processes for matching applications to assessors, the introduction of readers to ensure better engagement of panel members and supporting consensus building of applications with divergent scores.

The two-stage application process continues to be seen as a valuable feature, bringing efficiencies for applicants, UKRI and reviewers by saving time and effort through having a short outline application in the first instance. The introduction of demand management in Round 2 helped further reduce the workload for UKRI and IAC and improve success rates.

3. **CRCRM projects have higher interdisciplinarity and novelty compared to other UKRI IDR funding schemes**

   CRCRM applications and awards have higher interdisciplinarity compared to benchmark projects as exemplified by HHI and entropy indicators. These indicators like most other interdisciplinarity indicators do not differentiate between multi- and inter-disciplinarity, but the results suggest that diversity of fields is comparatively higher in CRCRM. Moreover, both HHI and entropy scores are much higher for the successful CRCRM applications compared to benchmark projects, suggesting a potential enabling mechanism in the selection process promoting interdisciplinarity.

   Analysis of SIFT scores, an alternative approach to assessing IDR, also suggests that that projects submitted to CRCRM are more likely to involve "interdisciplinary" approaches than multidisciplinary approaches. This is evident when comparing CRCRM successful applications with a random sample of UKRI awarded projects on the Gateway to Research portal. The difference was more prominent in Round 2 compared to Round 1. These results suggest that the assessment process identifies and funds applications that were able to demonstrate greater interdisciplinarity through integration of disciplines.

   CRCRM applications also have about two times greater odds of being novel compared to benchmark projects, as estimated by the relative novelty indicator score. This finding seems to support the view that IDR has the potential to lead to breakthroughs and even new disciplines.

Analysis also showed that research that was rated high for interdisciplinarity but was not considered high quality for other reasons, such as methodological flaws, did not get awarded funding. This also demonstrates that CRCRM assessment processes did not favour interdisciplinarity over excellence.

4. **The CRCRM pilot supported applications across a diverse range of fields**

Environmental challenges and health and wellbeing were the primary problem and impact areas targeted across both rounds and at all application stages (around 40-50% of all applications and awards).

> "There was a great deal of exciting work looking at the real worlding of theoretical issues or applying knowledge from different domains to real-world impact. The impactfulness of these IDR applications was very impressive." – Panel Chair

Technical, that is engineering, technological and biomedical, methods dominated Round 1 (39-40% of applications and awards), while participatory approaches and co-design were more prominent in Round 2 (around 40% in applications, 66% in awards).

Both IRC and Technopolis noted the disparities in the number of AHSS and STEM applications. Around 49% of Round 1 applications and 40% of Round 2 applications in both application stages fall within the STEM-STEM interfaces (EPS-EEB, BIOM-EPS and EEB-BIOM). This may be due to a shorter cognitive distance between STEM areas where IDR may be easier and/or more well-established, although it should be noted that the interfaces represent the predominant disciplinary areas and do not preclude involvement of other disciplines. In fact, it appears that social sciences are included to some extent in many of the CRCRM applications regardless of the interface; thus CRCRM may possibly be supporting novel disciplinary combinations and reciprocal learning that had previously not been supported. The focus on health and wellbeing and environment and sustainability problems and impacts in many applications may also have resulted in more applications assigned to STEM-STEM panels, which may still use social science methods and approaches.

The higher proportion of applications in STEM-STEM interfaces may reflect the relative UK researcher population in different disciplines. For example, there are many more STEM researchers in the UK compared to AH researchers (2.7 times more). This may also be the reason why applications with AH disciplines at the 'primary' interface were less well represented among the applications received (about a fifth and a quarter of applications in Round 1 and Round 2 respectively). Another possible explanation is that the two-council requirement in CRCRM may limit

applications from funding councils that represent a broader range of disciplines compared to the STEM research councils.

5.2 Recommendations for IDR opportunities and assessment processes

Based on the evaluation findings, the evaluators have the following recommendations for funders designing and implementing IDR funding opportunities.

1. **It is important to have a common understanding of what IDR means in the context of the specific IDR funding programme being implemented and its objectives**

   Throughout the early stages of implementing the CRCRM scheme, it became apparent that there were many differing interpretations of interdisciplinary, multidisciplinary and transdisciplinary research approaches. The scope of the CRCRM scheme was specifically to support IDR due to the challenges associated with its assessment. Therefore, considerable effort went into clarifying the definition of IDR according to the scheme and providing training and guidance on how to operationalise it in practice, particularly in applications and assessment.

   The key learning therefore is that it is important to articulate the specific definition of IDR being implemented in a funding programme at the outset and provide guidance on how this should be operationalised in applications as well as assessment. Guidance should clearly articulate how applicants can develop high quality interdisciplinary approaches and how interdisciplinarity should be assessed in applications.

   The following definitions of multi- and inter-disciplinary research were developed by the CRCRM Project Board as a result of the learning through the scheme.

   "Interdisciplinary research combines distinct components of two or more disciplines through the integration of separate disciplinary data, methods, tools, concepts and/or theories in order to address a complex issue, question, or problem. Interdisciplinary projects will be co-created/co-designed with input from all disciplines involved and the disciplines genuinely integrated with each discipline gaining something significant by being part of the project."

   "Multidisciplinary research involves different disciplines working towards the same challenge. The work packages do not require integration and remain largely separate and independent of each other, with the outputs of each combining to address the research challenge."

Multidisciplinary projects were not in scope for CRCRM but can be supported through other UKRI funding opportunities.

UKRI should consider adopting these definitions consistently across its activities to support shared understanding and consistent interpretation of these terms.

2. **IDR applications should be assessed by peers with interdisciplinary expertise, preferably those trained in reviewing IDR applications**

   Considering the diversity of topics and approaches covered in IDR, a trained IAC or similar group is a useful resource for identifying the right assessors for an IDR application. The evaluation findings show that the IAC is a key strength because it provides access to a broad range of expertise and trained, experienced IDR assessors.

   Interesting and high-quality IDR applications can be disadvantaged in traditional external written review, where assessments from discipline-specific perspectives may not capture the strength of the interdisciplinary whole. Proposals under the cross-council remit agreement (CCRA) may be judged by panels with limited experience of assessing IDR. This is where a trained college with experience of assessing IDR can ensure fair and robust assessment of IDR applications.

   The evaluation also demonstrates how experience of Round 1 led to improved understanding of the IDR definition, assessment criteria and processes among the IAC members, and how this led to improved and robust assessment in Round 2. It would be a shame to lose this resource and UKRI should look at finding ways to retain this expertise to support application assessment in other IDR funding opportunities. The training developed in CRCRM could also be used more widely.

3. **Deliberation should be a key element of the assessment process**

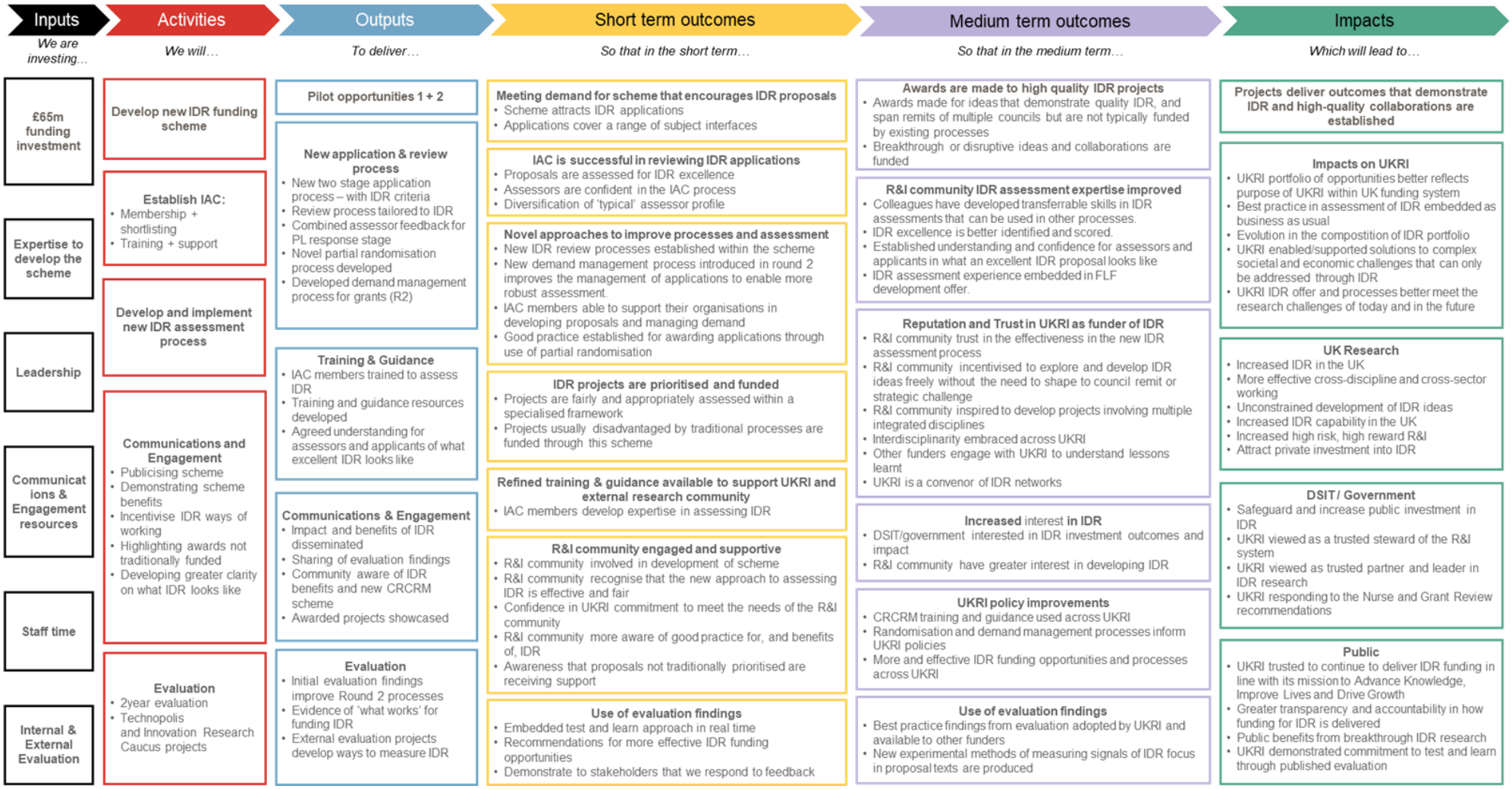
   Panel approaches and combined feedback processes that enable effective deliberation and consensus building, bringing together different interdisciplinary expertise and perspectives to better manage divergent opinions and scores allow for fairer and more robust assessment of IDR. Embedding deliberation within pre-panel feedback processes is also highly valuable to mitigate contradictory reviews.

   The CRCRM scheme developed an assessment process that enabled high quality applications with higher levels of interdisciplinarity and higher novelty to be funded compared to other IDR schemes. Therefore, UKRI and other funders should consider adopting the innovative approaches developed within CRCRM to ensure that deliberation is embedded throughout the assessment process in any future IDR funding opportunities.

4. **CRCRM has developed and tested some good practices that could be shared more widely across UKRI and other funders**

   Inclusion of readers in Round 2 full stage was well-structured and useful according to applicants and IAC members. Aspects such as scoring guidance on weaknesses, outline and generalised feedback processes, partial randomisation and demand management improved the efficiency, effectiveness and quality of the assessment. These could be taken up by UKRI and its councils as part of business-as-usual processes, not just for IDR assessment.

Appendix A Theory of Change



Appendix B Methodology

B.1. Evaluation questions

#	Question	Responsible owner
1.	**Process**	
	Understand if and how the scheme, and its processes are of high quality, are efficient and effective in the assessment of IDR applications across all disciplines to inform best practice and approaches for the assessment of IDR moving forward	
a)	i. How well has the programme addressed common challenge outline in the assessment of interdisciplinary research? ii. What further challenges related to doing and assessing IDR have surfaced as part of the evaluation of the pilot scheme? To what extent do they align with challenges identified in the literature?	Technopolis IRC
b)	Has this been done efficiently, in a manner that optimises the use of resource in UKRI and within ROs?	Technopolis
c)	Was the resource allocated to this scheme sufficient for its delivery?	Technopolis
d)	Are the assessment processes understood and regarded as credible, internally and externally?	Technopolis
e)	Did we provide effective, high-quality communications and 'user experience' for applicants, assessors, UKRI staff, broader research community	Technopolis
f)	How effective was the process of establishing the Interdisciplinary Assessment College (IAC)? Did the applications received cover the breadth of disciplines and expertise needed to build a diverse and effective IAC?	Technopolis
g)	Was the training for IAC members sufficient? What motivated IAC applicants to apply/participate?	Technopolis
h)	i. How effective was the Assessment College approach? How could this be improved? ii. What factors affect assessments of interdisciplinarity? Are there any discernible patterns to assessments based on assessor characteristics? To the extent that these (a) exist and (b) appear to influence outcomes, what measures can be suggested to mitigate issues?	IRC, Technopolis IRC
i)	i Did we have the suitable breadth of expertise in IAC for the assessment of applications? How could this be improved? ii. What are the characteristics of the pool of assessors and how do these align/differ from the pool of applicants/applications?	IRC, Technopolis IRC

j)	Does the assessment process allow for all voices in the room to be heard? Did panel discussions facilitate balanced integration and tensioning of cross-disciplinary perspectives?	IRC, Technopolis
k)	What does the college think and feel about the process and its value?	Technopolis
l)	How do applicants think and feel about the provision, quality and value of feedback process and its value?	Technopolis
m)	How effective was the partial randomisation process? How was it viewed by the college members on the panels? How was it viewed by the applicants?	Technopolis
n)	What level of support did applicants receive from host organisations in helping them to navigate the interdisciplinarity terminology in the call text (for example, what is meant by reciprocity and integration), in developing applications, and in facilitating conversations between interdisciplinary researchers? How clear was the guidance in enabling host organisations to support applicants in developing applications, with respect to eligibility, costings, scope etc.?	Technopolis
	Portfolio	
2.	Understand how and to what extent the scheme's processes have enabled the effective assessment and funding of new interdisciplinary research, and impacted perceptions of IDR support in the UK funding landscape	
a)	i. To what extent, if at all, has the programme funded what we intended- high quality research, transformative and well-integrated interdisciplinary research ideas? ii. To what extent does the funded/fundable research appear to be interdisciplinary, as compared to the unfundable research? What are the dimensions of interdisciplinarity (for example in the project and in the team) and how do applications suggest they will be integrated?	Technopolis IRC
b)	How and to what extent are the projects distinctive and complementary to interdisciplinary research funded through individual Council Responsive Mode?	Technopolis UKRI
c)	How diverse is the subject breadth– does the funded portfolio span a range of research areas/ disciplines/ councils (some limitations here given size of scheme, the 'unfunded' portfolio will be important indicator in round one)?	UKRI
d)	i. What is the nature of collaborations in awards? Do they contain a range of different types of collaborations including cross-institutional and international? ii. How is collaboration being proposed to underpin IDR? What indicators demonstrate that collaboration serves IDR goals and is not occurring in silos (for example)?	UKRI IRC

	iii. What impact does the diversity of organisations involved in the proposed projects have on assessments of interdisciplinarity (for example, interdisciplinarity scores) and final outcomes? Does having more organisations involved positively affect outcomes?	IRC
e)	Does funding span a range of career stages, diversity on four protected characteristics (sex, age, ethnicity, disability)?	UKRI
f)	Does funding span a range of organisations – the diversity of organisations involved (as lead and co-investigators).	UKRI
3.	**Impact of the process on the portfolio**	
	Understand the impact of the scheme on the nature of the applications and the portfolio of awards supported through the scheme, specifically focused on - the extent to which new research and people are being supported and how they differ to that funded by other UKRI schemes, to inform best practice and approach for supporting IDR moving forward - the extent to which and how the research and innovation (R&I) communities' perceptions and behaviours have changed, with respect to funding routes for IDR - how the outcomes and impact of the IDR funded differ (if at all) from other responsive mode schemes	
a)	i. To what extent is the programme incentivising, facilitating, and reducing barriers to IDR ii. What barriers does this analysis suggest might exist to meaningful IDR?	Technopolis IRC
b)	To what extent does the research supported differ for other responsive mode schemes	Technopolis UKRI
c)	Where the scheme has met its goals/ambitions, what conditions/factors seem to have enabled that?	Technopolis
d)	Where the scheme hasn't met its goals/ambitions, what conditions/factors have held it back?	Technopolis
e)	What is the nature of the cross-disciplinary research challenges being tackled?	Technopolis
f)	To what extent has this scheme enabled the research community to change their approaches to developing IDR applications and creating innovative ways of thinking	Technopolis
g)	What should we do the same/ differently going forwards	Technopolis

B.2. Interviews

Technopolis collected stakeholders' views on and experience of the CRCRM pilot and its processes across different stages of the programme through semi-

structured interviews. The interview programme was carried out in three stages during the implementation of the CRCRM pilot scheme:

- Round 1 (Winter 2024/25)
- Round 2 outline stage (summer 2025)
- Round 2 full application stage (Winter 2025/26)

At each stage, our interview sample IAC members in different roles, successful and unsuccessful applicants as well as a small number of interviews with UKRI staff. In total, the evaluation team interviewed a total of 67 stakeholders. The profiles of stakeholders of interviews are summarised in below.

Table 7 Breakdown of interviews

Role	**Round 1**	**Round 2 outline**	**Round 2 full stage**	**Total**	**Gender** (male:female)
Successful applicant *	8	4	4	16	9:7
Unsuccessful applicant	2	2	2	6	5:1
Applicants subtotal *	10	6	6	22	14:8
IAC Panel Chair / Co-chair	3	2	3	8	5:2
IAC Deputy Chair	2	1	1	4	1:3
IAC Panel member	5	1	2	8	4:4
IAC Rover	2	3	2	7	4:3
IAC subtotal *	13	7	8	28	15:12
UKRI staff	8	9 *	1	18	6:11
Grand total	30	22	15	67	35:31

Note: * Interviewee with multiple roles counted under both applicants and IAC subtotals

B.3. Surveys

Round 1 and Round 2 surveys were conducted by UKRI with applicants immediately after application submission deadlines, before assessments were completed and with IAC members (panel members, rovers, chairs and deputy chairs/co-chairs) after the completion of their assessment roles at each outline stage and full application stage. Technopolis analysed answers to closed questions using Excel and those to the open questions using Technopolis' in-house AI Policy Concierge (AIPC). The response rates for the survey are provided in the table below. The response rates were high especially for Round 1 surveys, with rates of 62% or more. In Round 2, both project leads and co-leads were invited to provide responses unlike in Round 1 where only project leads were approached; response rates were lower, at 24% and 33% for the outline and full application stages respectively but still led

to 900 responses for the two surveys. UKRI issued a survey to the IAC towards to the completion of Round 2 regarding the future of the IAC and their interest in remaining on the college if the scheme was to continue. Of the 80% who responded 96% were positive about continuing their membership. Members also shared testimonials regarding their experience of being part of the IAC, several of which are quoted in the report.

Table 8 Survey response rates

UKRI CRCRM Surveys - Stages and Recipients	**Total recipients**	**Total Responses**	**Response rate %**
Round 1 Outline Stage Applicants' Survey (Project Leads only)	899	650	72
Round 1 Outline Stage IAC Survey - Panel Members	213	158	74
Round 1 Outline Stage IAC Survey - Chairs and Deputy Chairs	39	31	79
Round 1 Outline Stage IAC Survey - Rovers	13	12	92
Round 1 Full Stage Applicants' Survey (Project Leads only)	170	152	89
Round 1 Full Stage IAC Survey - Panel Members	123	92	75
Round 1 Full Stage IAC Survey - Chairs and Deputy Chairs	21	13	62
Round 1 Full Stage IAC Survey - Rovers	12	11	92
Round 2 Outline Stage Applicants' Survey (Project Leads and Co-Leads)	3214	782	24
Round 2 Outline Stage IAC Survey - Panel Members	205	118	58
Round 2 Outline Stage IAC Survey - Chairs and Deputy Chairs	34	15	44
Round 2 Outline Stage IAC Survey - Rovers	18	13	72
Round 2 Full Stage Applicants' Survey (Project Leads and Co-Leads)	484	158	33
Round 2 Full Stage IAC Survey - Panel Members	87	51	59
Round 2 Full Stage IAC Survey - Chairs and Deputy Chairs	20	13	65
Round 2 Full Stage IAC Survey - Rovers	14	5	36
TOTAL RESPONSES ANALYSED BY TECHNOPOLIS		**2274**	
The Future of the UKRI Interdisciplinary Assessment College - IAC members	318	253	80

TOTAL RESPONSES OVERALL		**2527**	

Source: UKRI

B.4. Interdisciplinarity and novelty analysis

To identify early signs of added value from CRCRM and to understand the portfolio emerging from the implementation of the pilot's assessment processes, Technopolis compared CRCRM applications and awards to projects awarded under a group of benchmark programmes. To ensure we were able to provide a robust comparison against different configurations of IDR and assessment processes, we chose three schemes to cover a broad set of disciplinary mixes and both strategic and responsive mode schemes as benchmarks. These were

Strategic collaborative research scheme that is, Physics of Life (SPF), which is a top-down IDR funding scheme that invites applications for interdisciplinary projects combining the biological/biomedical and physical sciences specifically

Strategic challenge-led scheme that is, Global Challenges Research Fund, which is a top-down IDR funding scheme aimed at funding interdisciplinary collaborations to address key global challenges (not necessarily specifying the disciplinary mix)

Responsive mode scheme that is, projects funded under cross-council remit agreement, which is a policy for funding bottom-up, applicant-led IDR

We compared the novelty and interdisciplinarity of applications and awards for the CRCRM pilot with projects funded by the benchmark programmes. For this we computed interdisciplinarity and novelty indicators using advanced natural language processing tools. The CRCRM round 1 data (outline applications) provided by UKRI and benchmark data retrieved from Gateway to Research contain field labels and keywords; however, these tags do not present a systematic and harmonised approach across datasets. Considering all these factors, we opted to use a methodology (text embeddings) that identifies the semantic meaning of the text available (project titles and summaries) and, thus, its scientific content.

The analysis used only these textual descriptions enriched by a proxy of their respective prior art. Our prior art configuration is specific to each individual application or funded project. It consists of the most semantically similar scientific contributions among over 200 million works published the year before each application or funding year.

We computed the following **interdisciplinarity indicators**:

- **Herfindahl-Hirschman Index (HHI)**: The HHI measures concentration by summing the squared proportions of each field's counts relative to the total. A lower HHI reflects a more evenly distributed focus across multiple fields, reflecting higher interdisciplinarity. The HHI captures how balanced or unbalanced the distribution of prior art is across different fields. If all the prior art for a application came from a single field, the HHI would be maximal (equal to 1), reflecting no interdisciplinarity.
- **Entropy**: Entropy measures the degree of diversity or "spread" in a distribution. In the context of this analysis, a application's entropy value is higher when it draws on a balanced range of distinct fields, indicating that its prior art is widely spread across multiple research domains. Conversely, a low entropy value reflects a concentration of prior art in fewer fields.

It should be noted that the indicators are not able to differentiate between multidisciplinarity and interdisciplinarity.

The **novelty indicator**[21] stems from a machine learning model that learns from the novelty patterns in scientific articles that led to Nobel prizes (papers renowned for their high levels of novelty). The model's features comprise solely the text of an article and the text of the closest prior art, measuring the originality of each scientific contribution. Our machine learning approach demonstrates strong predictive power in identifying highly novel ideas across disciplines. In an econometric framework controlling for citations, journal venues, publication year and topics, the model remains a robust predictor of novelty, effectively disentangling it from quality and other scientific attributes.

Finally, we use a "Semantic fixed-effects" method to allow a better "like-to-like" comparison, while controlling for nuanced differences in topic content that are not be captured by conventional field boundaries. This is done by comparing each CRCRM application's indicators only against thematically related benchmark projects that are matched on the basis of the highest similarity of text embeddings. All the comparisons across all applications are then aggregated to assess overall differences in interdisciplinarity and novelty.

B.5. Signals of Interdisciplinary Focus in Text (SIFT)

SIFT is an indicator born of necessity. The ways that assessors were instructed to score interdisciplinarity changed several times across the different rounds and stages of the CRCRM process in response to the feedback from panels.

[21] Machado, D. (2024). "The meaning of novelty: predicting novelty with text embeddings". Presented at WOEP FULL4, TURIN, 2024; LIS-Bibliometrics, Brighton, 2024; and REvaluation Conference, Vienna, 2024; and upcoming at the ATLC25 and Global TechMining Conferences, Atlanta, 2025.

This led us to create a measure that could be applied consistently to each phase and compared over time. Our challenge was to develop a measure of interdisciplinarity that was more than just an indicator of diversity of disciplinary combinations and that could assess proposals without using AI tools. We also aimed to make it adaptable to other document types (e.g., future CRCRM rounds, other schemes, and beyond).

To develop a measure of interdisciplinarity for text we employed Natural Language Processing techniques (not AI) to identify sentences indicative of interdisciplinarity, which were then weighted and converted into a score reflecting each document's relative interdisciplinarity compared to other documents in the corpus of interest.

First, we developed a training dataset from 19 of the 169 Vision and Application (Full stage, Round 1) documents spread across A*, A and B Bands. Two raters identified and coded 95 sentences as low/mid/high indicators of interdisciplinarity. This involved pulling out all sentences that they judged were describing how the questions, challenge, research design, team, approach, or outcomes would be interdisciplinary, drawing on definitions of interdisciplinarity as presented in assessor training sessions/material and a review of the literature. Sentences were coded for "strength" of interdisciplinarity based on the degree to which they were exceptional or weak examples. The raters double coded 43 of these to confirm agreement, achieving statistically significant (p=0.0256) fair interrater agreement (Kappa = 0.317).

We then compared the similarity of all sentences to these training sentences, measured using a cosine similarity score (converted to a value from 0 t0 1). Compared to analysing text at a keyword level, sentence level analysis provides a strong indication of how concepts are used together (e.g. how two or more disciplines propose to interact) within the text.

Finally, we selected the top 5% of sentences similar to an example sentence and assigned them a score based on whether they were similar a low/mid or high scoring example sentence. Because only fair interrater agreement was achieved at sentence level, example sentences with double coding values were included with both these coding values. We then controlled for differing lengths of documents of by aggregating sentence scores per document, divided by the maximum possible score given then total number of sentences in the document.

The measure was validated by selecting 20 documents from the 150 Vision and Application (Full stage, Round 1) not used for training sentences, and asking two raters to code whether they displayed low/mid/high levels of interdisciplinarity. Statistically significant (p = 0.00063), interrater agreement

was achieved at a substantial level (Kappa = 0.754). Rater 1 achieved an F1 validation score of 0.6 compared to the SIFT predicted category, while Rater 2 achieved an F1 validation score of 0.455. Most disagreement lay between mid (rater coded) and high (SIFT predicted) document.

This is related to one of the current weaknesses of the SIFT score. It bases its scores on similarity alone and does not exercise judgement about how convincing a sentence is in context. Where human assessors differ from high SIFT scores is because the sentences are well written and say all the right things but assessors with experience and reasoning are not persuaded that what is being stated is genuine or feasible. This highlights the importance of including human assessors in all, and not only interdisciplinary, peer review. At the other extreme, SIFT will tend to assign lower scores to some proposals that make strong arguments about interdisciplinary integration across multiple linked sentences or paragraphs, where each component is not high scoring but the whole, taken together, presents a compelling case. While our approach is designed to account for and compensate for instances where multiple lower scoring sentence occur in clusters, this process may not yet be sensitive enough to catch these types of cases. We are currently working to adjust our method to mitigate these limitations.

These limitations are important to keep in mind when considering the appropriate application of SIFT. It was a useful tool to create a relatively accurate picture of relative levels of interdisciplinary potential within a sample of text (in this case, proposals). When applied to a random sample of publicly available proposals from the Gateway to Research (GtR) database of UKRI funded research, it showed that CRCRM applications had significantly higher scores and that the GtR proposals that scored highly were from specifically interdisciplinary funding opportunities. As such, SIFT could be used on incoming proposals to determine if any have the potential to be interdisciplinary and might require different assessment processes. It could also be used on portfolios of previously funded research to provide an overview of how much of it might be IDR. SIFT might also be useful as a step in evaluating completed research to explore its interdisciplinarity. There may be other applications as well. In each case, however, we recommend using SIFT as *an indicator* and not as the sole consideration for funding or evaluation of research success.

B.6. AI Policy Concierge (AIPC) analysis of application summaries

Technopolis' in-house AI tool, or AIPC, was used to identify terms relating to the following areas based on the application title and summary in both rounds and vision and approach and fit to opportunity text in round 2:

- Keywords

- Interdisciplinarity
- Problems being addressed
- Expected impacts
- Methods

However, the terms identified were too granular to analyse across all proposals so were clustered into higher level categories using the AIPC. The prompt requested that the bot reviewed the full list of words twice to understand the diversity before developing the categories, to remove any bias from creating the categories at the same time as processing the words. Both the categorisation of the terms and the categories themselves were manually reviewed and amended as necessary.

Once the categorisation was agreed, we were able to transform the original classification into a more standardised format.

For each interface (and all interfaces combined), the most frequently appearing categories of Methods, Impacts, and Problems were displayed as a word cloud or bar charts to observe changes in the categories with the different stages of assessment.

For disciplines, network diagrams showcasing interdisciplinary pairs were created, where the size (and colour) of the node represents the total number of proposals that incorporated that discipline and the thickness of connecting lines represents the number of proposals that involved the two connected disciplines.

Appendix C Supplementary data

C.1. Interdisciplinarity and novelty analysis

Table 9 Semantic fixed-effect regressions with CRCRM award dummy (Round 2)

	(1) HHI	(2) Entropy
CRCRM award dummy	-0.07***	0.15**
	(0.024)	(0.064)
Observations	**352**	**352**
Model	**Linear Fixed-effects**	**Linear Fixed-effects**
Semantic fixed-effects	**Yes**	**Yes**

Notes: Robust standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01
Source: Technopolis-group based on project data, Gateway to research and Semantic Scholar

C.2. UKRI analysis

Table 10 shows the data for the UKRI analysis on gender of Project Leads presented in 3.2.3. It should be noted that the applications for Round 1 were submitted to Je-S, where these data are collected under the 'Gender' section. Round 2 applications were submitted to the UKRI Funding Service, where these data are collected under the 'Sex' section, where there is a separate question on 'Gender assigned at birth'.

Table 10 Project Lead gender data for each assessment stage and awards for Round 1 (a) and Round 2 (b). Data is illustrated in Figure 20.

(a)	**Round 1**		
Gender	**Outlines**	**Full**	**Awards**
Male	580	110	30
Female	305	55	5
Not disclosed	15	5	<5
Total	900	170	35

(b)	**Round 2**		
Gender	**Outlines**	**Full**	**Awards**
Male	305	60	20
Female	195	35	10
Not disclosed	35	<5	<5
Total	535	100	30

Table 11shows the data for the UKRI analysis on ethnicity of Project Leads presented in 3.2.3. It should be noted that the applications for Round 1 and Round 2 were submitted to Je-S and the UKRI Funding Service, respectively, and therefore there are differences in the ethnicity categories between Rounds.

Table 11 Project Lead ethnicity data for each assessment stage and awards for Round 1 (a) and Round 2 (b). Data illustrated as binary in Figure 21.

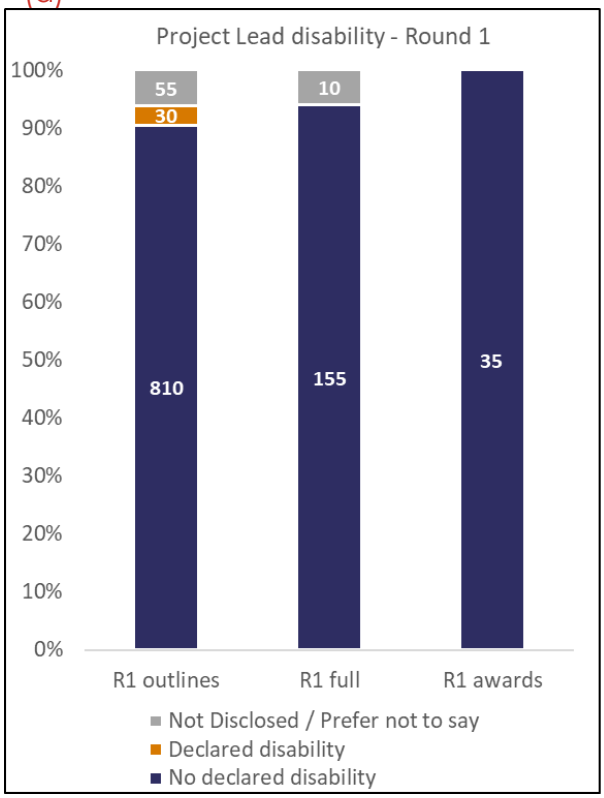
(a)	**Round 1**		
Ethnicity	**Outlines**	**Full**	**Awards**
Asian and Asian British	60	10	5
Black and Black British	5	<5	<5
Chinese and Other	65	5	<5
Mixed	30	<5	<5
White	675	125	25
Not Disclosed/Unknown	65	20	<5
Total	900	170	35

(b)	**Round 2**		
Ethnicity	**Outlines**	**Full**	**Awards**
Asian or Asian British	65	5	<5
Black, Black British, Caribbean or African	5	<5	<5
Mixed or Multiple ethnic groups	15	5	<5
Other ethnic group	10	<5	<5
White	395	75	30
Prefer not to say	35	<5	<5
Not Disclosed	10	<5	<5
Total	535	100	30

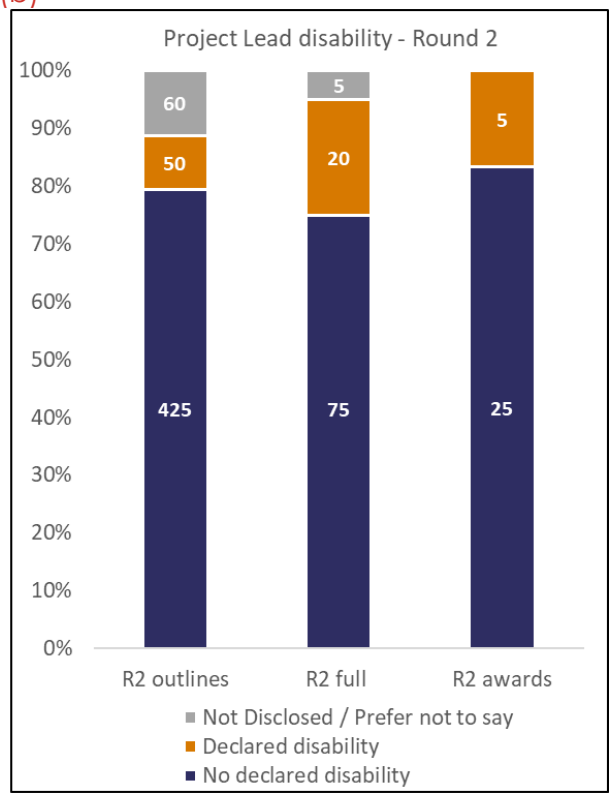
Figure 32 and Table 12 shows the data for the UKRI analysis on disability of Project Leads.

Figure 32 Project Lead disability for each assessment stage and awards for Round 1 (a) and (b)

(a)



(b)



Source: UKRI

Table 12 Project Lead disability data for each assessment stage and awards for Round 1 (a) and Round 2 (b).

(a)

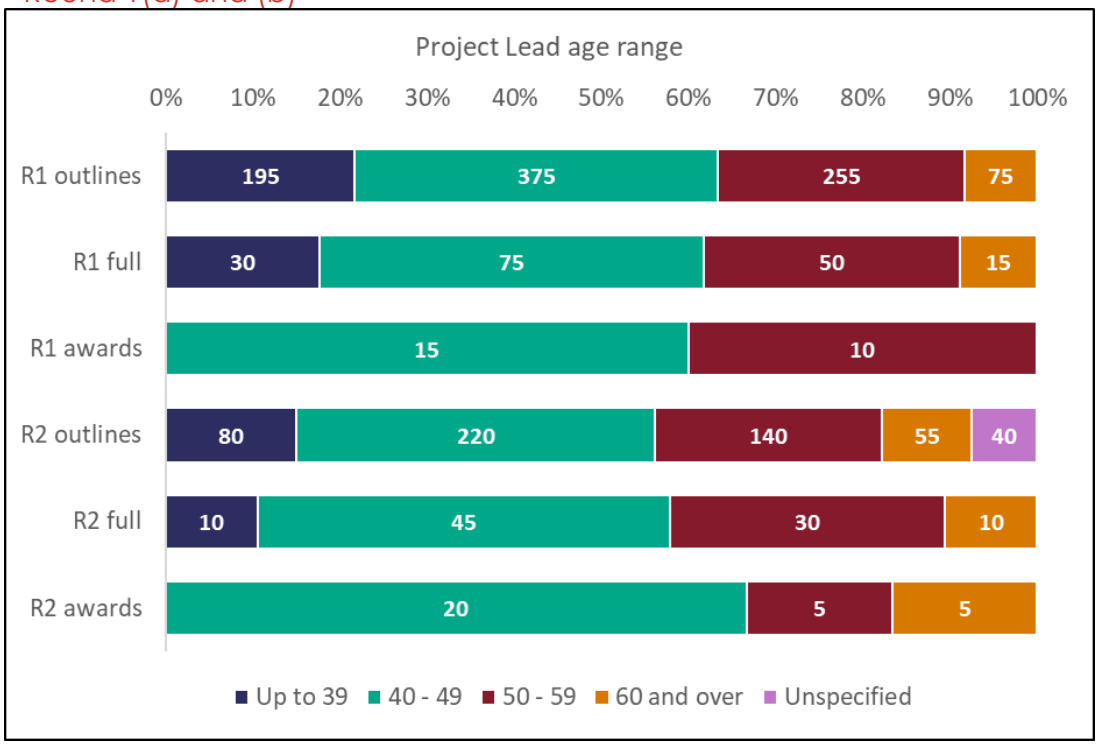
	Round 1		
Disability	**Outlines**	**Full**	**Awards**
No declared disability	810	155	35
Declared disability	30	<5	<5
Not Disclosed / Prefer not to say	55	10	<5
Total	900	170	35

(b)

	Round 2		
Disability	**Outlines**	**Full**	**Awards**
No declared disability	425	75	25
Declared disability	50	20	5
Not Disclosed / Prefer not to say	60	5	<5
Total	535	100	30

Figure 33 and Table 13 shows the data for the UKRI analysis on age range of Project Leads.

Figure 33 Project Lead age range for each assessment stage and awards for Round 1(a) and (b)



Source: UKRI

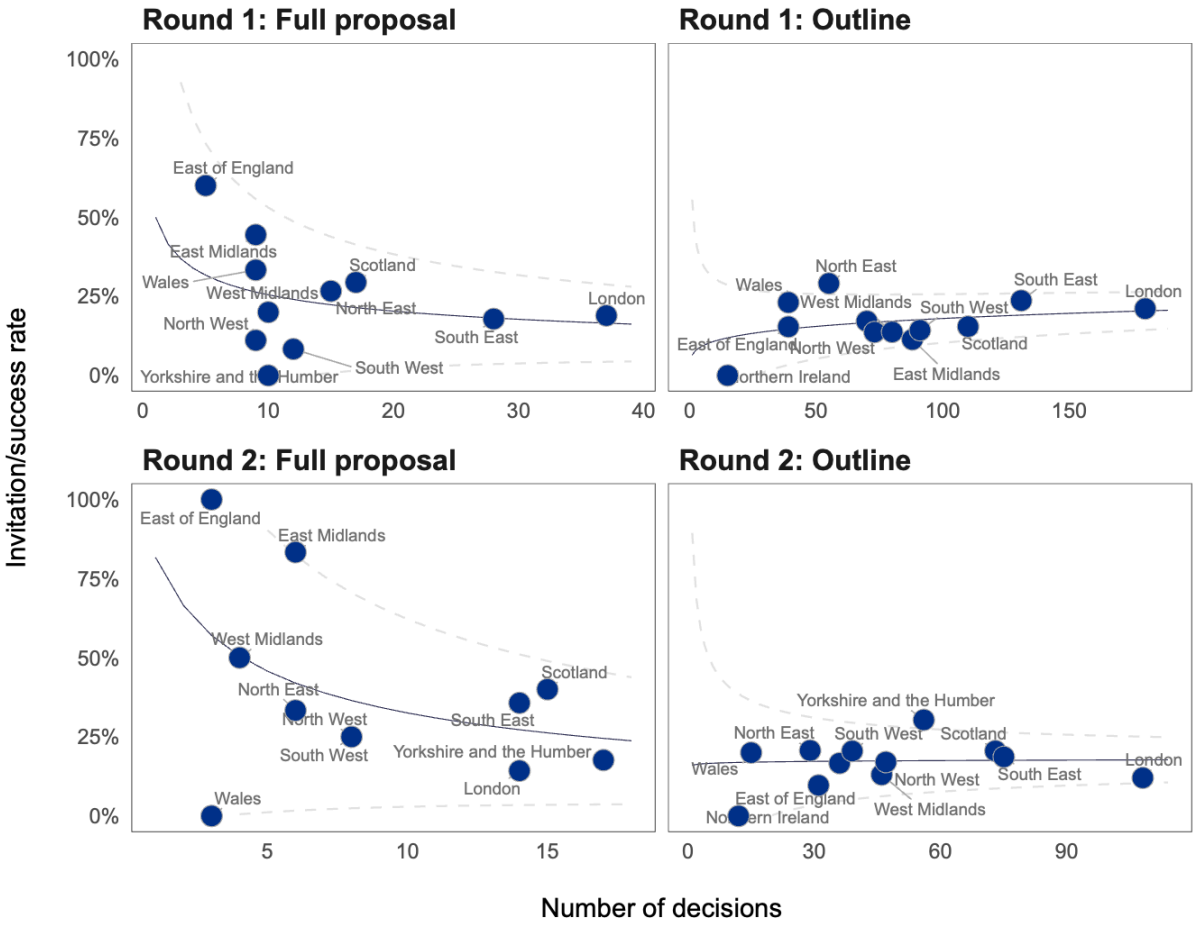
Table 13 Project Lead age range for each assessment stage and awards for Round 1(a) and (b)

(a)	Round 1		
Age range	Outlines	Full	Awards
Up to 39	195	30	<5
40 - 49	375	75	15
50 - 59	255	50	10
60 and over	75	15	<5
Not disclosed	<5	<5	<5
Total	900	170	35

(b)	Round 2		
Age range	Outlines	Full	Awards
Up to 39	80	10	<5
40 - 49	220	45	20
50 - 59	140	30	5
60 and over	55	10	5
Not disclosed	40	<5	<5
Total	535	100	30

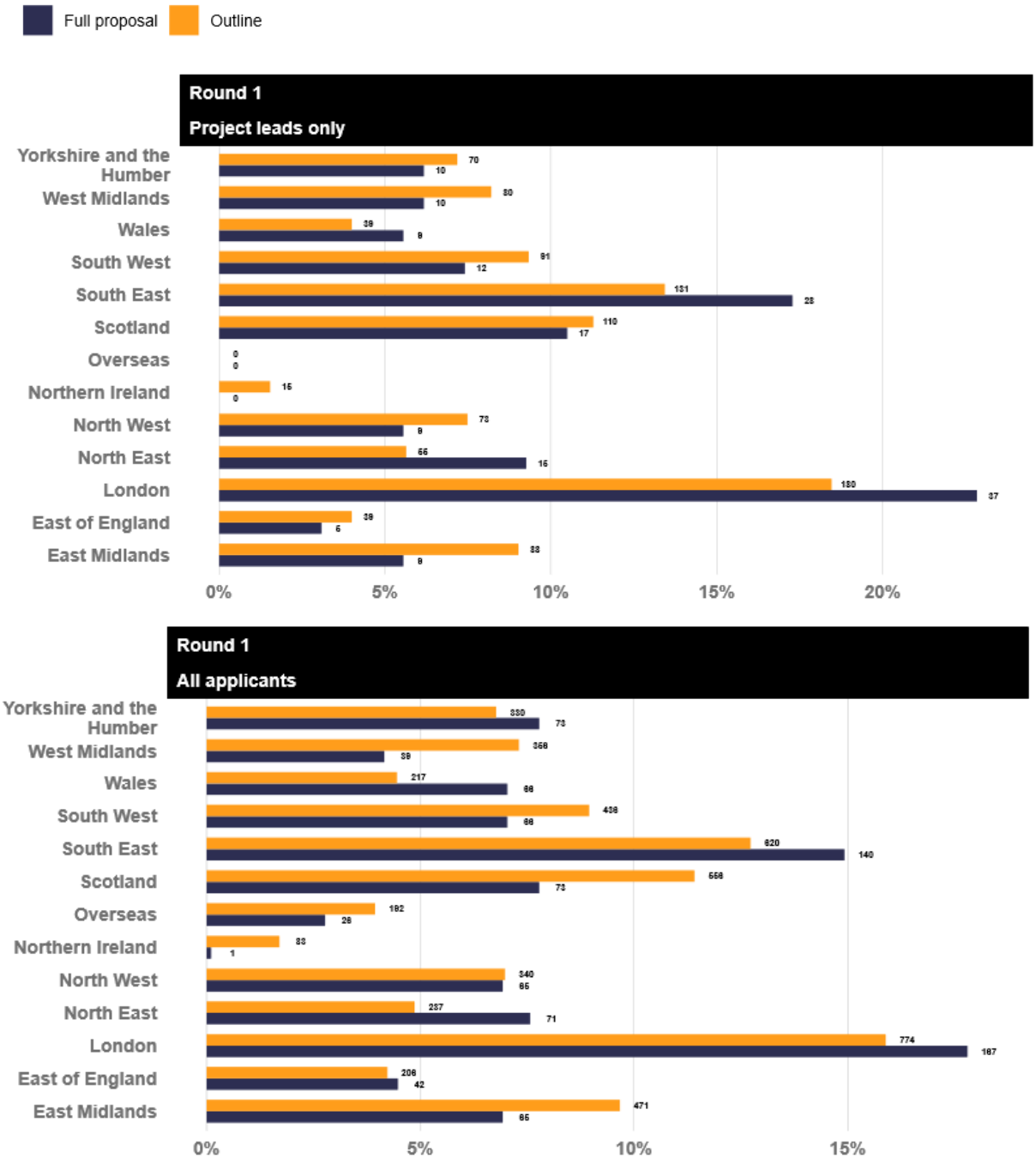
Figure 34 shows the UKRI analysis relating to regional success in the CRCRM pilot in four funnel plots. For each UK region/nation (of the lead applicant), the plots show the relationship between success probability (y axis) and number of decisions made (x axis) – this is the central solid curve in each plot. Note that the phrase 'decisions made' is based on the assumption that each application receives a funding decision. For full applications there is a negative correlation between decision volume and success rate. For outlines there is no strong relationship between decision volume and success. These plots show us that only two regions have unexpectedly high or low rates in any round. In Round 1 outline stage, the North East has a high rate, and in Round 2 outline stage, Yorkshire and the Humber has a high rate.

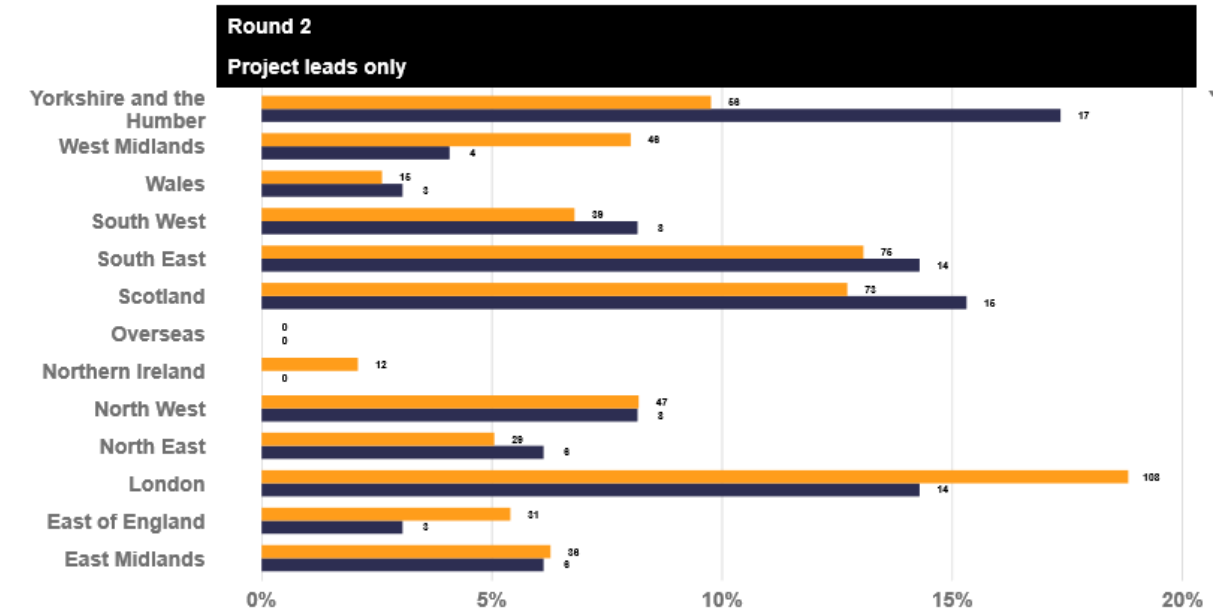
Figure 34 CRCRM application invitation and success rates by region by round and stage

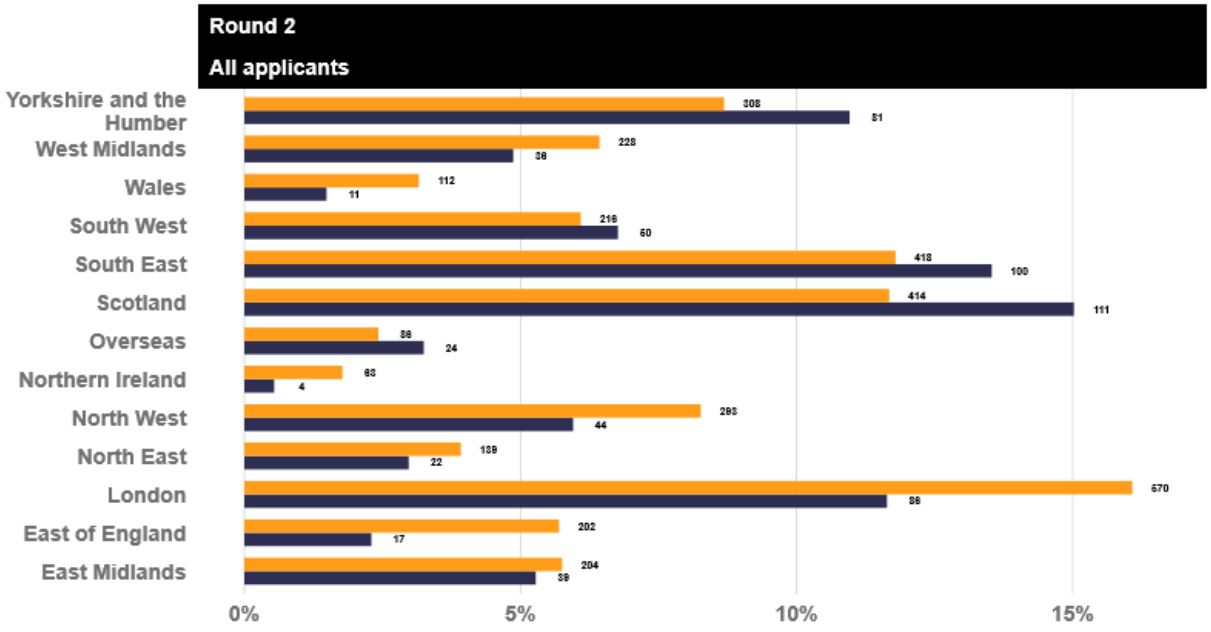


Source: UKRI

Figure 35 Regional representation of project leads and applicants

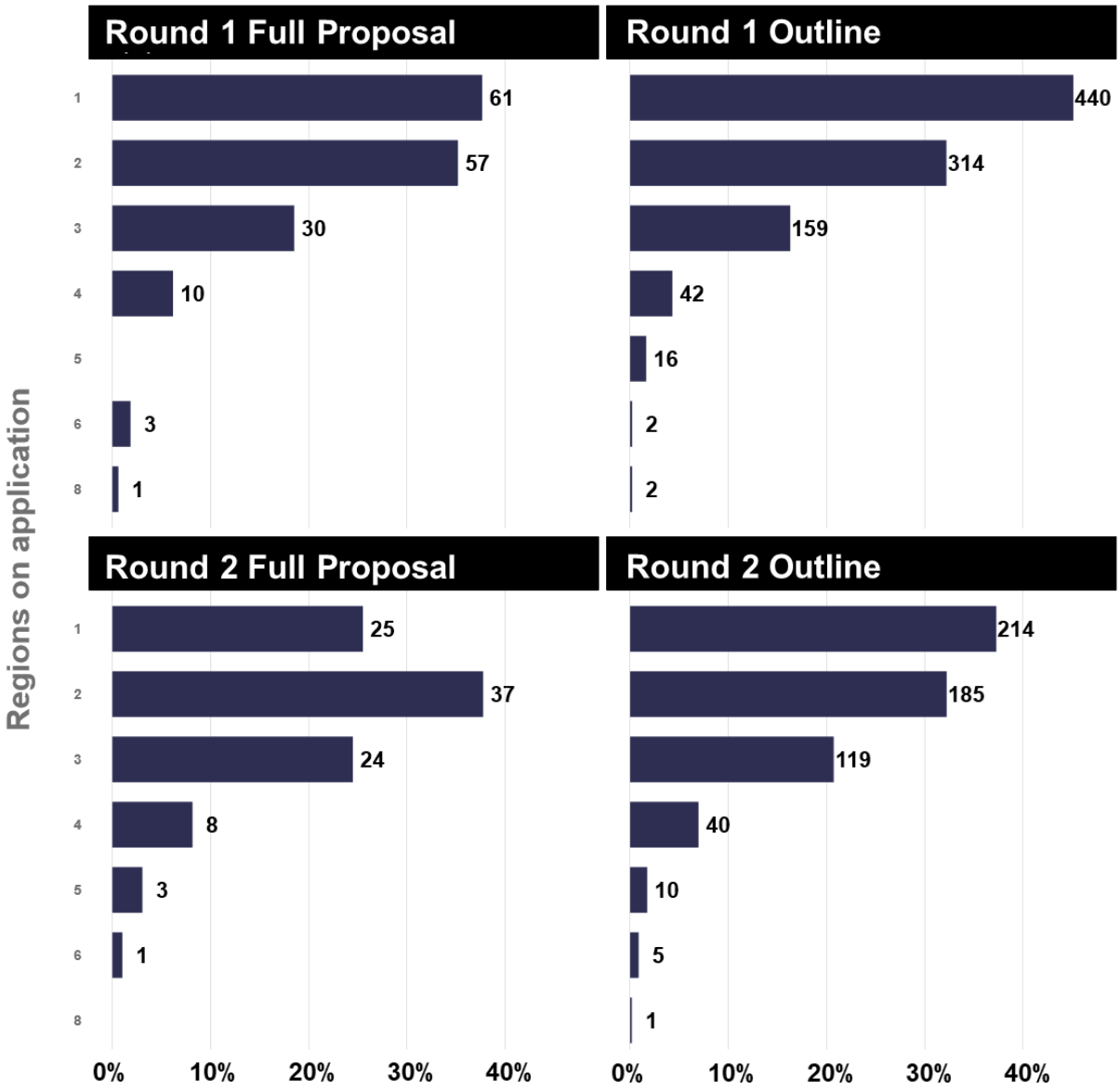



Round 2
Project leads only
Yorkshire and the Humber
West Midlands
Wales
South West
South East
Scotland
Overseas
Northern Ireland
North West
North East
London
East of England
East Midlands
58
17
48
4
15
3
39
8
76
14
73
16
0
0
12
0
47
8
29
6
108
14
31
3
38
6
0%
5%
10%
15%
20%


Round 2
All applicants
Yorkshire and the Humber
West Midlands
Wales
South West
South East
Scotland
Overseas
Northern Ireland
North West
North East
London
East of England
East Midlands
308
81
228
38
112
11
218
60
418
100
414
111
88
24
83
4
285
44
138
22
670
88
202
17
204
59
0%
5%
10%
15%

Source: UKRI

Figure 36 Number of regions on applications



Source: UKRI

C.3. Clarification Statements sent to Round 2 outline stage applicants

Fit to Scheme	**1**	The proposal demonstrated **clear fit to scheme** with good integration, co-creation and reciprocal benefits
	2	The proposal showed **some level of fit** to scheme
	3	The proposal showed a **low level of fit** to scheme and therefore did not meet this criterion
If selected #2 or #3 choose all that apply from #4-#7	**4**	The proposal did not fully evidence how **disciplines** would be integrated
	5	The proposal did not fully evidence how the project was **co-created**
	6	The proposal did not fully evidence what the **reciprocal benefits** would be for the disciplines involved
	7	The outlined proposal was not considered to meet the **scope of the call** by covering the **remit of a 2nd council by at least 10%**
Vision	**8**	The outlined proposal was considered to have **exceeded the quality threshold** required by the assessment process in terms of importance and the potential to advance current understanding, generate new knowledge, thinking or discovery beyond the field or area
	9	The outlined proposal was considered to have **met the quality threshold** required by the assessment process in terms of importance and the potential to advance current understanding, generate new knowledge, thinking or discovery beyond the field or area
	10	The outlined proposal was considered to have **not met the quality threshold** required by the assessment process in terms of importance and its potential to advance current understanding, generate new knowledge, thinking or discovery beyond the field or area
Choose 1 of the following if applicable:	**11**	The outline proposal was considered to be **very timely** given current trends, context and needs
	12	The outline proposal was considered to be **fairly timely** given current trends, context and needs
	13	The outlined proposal was **not considered to be timely** given current trends, context and needs
then		
Choose 1 of the following if applicable:	**14**	The outlined proposal was considered to have the **potential to impact** world-leading research, society, the economy, or the environment
	15	The outlined proposal **requires further clarity on how it could impact** world-leading research, society, the economy, or the environment
	16	The outlined proposal **demonstrated limited potential to impact** world-leading research, society, the economy, or the environment
Approach	**17**	The outlined **approach is effective and appropriate** to achieve the objectives
	18	There were some **minor concerns** about the effectiveness and appropriateness of the outlined approach in achieving the objectives
	19	There were some **significant concerns** about the effectiveness and appropriateness of the outlined approach in achieving the objectives or the outlined approach was not clear in how it would achieve the objectives
Team Capability	**20**	The proposed team was considered to have **high degree of the relevant experience and skills** to develop and deliver the proposed research programme
	21	The proposed team was considered to **have the relevant experience and skills** to develop and deliver the proposed research programme
	22	There were **some concerns about the relevant experience and skills** of the proposed team needed to develop and deliver the outlined proposed research programme or the application was considered to be **missing a particular area(s) of expertise**

Source: UKRIDo not delete anything after this (non-printing) line

technopolis
group

www.technopolis-group.com