



UK Research
and Innovation



Equality, diversity and inclusion in research and innovation: international review

Jess Moody and Dr Amanda Aldercotte,
Advance HE, for UK Research and Innovation



Contents

Chapter 1:	Executive summary	5
	Overview	6
	Key findings	6
	Key recommendations	7
Chapter 2:	Introduction	8
	2.1 Project background, scope and research questions	9
	2.2 The international context	9
	2.3 Reflexivity	10
Chapter 3:	Methodology	11
	3.1 Overview	12
	3.2 Data reduction	13
	3.3 Evaluation framework	14
	3.4 Descriptive analysis of the international dataset	15
Chapter 4:	Previous studies	20
	4.1 Overview	21
	4.2 Key themes emerging from organisational reviews	22
	4.3 Conclusions and recommendations	24
Chapter 5:	What works?	26
	5.1 Overview	27
	5.2 Training	27
	5.3 Strategies, policies or processes	30
	5.4 Career development programmes	33
	5.5 Employer engagement and outreach	35
	5.6 Why were these interventions more or less effective?	36
	5.7 Conclusions and recommendations	37
Chapter 6:	Measuring success	39
	6.1 Overview	40
	6.2 Methodological hierarchies	40
	6.3 Outcome measurement and reporting	41
	6.4 Evaluation methods	42
	6.5 Conclusions and recommendations	52
Chapter 7:	Enhancing data and disclosure	53
	7.1 Overview	54
	7.2 Categorisation in the international context	54
	7.3 Encouraging disclosure	56
	7.4 Improving datasets	57
	7.5 Data for decision making	58
	7.6 Conclusions and recommendations	58

Chapter 8: Who is leading?	60
8.1 Overview	61
8.2 Quantity	61
8.3 Innovative approaches	63
8.4 Wider impact	65
8.5 Conclusions and recommendations	66
Chapter 9: Conclusions	67
9.1 Conclusions	68
9.2 Gaps in the evidence base	68
9.3 Review limitations and avenues for future work	68
Chapter 10: Recommendations	69
 Bibliography	 72
Glossary	77
Appendix	79
A. Database search terms, restrictions and limitations	80
B. Organisational sources included from the targeted grey literature search or responding to the Call for Evidence	81
C. Call for Evidence	81
D. Working understanding of identity characteristics within the international context	84
E. Data refinement	85
F. Evaluation framework	87
 Acknowledgments	 93
About Advance HE	94

The background of the slide is an abstract composition. It features a dense network of thin, glowing lines in shades of blue and red, creating a sense of depth and complexity, reminiscent of a digital or architectural structure. A large, solid yellow triangle is positioned on the right side, pointing towards the bottom right corner. In the top left corner, there is a dark blue rectangular area containing the chapter title in white text.

Chapter 1: Executive Summary

Overview

UK Research and Innovation (UKRI) commissioned Advance HE to review the equality, diversity and inclusion (EDI) challenges in the research and innovation (R&I) sector. This review examines the international context (a concurrent review examines the UK context) and sits within a wider context of understanding the evidence base of EDI interventions and activities and identifying further areas for research.

A review of academic and unpublished, non-commercial ('grey') literature, supported by a small-scale Call for Evidence from organisations in the R&I sector, aimed to address five research questions:

- **previous studies** - which organisations have previously reviewed and explored the key challenges for EDI in the international R&I landscape?
- **what works?** - of the international interventions that have been implemented by organisations comparable to UKRI, which have proven effective, or less effective, and why?
- **measuring success** - how is the effectiveness of EDI interventions measured and are there methods that are particularly useful for the international R&I landscape?
- **enhancing data and disclosure** - how can EDI data capture and disclosure rates in the international R&I landscape be improved?
- **leading organisations** - which organisations are leading in terms of EDI in R&I?

To understand 'what works' across a range of R&I contexts (such as research funding, policy, employment, doctoral study) an evaluation framework was applied to a total of 109 sources covering 130 interventions. Our framework particularly focused on 'mapping' the range of activities related to different contexts and identity characteristics as well as understanding which evaluation methods had been used (if any) to understand the impact or outcomes of those activities.

Key findings

Focus on gender and 'general EDI'

A large proportion of interventions focused on gender (or sex) equality, or wider EDI issues (such as 'diversity training'). Other characteristics such as disability, religious inclusion or age were less likely to be the primary target of interventions within our sample.

A diversity of EDI interventions

A range of EDI interventions were evaluated, grouped loosely as:

- training
- strategies, policies or processes

- career development programmes
- employer engagement and outreach

Regional differences in literature on interventions

Literature evaluating EDI interventions in North America was more prevalent, with far fewer sources from Africa, Asia and South America. Limitations of the research approach that might have impacted on this are discussed, though this echoes findings of other reviews (such as Chambers et al., 2017) and suggests further targeted reviews of EDI in R&I in the Global South may be beneficial.

Factors affecting the effectiveness of interventions

There was wide variation in the way EDI activities were evaluated and how outcomes were reported. This presents challenges for understanding effectiveness (in terms of impact and sustainability). However, there appeared to be:

- evidence that positive or affirmative action measures can improve the representation of women in funding award schemes and access to higher education, as well as reducing bias towards women in recruitment (for example shortlisting)
- more confidence in the evidence around diversity training programmes and diversity management policies; in contrast, the effectiveness of family-friendly policies, career development programmes and employer engagement and outreach through EDI committees and advisers was mixed; there was evidence of positive impact in some areas (for example, improving the experiences and support for women and underrepresented staff through networks and affinity groups) but not in others (for instance, the lack of senior managers in such affinity groups limits the utility of these networks in improving career progression); these examples suggest the importance of understanding the impact of not only creating new policies and programmes but also considering their uptake and sustainability
- successful interventions tended to involve collaboration across and within organisations (such as forming internal support networks or creating communities of practice around an external scheme or framework); they also have commitment from senior management and embed EDI awareness and initiatives into organisational culture
- common features of less effective interventions were a lack of clarity around policies or objectives and how they are implemented, and a reliance on the 'business case' for diversity to drive change.

Evaluation methods need strengthening, but mixed methods have value

Only a third of publications examined reported a clear effect size and few evaluations clearly defined and measured outcomes of interventions beyond correlation with high-level indicators of equality (such as staff recruitment rates for underrepresented groups). However, we note the value of qualitative studies in understanding the nature of EDI work and analysing the 'why' behind successes and limitations of approaches. We also note the practical challenges of resourcing in-depth evaluation alongside diversity work.

Challenges in measuring sector progress

Variation in practice of collection, comparison and use of EDI data impacts on the ability to paint a 'global' and holistic picture of EDI across R&I sectors, as well as the rigour of evaluation. Continued investigations into different models of 'measuring' diversity and inclusion are recommended, whilst acknowledging the importance of cultural and legal contexts.

Key recommendations

This report details specific recommendations for policy makers, R&I funders, employers and researchers. These range from actions or approaches to support successful practice, to more effective evaluation. Key recommendations include:

- co-ordination of further reviews and research to fill 'gaps' in the evidence base (Chapter 4: Previous studies)
- investigate new tools and resources for comparative analysis of EDI progress, sensitive to varied disciplinary, sector and national contexts; an exploration of attitudes to data collection across a wide range of characteristics in international research and innovation may be required (Chapter 7: Enhancing data and disclosure)
- foster collaboration within and between organisations to fill skills and resource gaps relating to evaluation, avoid duplication and share best practice (Chapter 6: Measuring success; Chapter 8: Leading organisations)
- ensure alignment between organisational EDI strategy, messaging and practice, and between the attitudes of senior leaders and those delivering EDI interventions (Chapter 5: What works)
- reflect on definitions of 'leadership' and future methods for recognising and rewarding innovation, commitment and collaborative practice (Chapter 8: Leading organisations).

Detailed recommendations for policy makers, research funders, employers and researchers are included at the end of this review.

Chapter 2: Introduction

The background of the slide is a dark blue field filled with numerous thin, glowing lines in shades of cyan and magenta. These lines are mostly diagonal, creating a sense of motion or data flow. In the upper right corner, there is a complex, swirling pattern of these lines. A large, solid yellow shape, consisting of a rectangle and a triangle, is positioned in the lower half of the slide, partially overlapping the light streaks.

2.1 Project background, scope and research questions

UKRI commissioned Advance HE to undertake a review of challenges and interventions (strategies, policies and programmes) used to address current EDI challenges in the international R&I sector.

This exploratory study, conducted over 15 weeks, provides an overview of what is known about EDI interventions and the antecedent challenges they were designed to address. Research questions were identified as:

- **previous studies** - which organisations have previously reviewed and explored the key challenges for EDI in the international R&I landscape?
- **what works?** - which organisations are leading in terms of EDI in international R&I?
- **measuring success** - of the international interventions that have been implemented by organisations comparable to UKRI, which have proven effective, or less effective, and why?
- **leading organisations** - how is the effectiveness of EDI interventions measured and are there methods particularly useful for the international R&I landscape?
- **enhancing data and disclosure** - how can EDI data capture and disclosure rates in the international R&I landscape be improved?

The review did not intend to offer a comparative assessment of different regional equality protections or 'levels' of inclusivity and diversity (although the challenges of doing so are discussed later). Instead the focus is on understanding and identifying key practices which may be of interest to the work of:

- R&I policymakers (such as government bodies or sector agencies)
- R&I funders
- employers (industry or academia)
- Researchers and innovators (in industry or academia).

The focus of the review was broad and considered work that had taken place internationally within the context of higher education and research institutes, learned societies, government agencies, charities and the voluntary sector, and private companies. Data collected included recent academic papers and grey literature, supported by a short targeted Call for Evidence. Research focused on work that has taken place since 2010 in organisations that share UKRI's role as a research funder, a leader in R&I policy, outreach and public engagement, and a large employer.

This review collated the evidence gathered and applied an evaluation framework to enable analysis and synthesise findings from across the different types of data source. Results from this synthesis

present an evidence base for 'what worked?' and 'what did not work?' in response to a range of EDI challenges.

UKRI also commissioned a concurrent review which examined these issues from a UK perspective. The Global Institute for Women's Leadership at King's College London undertook a third project focused specifically on 'bullying and harassment'. As it is not always possible or appropriate to separate bullying and harassment from wider issues of EDI, or the UK from the international context, we recommend that the three reports are read in conjunction with each other. The full background to all three reviews can be found at www.ukri.org/about-us/equality-diversity-and-inclusion/strengthening-our-approach/

A fixed-term Advisory Group was recruited to provide external advice on the scope and methodology of the two reviews. The Advisory Group included members with backgrounds in research and innovation and/or EDI, who worked across the higher education, voluntary and public sectors. For a full list of Advisory Group members see the acknowledgments.

2.2 The international context

Understandings and definitions of 'equality, diversity and inclusion' will vary across national, legislative and cultural boundaries, and we acknowledge this throughout this international review. Informed by engagement with stakeholders at a Challenge Workshop, as well as working with our Advisory Group, we focused on an understanding of EDI work which seeks to:

- **tackle instances of underrepresentation, differential needs and systemic disadvantage:** in R&I this could present as unequal representation compared with local or 'pipeline' populations in senior academic positions, leadership and research grants and citations, and by discipline
- **support inclusion and reduce the impact of bias and discrimination on individuals and groups:** this may mean addressing different experiences of discrimination, bias and harassment within employment, postgraduate study or R&I.

Some of these challenges will present globally or across regions. For example:

- underrepresentation of women in research (for example, She Figures 2018 from the European Commission, 2019)
- underrepresentation or differential experiences of specific ethnicities within doctoral study (for example, ACE 2019; Canadian Institutes of Health Research (CIHR) Institute of Gender and Health, 2012; Webber and González Canché, 2015)
- experiences of bias, harassment, racism and sexism (Arday and Mirza, 2018; Brunnsma, Iyall Smith and Gran, 2015; National Academies of Sciences, 2018)

- additional labour on disabled and/or Deaf/deaf people arising from structural inequalities and stereotypes in academia (Inckle, 2018; Sang, 2017; Woodcock, Rohan and Campbell, 2007)
- gendered differences in research publications (e.g. Mayer and Rathman, 2018; Symonds et al., 2016) or patent applications (Hunt et al., 2013; OECD, 2017).

Other challenges may be more specific to context, for example the underrepresentation of indigenous peoples in research and academia, or unequal caring responsibilities in the context of different national infrastructures.

Different historical and legal contexts also impact our understanding of equality 'monitoring', culture change and majority or minority experiences. A wide range of individual and group identities and interactions may be considered through an EDI 'lens'. Therefore this review did not specifically restrict itself to challenges relating to characteristics as conceived of and protected in UK law. However, a core focus on characteristics of gender and sex, age, disability, race and ethnicity, sexual orientation, transgender identity and history, pregnancy, maternity and childcare, and socio-economic status formed a useful starting point and framework.

As understandings of EDI aims and objectives vary, so too will the protections for different identities and backgrounds, and the tools available to achieve these. For example, some jurisdictions will permit 'targeted' actions (also known as 'affirmative' or 'positive actions') to ensure equal outcomes for certain groups, but such interventions may not be considered welcome, appropriate or legal in other contexts (see Archibong et al., 2009).

The drivers for EDI change may of course vary within sectors and organisations, as well as regional and national contexts. Common drivers include: the 'business case' (increased productivity, staff retention, reputation); ethical and human rights arguments; and legal or regulatory provisions. In R&I the role of the research funder and policy maker as a key driver of change is of particular interest and will be explored throughout this study. In the R&I context, staying informed about EDI (beyond diverse and inclusive teams) can also support the efficacy of R&I: from gendered implications of health studies (CIHR, 2012) to racialised bias in technology innovations (Barocas and Selbst, 2016).

This review sought to understand not only the range and effectiveness of existing EDI interventions in international R&I but also how that effectiveness is evaluated, while staying mindful of the limitations of comparability when interventions take place in different cultural contexts (Jonsen, Maznevski and Schneider, 2011; Peretz, Levi and Fried, 2015).

Findings from this review will shape the development of UKRI's EDI strategy and be shared widely with others in the sector to expand the evidence base on EDI interventions and facilitate the

sharing of good practice. To aid future directions we provide a summary of key recommendations for a range of stakeholders, including areas for future research, in the final section.

2.3 Reflexivity

The practice of EDI research is not value-neutral and, as with other organisations, Advance HE's work in this area will invite a degree of subjectivity and bias. We also note Advance HE's influence or engagement with some of the organisations or interventions discussed in this report as the organisation which also owns the Athena SWAN gender equity charter.

As with most EDI research, the frames of reference brought to this review were also likely affected by researchers' identity characteristics and organisational and academic backgrounds. Taking account of these potential limitations, we recognised that:

- the review cannot present a fully objective or unbiased account of EDI interventions in R&I
- measures were required to address, as far as possible, the research team's inherent subjectivities and biases.

This review therefore followed a rigorous methodology (see below) that was intentionally designed to counter subjectivities and biases as much as possible during the short timeframe for the review. As a way to further diversify the design of the methodology and ensure that input went beyond Advance HE staff and associates, an Advisory Group was recruited and two meetings were held to discuss the review's search terms, inclusion and exclusion criteria, targeted grey literature search and evaluation framework design.

However, even with these measures in place, the methodology followed meant that this review would not identify and analyse all possible sources related to EDI interventions internationally. Rather, the 15-week study identified a range of EDI interventions, via three data collection streams, and then used an evaluation framework to assess their effectiveness, evaluation methods used and approaches to data collection. Many activities and reports will raise questions (of long-term evaluation and of 'missing' data) which could only be fully answered by further primary research beyond the scope of this review. Final reflections on methodological limitations and opportunities for future research are highlighted throughout the report and in chapter 9).



Chapter 3: Methodology

3.1 Overview

This review aimed to answer a range of research questions relating to understanding the evidence base around EDI interventions in the R&I landscape.

Building on Advance HE's extensive knowledge and expertise in EDI, our methodology consisted of the following elements:

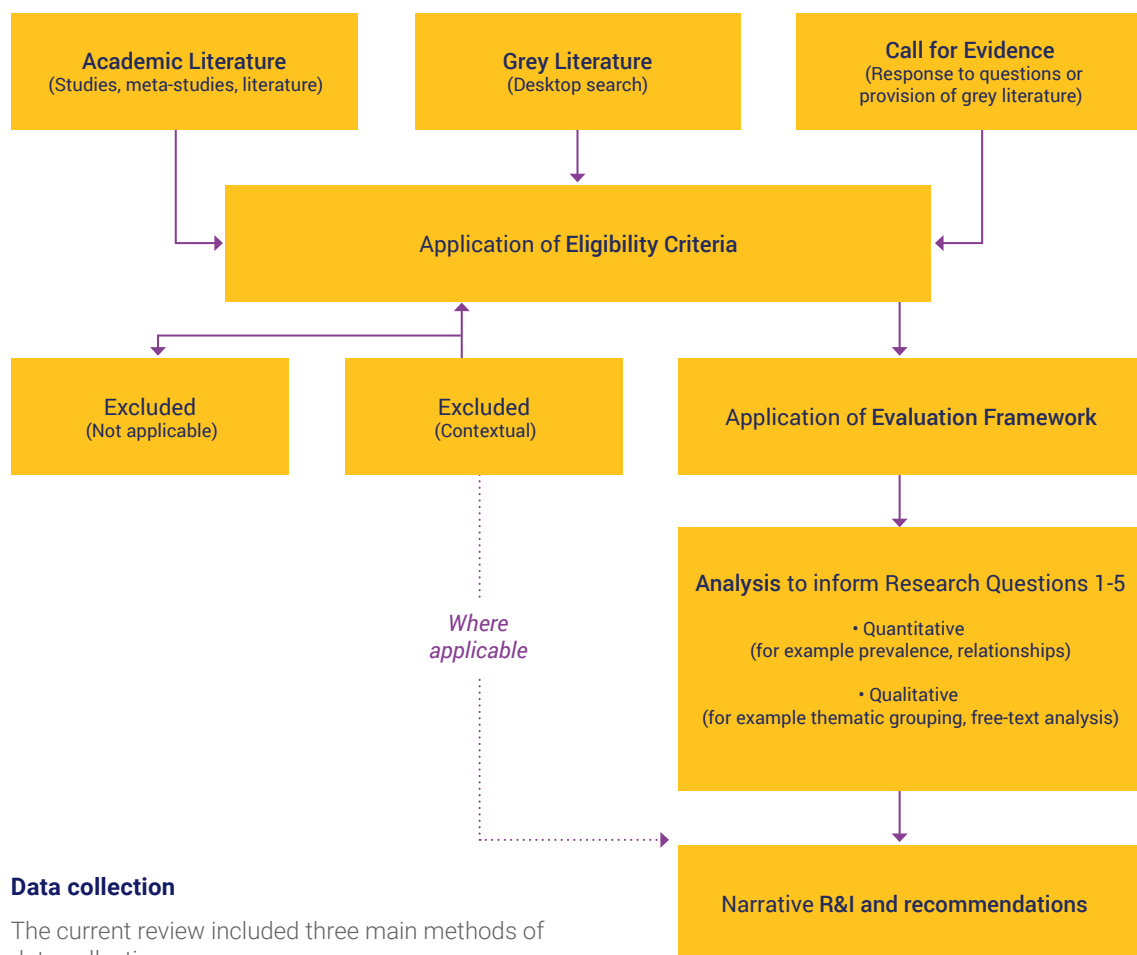
- a desktop search and curation of available literature (academic and grey)
- collection and development of further data through a targeted Call for Evidence with organisations in the UK R&I base.

- The formation of an evaluation framework to qualify existing literature.
- Analysis (quantitative and qualitative) informed by the evaluation framework.
- Synthesis, followed by development of recommendations.

Our approach was informed by our Advisory Group and some piloting of methods.

A high-level overview of data streams is provided below (figure 3.1), followed by a detailed description (with references to appendices where appropriate).

Figure 3.1. Summary of methodological approach.



Data collection

The current review included three main methods of data collection:

- an extensive search of existing academic and grey literature using online search databases
- a targeted search for grey literature which included mining the websites of organisations known to have focused on EDI issues for published grey literature
- a Call for Evidence which involved primary data collection from international R&I organisations.

3.1.1 Academic and grey literature database search

The search for existing literature was conducted via three main databases (EBSCO, Scopus and OpenGrey) and used Boolean search terms related to (i) EDI, (ii) interventions, and (iii) R&I

(see appendix A). Due to the short timeframe for data collection, it was most practical for the UK and international research teams to use the same search terms and then sort eligible sources across the two reviews.

Table 3.1. Summary of the total number of sources identified through each database search

Database	Total number of UK and international sources identified
EBSCO	3,011
OpenGrey	684
Scopus	2,295
Total	5,990

3.1.2 Targeted grey literature search

Alongside the database searches, a targeted search of around 60 organisational websites was undertaken to locate publications related to EDI interventions in the international research base (see appendix B for the list of websites). While the search could not be exhaustive, it was designed to encompass a broad sample of organisations involved in the R&I landscape and private or public sector organisations doing work on EDI. The following types of organisation were included in the search:

- research councils and other funding bodies
- charities and non-governmental organisations (NGOs)
- government or quasi-governmental science, research or innovation bodies
- HE sector agencies
- learned societies
- private sector companies

Searches on these websites involved using any search function available with simple search terms such as 'equality', 'diversity' or 'inclusion'. Where a search function was not available, publications lists were scanned for relevant material.

3.1.3 Call for Evidence

The Call for Evidence was circulated for a short period with Advance HE and Advisory Group contacts and relevant mailing lists. A form was developed to capture key information from institutions about interventions they had undertaken (see appendix C.2). Questions were designed to be flexible so that respondents could share different types of interventions and to prompt institutions to return information that would help answer the project's research questions.

3.2 Data reduction

3.2.1 Inclusion criteria

To limit the scope of the review and most effectively answer the review's research questions, inclusion and exclusion criteria were applied to all sources across all strands of data collection (see appendix E.1 for a detailed list of inclusion criteria and details regarding their application).

Sources that met the following criteria were included in further analysis:

- published on or after 1 January 2011 (to capture the latest innovations and to mirror the scope of the UK report)
- discusses at least one identity characteristic from a list which used the UK Equality Act 2010 protected characteristics as a base, in addition to other factors
- published in an authenticated source (an academic journal, book, organisation website etc.)
- evaluates an EDI intervention in an empirical manner, or is a review or meta-analysis of EDI interventions
- relevant to R&I or to the funding, practice or communication of R&I
- available in English
- discusses interventions conducted outside of the UK, or includes the UK in a wider international scope.

The research team acknowledges that some of the excluded publications (blogs, book reviews, legal cases etc.) may include academic or empirical content. However, the high degree of variability in the quality and quantity of information present in these sources placed them beyond the timeframe and rigour of the current review. It is worth noting

that review articles (such as those describing current EDI challenges or barriers) were excluded from the following analysis using the evaluation framework (described in section 3.3) but included in our discussion of which organisations have undertaken reviews of EDI challenges and interventions in section 4.3.

A team of four researchers (two from the UK review, two from the international review) manually reviewed the 5,990 sources identified. In total, 5,881 sources were excluded from the final international dataset. Of these, 1,494 were duplicate sources (sources that were identified through more than one of the data collection methods). An additional 886 sources were inaccessible to the research team (that is, behind a paywall or in a database that the research team could not access) and thus excluded from further analysis. Finally, 2,126 sources were excluded for not being an empirical evaluation

or a review or meta-analysis and 1,273 were excluded for not discussing at least one protected characteristic identified within the Equality Act 2010, or EDI in general. Finally, during the application of the evaluation framework (described in section 2.4.3), an additional 24 sources were removed from the list of eligible sources because they did not include an evaluation of an EDI intervention or were not accessible (in other words, the articles were published in journals that were not available to the researchers through their current EBSCO subscription, were not published online or were archived and no longer available online). Of the remaining 187 sources, 78 applied to the UK context and were as such removed from this analysis. This resulted in a final sample of 109 sources representing 130 interventions (see table 3.2 for a breakdown across data collection strands).

Table 3.2. Eligible sources and interventions

across data collection strands

Strand	No. of eligible sources	No. of interventions
Academic and grey literature database search	82	98
Targeted grey literature search	13	15
Call for Evidence	14	17

3.2.2 Reliability of inclusion criteria

To ensure that the eligibility criteria had been applied in a similar manner, a subsample of 10% of all identified sources was double-coded by a the fifth researcher who was blind to which sources had been labelled as eligible by the research team. Overall, each criterion was applied in a similar manner with both the research team and fifth researcher disqualifying: roughly 11% of subsample sources as duplicates (as the same source could be identified through the different streams of data collection); 9-10% as being inaccessible; and 63-64% as not referring to an agreed identity characteristic, socio-economic status or EDI in general, or being an empirical evaluation, review, meta-analysis or gap analysis of EDI interventions, or evidence-based recommendations (see appendix E.3 for a summary). It should be noted that the last two criteria were combined as many of the sources met both of them.

3.3 Evaluation framework

3.3.1 Design and application

We developed and applied an evaluation framework to extract the information in each source that was pertinent to the five research questions addressed in the current review. The evaluation framework (see appendix F.2) involved applying labels or descriptors to each source's content, such as which identity characteristic was examined, what type

of intervention was evaluated and which area of UKRI's work this intervention could be applied to. By converting the sources into a common rubric, this quantification facilitated the application of a number of synthesis techniques including tabulation (in other words, how many sources employed a given evaluation method, or discussed a specific protected characteristic), as well as grouping and clustering studies according to their applicability to each of the current research questions (for example, sorting the database by evaluation method to identify most frequently applied methods in chapter 5). The evaluation framework also provided us with the space to highlight important information or discussions within each source and tease out content for additional qualitative analysis. The qualitative approaches applied within the current review not only included the identification of themes (such as what types of intervention are presented in the current database), but also the triangulation of methodologies and concepts across both the qualitative and the quantitative information present in the evaluation framework to determine how the interventions work, why they work and for whom.

As many of the sources described more than one intervention, we applied the evaluation framework to the individual interventions rather than the source. In other words, while the final international sample included 109 sources, the total number of interventions evaluated in the current review was 130.

3.3.2 Reliability

A similar process to that described for the reliability analysis of the inclusion criteria was undertaken for the application of the evaluation framework. Specifically, a subsample of 10% of all eligible sources (including some from the academic and grey literature, and the Call for Evidence) was double-coded by a fifth researcher. We used percent agreement between the research team and the fifth researcher to establish reliability, with a cut-off point for satisfaction being at least 80% agreement. Overall, percent agreement was acceptable for all categories except type of data and the Maryland Scientific Method Scale (see appendix F.3 for the percent agreement for each variable). This was likely due to a large degree of variability in how methods and results were reported (or not) in many of the sources. Specifically, 42 (32.3%) of sources did not measure or report an outcome, making it difficult to ascertain what kind of data was collected and how it was analysed. Given the low percent agreement on these two categories, this data was reviewed and recoded by the research team as a unit prior to further analysis.

3.4 Descriptive analysis of the international dataset

The following analyses are based on the 130 interventions presented in the 111 academic,

grey and Call for Evidence sources. An initial series of descriptive analysis of the quantitative components in the evaluation framework was undertaken to inform the content of our results sections. Many of the categorical variables in the evaluation framework were not mutually exclusive; as such we relabelled the raw data to represent the original categories as well as how these categories were examined in combination (for example, an intervention that is relevant to UKRI's work in public engagement and R&I policy).

3.4.1 Geographic coverage

A large number of interventions stemmed from research conducted within the US (52 out of 130, or 40.0%). Roughly one third of the remaining interventions were from Europe (31 interventions, or 23.9%), whether this be from research conducted within a single European country (25 sources) or multiple European countries (six sources). There was representation of work from Australia (six interventions), Canada (five interventions) and New Zealand (four interventions). However, the representation of sources from other regions including Africa (five interventions, three in Tanzania), South America (no interventions) and Asia (four interventions, all from India), was relatively low in comparison (see table 3.3 for a summary).

Table 3.3. Summary of geographic regions covered in current sample

Geographic region or country	No. of interventions	%
Australia	6	4.6
Austria	1	0.8
Belgium	1	0.8
Canada	5	3.8
Multinational (that is, referring to two or more countries, with at least one country outside of Europe)	23	17.5
Ethiopia	1	0.8
Europe	6	4.6
France	2	1.5
Germany	10	7.7
India	4	3.1
Ireland	1	0.8
Kenya	1	0.8
Netherlands	2	1.5
New Zealand	4	3.1
Norway	1	0.8
Spain	1	0.8
Sweden	2	1.5

Geographic region or country	No. of interventions	%
Switzerland	4	3.1
Tanzania	3	2.3
US	52	40.0
Total	130	100.0

3.4.2 Coverage of identity characteristics

- Over half of the interventions analysed were primarily related to promoting gender or sex equality (57.7%, 55 interventions), with other identity characteristics receiving considerably less attention in the eligible academic literature.
- 17 interventions (13.1%) looked at ethnicity, race or nationality.
- 10 interventions (7.7%) examined sexual orientation.
- 10 interventions (7.7%) considered disability (including mental health).
- Nine interventions (6.9%) focused on age.
- Five interventions (3.8%) evaluated an intervention related to pregnancy or maternity leave, or childcare.
- Two interventions (1.5%) considered religion and belief.
- 21 interventions (16.2%) looked at EDI in general.

It is worth noting here that an individual intervention could cover more than one identity characteristic (that is the categories were not mutually exclusive).

For example, an additional six interventions looked at both gender and race, eight interventions investigated general EDI combined with a specific protected characteristic, and four considered three or more protected characteristics. We also note a limitation here for 'general' EDI: the interventions' definitions of this varied, either referring to an aspect of EDI without detail (such as 'diversity training') or as an umbrella term for a range of identity characteristics. It is possible that some also may be 'euphemistically' using this term to relate to ethnicity or race (and indeed, some of our sources discussed the implications of 'colourblind' EDI approaches, for example, Williams, 2018; Wilton, Good, Moss-Racusin and Sanchez, 2015).

3.4.3 Area of EDI investigated and sector or disciplinary focus

The majority of interventions investigated aspects of an individual's career (for instance, how they were recruited or the factors that influence applying for or taking up a post, or factors related to promotion or leave policies) either on their own (33 out of 130 interventions, or 25.4%) or in combination with aspects of an organisation's culture or wellbeing of staff (an additional 38 interventions, or 29.2%) (see table 3.4 for a summary).

Table 3.4 Summary of EDI areas covered within current sample

Area of EDI focus	No. of interventions	%
Careers (recruitment, promotion, leave policies etc.)	33	25.4
Culture and wellbeing (inclusion, experiences etc.)	10	7.7
Outreach and public engagement (community work, events etc.)	4	3.1
Data (equality monitoring, increasing disclosure etc.)	4	3.1
Funding (scholarships, grant awards etc.)	5	3.8
General policy, practice or governance	6	4.6
Careers and culture combined	38	29.2
Three or more areas covered	15	11.5
Other: mixed range of EDI areas	6	4.6
Other: access, retention and employability of students	3	2.3
Other: reduction of bias, improved perceptions	6	4.6
Total	130	100.0

The majority of interventions (54.6%, or 71 interventions) focused on EDI within the context of higher education, or research (all disciplines) or STEM sectors. 17.7% (23 interventions) explored EDI in sectors related to business, management and leadership. Only one intervention looked at EDI issues in healthcare (0.8%), one within a charity, community or public service (0.8%), and none considered aspects of wider education, teaching and learning, though the latter could be due to lack of clarity between categories (for example, teacher training recorded as 'higher education'). Twelve

interventions (9.2%) covered EDI in multiple sectors or disciplines, and 16 interventions (12.3%) did not apply to a specific sector or discipline.

3.4.4 Relation to UKRI and R&I landscape

The evaluation framework also aimed to identify how each intervention may contribute to UKRI's EDI policies and initiatives, and where these contributions would fit with regards to UKRI's constituent organisations. The frequencies of interventions across the different areas of UKRI's work are summarised in table 3.5.

Table 3.5. Summary of UKRI's areas of work covered within current sample

Area of UKRI work	No. of interventions	%
Research funding or funder(s)	9	6.9
Employers	22	16.9
Research funding or funder(s) and employers	9	6.9
Research funding or funder(s) and R&I policy	9	6.9
Employers and R&I policy	11	8.5
R&I policy	5	3.8
R&I policy and public engagement/outreach	5	3.8
Employers and public engagement and outreach	7	5.4
Public engagement and outreach	2	1.5
Combined three areas	48	36.9
Combined four areas	1	0.8
Other	2	1.5
Total	130	100.0

3.4.5 Type of intervention

As seen in table 3.6, different types of intervention were represented in the international dataset, with multiple interventions covering training

and development, strategy or policy change, positive or affirmative action, and learning and teaching resources. Notably, a large proportion of interventions (49, or 37.7%) included more than one type (as defined in the framework).

Table 3.6. Summary of intervention types covered within current sample

Type of intervention	No. of interventions	%
Training and development	17	13.1
Mentoring and coaching	3	2.3
Strategy or policy change	16	12.3
Awareness raising	2	1.5
Organisational review/assessment of EDI	2	1.5
Learning and teaching resources	6	4.6
Positive or affirmative action	7	5.4
Networks, sponsors and champions	3	2.3

Type of intervention	No. of interventions	%
Family-friendly policies	4	3.1
Data collection and tools	4	3.1
Other	17	13.1
Two intervention types	4	3.1
Three intervention types	29	22.3
Four intervention types	7	5.4
Multiple intervention types (total unclear)	9	6.9
Total	130	100.0

3.4.6 Methodology employed and data captured

For interventions citing empirical results, the evaluation framework extracted two key pieces of information related to their methodology: (i) what type of design was employed and (ii) what type of data was collected. The evaluation framework listed eight explicit types of study design as well as options for interventions that did not state a clear methodology, were not an empirical evaluation, or

were less frequently adopted methods in academic literature, such as national figures, audits and document or discourse analysis (summarised in table 3.7). One out of 10 interventions did not include sufficient information to categorise their methodology (10.8%). However, of those that did provide information on their method, a large proportion adopted either a between-groups (cross-sectional) design (30 interventions, or 23.1%) or a mixed-methods design (23 interventions, or 17.7%).

Table 3.7. Summary of study designs covered within current sample

Study design	No. of interventions	%
Within-groups or longitudinal	4	3.1
Between-groups or cross-sectional	30	23.1
Time series analysis	4	3.1
Case study (or case studies)	9	6.9
Qualitative analysis of interviews or journals	12	9.2
Meta-analysis	5	3.8
Other quantitative analysis (for example, analysis of sector- or staff-level data)	12	9.2
Other qualitative analysis (for example, discourse analysis)	2	1.5
Mixed methods	23	17.7
Conceptual article or not applicable	15	11.5
Unknown	14	10.8
Total	130	100.0

With regard to the type of data collected and analysed, the majority of interventions included either quantitative data (52 interventions, or 40.0%) or both qualitative and quantitative data (34 interventions, or 26.2%). An additional 24 interventions (18.5%) collected qualitative data only. There were 20 interventions for which the type of data collected was unclear or not presented.

3.4.7 'Robustness' of methodology

Finally, in order to quantitatively assess the 'robustness' of the evidence base of each intervention employing an empirical design, the evaluation framework included a simplified version of the Maryland Scientific Method Scale, along with open-text descriptions of the method, data and outcomes. This scale is intended to rate the level of scientific 'rigour' in the methodology adopted by each intervention. The frequency of interventions across scale levels is summarised in table 3.8.

Table 3.8. Summary of the Maryland Scientific Method Scale within current sample

Level of the Maryland Scientific Method Scale	No. of interventions	%
1: Correlation (for instance, departments with a female leader have more female staff)	29	22.3
2: Before and after assessment, with no control of conditions (for instance, female staff in a department increased after the appointment of a female leader)	21	16.2
3: Before and after assessment, with experimental conditions (for instance, female staff in a department increased after the appointment of a female leader, female staff in a department did not increase after the appointment of a male leader)	7	5.4
4: Before and after assessment, with multiple experimental conditions (for instance, as with level 3 but with additional controls for gender culture in department and individual backgrounds of staff)	11	8.5
5: Randomised control trial (RCT)	4	3.1
Not applicable, unclassifiable or unclear	58	44.6
Total	130	100.0

The large number of interventions classified as 'not applicable or unclear' (44.6%, or 58 sources) highlights the limitations of this scale. Specifically, the scale does not:

- cover all methodological approaches (for example, qualitative approaches, or other quantitative methods such as the analysis of sector-level figures)
- distinguish between studies that employ qualitative versus quantitative methods
- consider whether a study has adopted a mixed-method approach.

The limitations of using this scale (including equality-related critiques of such hierarchical approaches to assessing research) are discussed in further detail in chapter 6 alongside our presentation of methods for evaluating intervention effectiveness.

Chapter 4: Previous studies

The background features a complex network of glowing blue lines and nodes, resembling a molecular structure or a data network. The nodes are small spheres in various colors (yellow, orange, red, white) connected by thin blue lines. A large, solid yellow geometric shape, resembling a stylized 'Y' or a corner, is positioned in the lower right quadrant, partially overlapping the network.

Which organisations have previously reviewed and explored the key challenges for EDI in the international research and innovation landscape?

4.1 Overview

Exploring EDI challenges in employment, academia and R&I is not a new endeavour. Organisations that have undertaken this research have taken different approaches. For example

- **perspective:** some will be outward-facing (learning from others), others more internal (understanding the organisation's own challenges)
- **scope:** for example, or example, regional, national or international, or focusing on an aspect of R&I such as leadership or study

- **EDI focus:** for example, on specific identity characteristics (such as gender) or wider 'diversity' issues or measures
- **aims:** sharing best practice, identifying challenges, monitoring 'progress'.

Examples of recent reviews are included below to demonstrate some of the breadth of such studies, with some findings and outputs.

Data collection on EDI issues on a global scale is challenging due to different data collection and definition methods (see further discussion later). Attempts to collate useful information for both an overview and (potentially) benchmarking:

OECD: Technology and Industry Scoreboard (2017/biennial) (OECD, 2017)

International	Various demographics (including gender and education)	Science, Technology & Innovation
---------------	---	----------------------------------

Collates key datasets on the size and certain demographics (primarily gender) in research and industry knowledge production sectors. It highlights, for example, that women made up less than a third of all 'tertiary' graduates in natural sciences, engineering and ICT globally between 2005 and 2015, and provides gendered breakdowns of the doctoral holders within a population (all subjects).
3,011

Link: <http://www.oecd.org/sti/scoreboard.htm>

Organisation for Women in Science in the Developing World: National Assessments on Gender and STI (from 2012)

International	Gender (women)	Science, Technology & Innovation
---------------	----------------	----------------------------------

National and regional gender reviews providing qualitative and quantitative data on different regions, focusing on key areas of women's participation in the economy, education, social status, health and technology, all as important contextual

Link: <https://owsd.net/resources/data-collection-and-analysis>

Comparative studies across regions and identity characteristics are less common:

Innovate UK: Global Review of Diversity and Inclusion in Business Innovation

International	Ethnicity and race, disability, LGBTQ+, gender	Business innovation: all disciplines
---------------	--	--------------------------------------

Reviewed EDI challenges in business innovation and entrepreneurship in 10 countries. Key synergies and differences in approaches (for example targeted initiatives, legal mandates, industry aims) were discussed. The report also contains recommendations for practical approaches (networks, summits) to enable knowledge exchange on EDI programme implementation and evaluation.

Link: Global review of diversity and inclusion in business innovation

Forbes: Global Diversity and Inclusion: Fostering Innovation through a Diverse Workforce (2011)

International

Ethnicity and national origin, race and colour, age, disability, sexual orientation, nationality, gender, veteran status

Large global enterprises

Case studies and responses from over 200 organisations, highlighting the link between staff diversity and innovation of practice, products and decision making, and noting:

- internal challenges facing EDI efforts (including attitudes to EDI)
- external challenges facing EDI efforts (financial climate)
- noted that most organisations felt that most EDI progress had taken place around 'gender' (54%), whereas 'disability' and 'age' were felt to be areas needing most improvement.

Link: https://www.forbes.com/forbesinsights/innovation_diversity/

Sharing of best practice across national systems is more common:

Te Kāhui Amokura: Tauira Māori Initiatives - Sharing Good Practice in New Zealand Universities (2018)

New Zealand

Indigenous identity

Study: all disciplines

Overview of best practice and approaches for tauira Māori (Māori students) across New Zealand universities.

Link: <https://bit.ly/2Osdgfg>

Organisational reviews examining a holistic concept of EDI (importing a wide range of identity characteristics, the broad range of R&I careers and environments, and the 'pipeline' from postgraduate study) are more limited. Literature has predominantly focused on a smaller number of identity characteristics (for example, gender or, more specifically, women) or on specific elements of R&I such as academic leadership (Nyoni, He and Yusuph, 2017). There are, however, examples of organisational or collaborative self-evaluations and reviews from research funders or agencies, and examples of particular frameworks which support such exploration. Some of these will be explored in greater detail in following chapters, but examples of work under different themes are presented below.

4.2 Key themes emerging from organisational reviews

4.2.1 Gender Focus

'Gender equality' remains a priority focus of much of the literature, as well as of relevant R&I organisations and institutions, though we note that gender equality efforts within our literature review have tended to focus more specifically on women's experience, progression and reputation (and with a binary conception of sex and gender, that is to say men and women), with fewer examples of specific references to the experiences of men and/or wider gender identities. Reports that are clear on their focus on women include:

Vitae (with Research Councils UK and the Science and Engineering Board of India): Global Survey on Equality and Status of Women in Research (Metcalf and Day, 2016)

International

Gender (women)

Research: All disciplines

Aimed to inform the Global Research Council (GRC) meeting in 2016. The survey incorporated five regions and noted (among other things): awareness of gender equality was high but (with the exception of the Africa region) leadership boards were still predominately male; countries with 'less developed' research systems recognised inclusion as a means of increasing recruitment; and 'more developed' research systems were more likely to have clearer policy positions and equality plans. A wide range of practices were identified, and key recommendations included greater commitment to data collection and setting 'key indicators' on gender equality. The GRC launched its Statement of Principles and Actions Promoting the Equality and Status of Women in Research the same year.

Link: <https://www.globalresearchcouncil.org/library/>

The focus on gender over (or in isolation from) other characteristics may be attributed to a number of factors. The generally binary conception of gender in many of the data sources may mean that key differentials between the representation of women and men are easier to analyse and communicate as a cross-border issue: for example, in leadership (Nyoni et al., 2017) and governance (Casey, Skibnes and Pringle, 2011), specific disciplines (Wang and Degol, 2017), or patent applications (OECD, 2017). This in turn may lead to a wider evidence base for a policy focus. Underlying global frameworks such as the UN Sustainable Development Goal for 2030 for gender (with specific indicators and targets) also speak to a shared global consensus on the direction of change needed.

4.2.2 New horizons: intersectionality, disciplinary focus

A large number of the sources examined focus on STEM industries and subject knowledge, or on wider organisational perspectives on diversity. Explorations which highlight and examine disciplinary differences, however, can help understanding of issues such as career pipeline and provide evidence for considering the context of different initiatives (as we will discuss later). Reports which seek to take an intersectional approach can also highlight compounding inequalities and underrepresentation:

American Council on Education (ACE): Race and Ethnicity in Higher Education (2019)

US	Race and ethnicity (with gender, age)	Study: all disciplines
----	---------------------------------------	------------------------

Overview of race and ethnicity and access to or success in higher education in the US. Notable findings relevant to the R&I researcher career pipeline include the following:

- key differentials indicated by intersectional (ethnicity and gender) analysis of doctoral holders (all disciplines) indicated key differentials; for example, analysis shows that while 2.5% of men aged 25 and under held a doctoral degree compared to just 1.5% of women, there are wide variations across racial and ethnic groups (for example, the largest gap is between Asian men at 6.3% and 0.1% for Native Hawaiian or other Pacific Islander women, though caveats were in place on these numbers)
- ethnic differences in choice of graduate college (university) and loan debt (for example, black students were overall most likely to take out study loans regardless of income, and a smaller share of black doctoral students received grant aid compared to other racial or ethnic groups)
- ethnic and national trends in choice of doctoral discipline (for example, the majority of international students studying in the US were focusing on STEM fields and a larger proportion of Hispanic students and 'students of more than one race' were focusing on 'social and behavioural sciences' than white, black or international students).

Link: <https://www.equityinhighered.org/resources/report-downloads/>

4.2.3 Organisations sharing and generating knowledge through networks

Many organisations are seeking to produce, extend or share their EDI knowledge through formal collaborations or participation in shared frameworks and can provide insight into specific contexts.

Further to the Organisation for Women in Science in the Developing World review above, other

networks or research funders have developed a range of resources (again, gender is a priority) including the GENPORT European portal for gender diversity literature, the Te Kāhui Amokura Resource Inventory (2017) for Maori and Pasifika academic and research support, and the German Research Foundation (DFG) toolbox:

DFG: Gender Equality in Research (DFG, 2017)

Germany	Gender	Research & PGR study: All disciplines
---------	--------	---------------------------------------

Compilation of a searchable database (as a sector resource) on a range of EDI best practice in gender equality across a range of disciplines and target groups. Includes an element of review as it makes a judgment on particularly 'innovative models'.

Link: https://instrumentenkasten.dfg.de/index_en.html

GENOVATE: Gender, Culture and Climate Assessment (2016) and Developing Learning Circles (GENOVATE, 2016a)

Europe

Gender

R&I and scientific decision making bodies

Summary reports on organisational approaches to gender equality under the GENOVATE consortium, discussing issues of gender representation, allocation of resources and work-life balance. Further areas for development included the need to integrate gender equality within organisational functions, provision of data to support policy making, and ensuring sustainability and long-term actions.

Other outputs from GENOVATE included a synthesis of practitioner and stakeholder perspectives under a 'learning circles' approach.

Link: www.genovate.eu/resources/gender-culture-and-working-climate-assessment-report/

Link: www.genovate.eu/media/genovate/docs/deliverables/GENOVATE-National-Learning-Circles-Brochure.pdf

4.2.4 Beyond organisational perspectives

In acknowledging the role of collaborations and networks, we have challenged the formation of our original research question asking 'which organisations have previously reviewed and explored' EDI in R&I. While some of these collaborations are more formalised through joint reports or collection of initiatives, we acknowledge that explorations and knowledge exchange may also be undertaken more informally or with actors outside of R&I organisation settings. For example, individual practitioners or researchers undertaking 'diversity work' (Ahmed, 2015) and those with lived experience of key challenges (for example, Chronically Academic, the Cite Black Women collective) may be a key source of knowledge on EDI challenges and interventions. Whilst not confined to R&I, perspectives from NGOs or other representative bodies in the workplace (such as Stonewall and Kaleidoscope for LGBTQ+ rights) can provide useful context for undertaking specific interventions. It is recommended that ongoing work seeks to connect, engage and value these perspectives and labour.

4.2.5 Looking for answers

The question of what type of intervention is most effective in achieving EDI-related goals is not new and has attracted both academic and business interest. A range of comparative studies, meta-analyses and literature reviews have been conducted, particularly looking at intervention types, successes and cross-national contexts, and these will be discussed later in this report.

4.3 Conclusions and recommendations

Organisations are taking a **range of approaches** to exploring EDI challenges in R&I, including production of reports and datasets, monitoring of trends, creating platforms to share best practice and embedding EDI analysis into everyday functions. Collaborative approaches (for example, between funders or universities) are common.

Key challenges of underrepresentation in leadership, specific research disciplines and career progression in R&I are generally well known for some identity characteristics (gender, race and ethnicity) but others appear less frequently explored by R&I organisations (for example, disability, religion, LGBTQ+ and ageing workforces). Supporting knowledge exchange with informal networks and underrepresented communities may be advantageous.

Recent trends focus on understanding change including 'progress' on EDI and on sharing best practice case studies etc.

Recommendations from this section	Policy makers	Funders	Employers	Research
Coordinate reviews to focus on key gaps with regard to EDI challenges and interventions, and to include: ■ disability, religious identity, language in minority or multilingual workplaces, gender and sexual diversity, and age; gender research also to ensure cross-national knowledge-sharing on R&I areas of research or job function where men are a minority. ■ disciplines beyond STEM. ■ R&I in the Global South. ■ intersectional approaches to understanding EDI challenges and implementing change.	✓	✓		✓
Explore repositories for wider EDI literature in R&I (beyond gender), with consideration of accessibility and language options.	✓	✓	✓	
Consider catalyst/grant support for ongoing EDI research or knowledge exchange (including underrepresented communities).	✓	✓		
Interdisciplinary approaches and cross-disciplinary communication.				✓
Basic toolkit or support for researchers looking at comparative EDI across multiple organisation types in R&I.	✓	✓		✓



Chapter 5: What works?

Which interventions have been proven effective, or less effective, and why?

5.1 Overview

This chapter examines EDI interventions implemented at organisations similar to UKRI which have demonstrated a positive impact and which have been less successful, and considers which factors were related to their relative success (or lack of success). To address this research question, we extracted the following information from each of the interventions in our sample:

- where was the intervention developed and implemented?
- what type of intervention is being discussed (for example, is it related to: training or development; mentoring or coaching; strategy or policy change; awareness raising; organisational review or assessment of EDI; learning resources or tools; outreach; or something else)?
- a brief description of the intervention
- what type of methodology was used to evaluate the intervention?
- if reported, the outcomes of the intervention
- do the authors present reasons for success?
- do the authors present reasons for failure?

The evaluation framework allowed us to categorise interventions across different types (and their subtypes, where applicable). However, in the original framework, intervention types were not mutually exclusive, as some EDI interventions:

- had multiple aims within a single intervention (for example, a diversity training programme that also includes an awareness raising campaign)
- were analysed as a bundle, with the impact of multiple EDI interventions (for example, a mentorship programme, diversity training, changes to family-friendly policies) being evaluated in a single analysis.

We have disentangled the results of such analyses where possible, but note that this was a common issue with how evidence supporting EDI interventions was measured and presented (that is 37 out of the 109 sources looked at more than one type of intervention in a single analysis).

There were insufficient sources or clarity of aims to draw conclusions as to whether ‘bundled’ or multiple interventions were more or less effective as a general rule. This could be an area for future research. Findings are therefore organised by intervention type, with examples to highlight interventions that were effective and less effective (where possible) within each type as below:

- training (such as diversity training and comprehensive training programmes)

- strategies, policies or processes (such as diversity management, family-friendly policies and positive action)
- career development programmes (such as networks and affinity groups, and mentorship programmes)
- employer engagement and outreach (such as EDI committees and advisers).

5.1.1 Rating interventions’ effectiveness

The definition of effectiveness was based on the following factors:

- the outcome examined was measured, evaluated and reported in a clear and transparent manner
- there was corroborating evidence across multiple publications supporting the impact of this type of intervention (in other words, using a clustering approach and triangulating evidence)
- whether the authors described reasons for the intervention’s success or failure.

We present a more detailed discussion of rating effectiveness across different evaluation methods in chapter 5 but it is worth noting here that we have highlighted examples that:

- adopted a method rated highly on the Maryland Scientific Method Scale
- used alternative methods that were appropriate to the given research question (for example, using interviews with women in the oil and gas industry to explore how different EDI interventions influenced their experiences) and provided additional insight into the mechanisms underpinning the intervention’s effectiveness.

Interventions that did not report an outcome were not highlighted as examples.

5.1.2 Comparability to UKRI

As far as possible, examples are included from organisations comparable to UKRI in their functions or interests. Where possible, examples that focused on protected characteristics other than gender are foregrounded to counter the overrepresentation of gender-related interventions present in the current sample (75 out of 130 interventions looked at gender in some way, 26 of these focusing solely on this protected characteristic).

5.2 Training

Of the 39 training interventions present in the current sample, around half (20, or 51.3%) focused on diversity training, including two meta-analyses (Bezrukova, Spell, Perry and Jehn, 2016; Kalinoski et al., 2013) and a number of review articles (Alhejji, Garavan, Carbery, O’Brien and McGuire, 2016; Chambers et al., 2017; Phillips, Deiches, Morrison, Chan and Bezyak, 2016), all of which were academic. An additional three articles reviewed

multiple EDI interventions, which included diversity training or bias training as part of these 'bundles' of EDI interventions (for example, Gill, McNally and Berman, 2018; Klingler-Vidra, 2019; Madera, 2013). Of the remaining 11 training interventions, only two directly investigated the impact of a training programme that was not related to general diversity training.

5.2.1 Diversity and bias training

Within the current sample, diversity training referred to organisation-led seminars or training sessions meant to enhance cultural sensitivity and awareness of diversity-related issues (such as structural barriers, racism, sexism or bias)

Many of the initial investigations and reviews of diversity training focused on how this training could improve the diversity of an organisation and address underrepresentation of specific groups (Williams and O'Reilly, 1998). However, other investigations of the impact of diversity training have adopted a multidimensional approach, rather than looking only at the effect of training on increasing representation (for example, Kraiger et al., 1993). Research adopting the multidimensional approach has revealed that diversity training can help improve individuals' attitudes, self-efficacy, motivation, knowledge and behaviour (Kalinowski et al., 2013).

The theory behind this is that diversity training helps individuals identify biases (unconscious or otherwise) within themselves and the barriers within the current structure of organisations and majority culture. In doing so, diversity training is thought to motivate participants to modify their attitudes and behaviour, improving the diversity climate within an institution and slowly progressing towards change at the organisational level.

In their examination of 65 studies, representing 8,465 participants, Kalinoski et al. (2013) identified the average strength of the association between diversity training and participants' (i) affective-based outcomes (that is, attitudes and motivation), (ii) cognitive-based outcomes (that is, verbal knowledge, knowledge organisation and cognitive strategies), and (iii) skill-based outcomes (that is, measures of behaviour and behavioural intentions). Overall, diversity training exhibited a medium to large positive effect on cognitive-based and skill-based outcomes, whereas it exhibited only a small to medium-sized effect on affective-based outcomes.

In addition, Kalinoski et al. (2013) explored which factors influence the strength of the association between diversity training and these multidimensional outcomes. Diversity training had a stronger influence on participants' attitudes and motivation when:

- participants worked together on tasks, rather than work independently
- both active and passive forms of instruction were employed, rather than just passive

- it was delivered by a human instructor, rather than mediated by a computer
- it lasted four or more hours
- the content was delivered over multiple sessions, rather than within a single training session
- the content was delivered by participants' direct manager or supervisor, rather than a general member of staff.

Notably, other factors, such as whether the training was voluntary rather than mandatory, or whether it was an awareness or an awareness-plus-skills training programme, did not strengthen (or weaken) the impact of the training programme on participants' attitudes and motivation. Beyond this additional analysis of factors that might make diversity training more or less effective, Kalinoski et al. (2013) only included studies that employed a degree of experimental control (that is, training vs. control groups with pre-tests and post-tests, training vs. control groups with post-tests only, and single training groups with repeated measures) when investigating the impact of diversity training.

A more recent meta-analysis by Bezrukova et al. (2016), which included effect sizes from 260 independent samples, corroborated the impact of diversity training reported by Kalinoski et al. (2013); it showed that diversity training had a larger impact on participants' cognitive learning than on their behavioural and affective learning. Similar to Kalinoski et al. (2013), only studies that included a pre- or post-test design or control groups were included in the analysis. However, an additional strength of Bezrukova et al. (2016) is the inclusion of unpublished literature in their sample of studies, addressing the issue of publication bias (in other words, the fact that only studies with statistically significant or reporting larger effects are published). In addition to replicating the overall effect reported by Kalinoski et al. (2013), Bezrukova et al. (2016) found that the effects of diversity training on participants' attitudes and motivation (that is, affective learning) decayed over time, while the effects on cognitive learning tended to remain stable or increase. The positive effects of diversity training were greater when the training was delivered over multiple sessions and as part of a comprehensive diversity curriculum (for example, integrated and implemented alongside other diversity initiatives). In contrast to Kalinoski et al. (2013), Bezrukova et al. (2016) found that diversity training programmes that included both awareness and skills development were more likely to yield positive results.

Both of these meta-analyses highlight the large degree of variability in how diversity training is implemented across institutions, as well as variability within the content and how the content is presented. As such, in addition to discussing these two meta-analyses in detail, we chose to highlight interventions that describe the elements or components of diversity training that augment the impact of training on participants' learning, attitudes

and behaviour, including the use of: (i) real data and empirical findings in course materials (for example, Jackson et al., 2014); (ii) video presentation methods (for example, Hennes et al., 2018; see also Moss-Racusin et al., 2016, Pietri et al., 2018 and Peitri et al., 2017 for additional investigations of this video-based training programme); (iii) experiential learning (for example, Zawadzki, Danube and Shields, 2012); and (iv) self-reflection and problem-solving (for example, Carnes et al., 2012).

Diversity training with additional focus on specific protected characteristics

Although addressing a common intervention type, particularly in relation to women's careers, the sources reviewed did not demonstrate clear evidence of the impact of mentorship in the R&I context. This suggests the need for longitudinal evaluations to determine the efficacy of this type of intervention. Sponsorship appears to have positive impact in the corporate sector and could potentially be transferred to other R&I contexts.

Returnships for women are a relatively new intervention within the R&I landscape and have delivered promising results.

Leadership development schemes for groups who share a protected characteristic have also demonstrated some positive results.

As with previous sections, gender is the predominant area of focus within career development programmes in the R&I landscape.

5.2.2 Comprehensive training programmes and specific skills training

Of the remaining 11 training interventions, only two empirically investigated the impact of specific training programmes (Behaghel, Caroli and Roger, 2014; Saetermoe, Chavira, Khachikian, Boyns and Cabello, 2017). The other 10 interventions considered training from either a theoretical perspective (for example, Githens, 2012, outlined the different approaches to diversity a training programme might adopt) or in terms of how training staff in general can be linked to improving diversity. For example, Forbes (2011) described the link between a company's ability to provide employees with opportunities for promotion and career development, such as skills training and mentorship programmes, and its ability to recruit a diverse workforce.

In contrast, the two interventions that discussed specific training programmes looked at how a small-scale, inoculation-type training programme (Behaghel et al., 2014) or a large-scale comprehensive suite of training and other diversity

initiatives (Saetermoe et al., 2017) can improve employability, progression and retention outcomes for staff and students. For example, through a complex series of longitudinal modelling using employment data on more than 2,000 French businesses, Behaghel et al. (2014) found that firms that provided ICT training or sponsored employees to complete this training were more likely to retain employees aged 45 and older on their workforce, suggesting that this type of skills training helps protect older workers in terms of employment and earnings.

While Saetermoe et al. (2017) also looked at change at the organisation level, the authors describe a training programme that is embedded in a comprehensive series of diversity initiatives targeting racial discrimination and built around the tenets of critical race theory (CRT). The initiatives encompassed by the Building Infrastructure Leading to Diversity (BUILD) Promoting Opportunities for Diversity in Education and Research (PODER) programme used CRT as a framework for:

- implementing individual training programmes such as eight-week summer research programmes or four-week intensive research training to help students transition to the research culture, enrolling students in advanced research methods courses, etc.
- building faculty research and mentoring skills by training staff to challenge negative group stereotypes and support positive racial, ethnic and science identities
- developing organisational infrastructure and improving campus culture through centres that support community-based research and train new researchers to conduct socially-relevant research.

In this way, the BUILD PODER programme supports and empowers the whole student, rather than simply attempting to fill-in gaps in their own knowledge (in other words, taking a holistic approach as compared to using a deficit model). Preliminary evidence from the five participating community colleges suggests not only that the BUILD PODER programme has improved racial and ethnic consciousness within campus culture, but also that it encourages more egalitarian and respectful faculty-student relationships and supports biomedical research students as they achieve their goals. However, while Saetermoe et al. (2017) list the evaluation methods employed, including pre- and post-test questionnaires, analysis of existing data and documents, individual and focus group interviews, and ethnographic observations and virtual ethnographies, the authors do not present the analysis of this data or reveal what this data tells us about which elements of the BUILD PODER programme have demonstrated the greatest impact.

Summary

Diversity training programmes were more effective when:

- participants worked together and were involved in self-reflection and problem-solving
- a human instructor delivered training over a series of sessions
- course materials included empirical evidence and were part of a comprehensive diversity curriculum

In contrast, diversity training was less effective when:

- delivered by a general member of staff instead of participants' line manager, and when they employed passive forms of instruction.

While the majority of training examples stemmed from research within the US, the two meta-analyses included studies from multiple countries. The examples also covered a number of characteristics, including gender, race, sexual orientation and age, in addition to EDI issues in general. Finally, although these examples were all pulled from academic sources, the interventions themselves tended to take place in higher education institutions (HEIs) and may be particularly useful for the work that UKRI does with universities and employers.

5.3 Strategies, policies or processes

This overarching theme encompassed a number of subtypes of intervention, including diversity management policies (Madera, 2013; Williams, Kilanski and Muller, 2014), policies related to promoting work-life balance or considered 'family-friendly', and stricter policies such as the implementation of quotas and adopting positive action strategies.

It is worth noting that, while a number of the Call for Evidence interventions described policies supporting diversity and inclusivity management (for example, CIHR's implementation of a gender-based analysis approach, or the Max Planck Society's positive action initiatives and 'reconciliation of family and career' measures), these interventions have not been evaluated (or their evaluation was not described in the submitted Call for Evidence form). As this chapter is devoted to describing which interventions have been more or less effective, we have decided to highlight academic examples that have included an impact evaluation. Nonetheless, we felt it important to mention the above yet-to-be-evaluated examples within the international R&I landscape.

5.3.1 Diversity management policies

Diversity management refers to organisational policies and strategies that aim to improve the diversity of staff, such as, by instituting programmes to attract and retain workers from different backgrounds. This is often informed by the 'business case' for diversity, which indicates that employing people from diverse backgrounds can increase an organisation's profits through improved productivity, creativity, innovation, and problem solving (for example, see Hemphill and Haines, 1997). Within the current sample of academic sources, Wilton et al. (2015) showed that an organisation's communications around their commitment to diversity was related to participants' perceptions of that organisation and subsequent performance in a series of two experiments. Women of colour who were presented with marketing information describing an organisation as 'colour-blind' were more likely to see that organisation as less diverse and more susceptible to bias than women of colour who viewed marketing information with a multicultural message. Moreover, women of colour in the colour blind condition tended to perform worse in a maths test after viewing this material compared with women of colour in the multicultural condition, an effect the authors attribute to 'identity threat'.

Diversity management strategies include a range of programmes such as targeted recruitment, mentoring programmes, affinity groups or networks and diversity training programmes. While this review discusses each of these types of intervention individually where possible, there were a number of interventions within the international database of publications that discussed diversity management practices in general. These interventions tended to describe the implementation of diversity management policies and their impact on individual members of staff (Janssens and Zanon, 2014; Webster, Adams, Maranto, Sawyer and Thoroughgood, 2018; Wilton et al., 2015), the organisation as a whole (Johnson, Warr, Hegarty and Guillemin, 2015) or the effects at both the individual and organisation level (Bieling, Stock and Dorozalla, 2015).

Impact of diversity management policies on staff

In a recent meta-analysis of 27 studies, Webster et al. (2018) found small but significant links between organisations' formal LGBT-supportive policies and practices (such as formal written statements barring discrimination based on LGBT status; including sexual orientation and gender identity in diversity training initiatives; creating LGBT and allies-related employee resource groups) and LGBT employees' work attitudes, psychological strain, disclosure and perceived discrimination. It is worth noting, however, that diversity management policies were not as strongly related to employee outcomes as other factors such as supportive workplace relationships and an LGBT-supportive workplace climate (discussed later in this section). However, this does not mean that these policies are

not important: the results presented by Webster et al. (2018) support the interpretation that the mere presence of diversity management policies may not be enough to bring about change, and instead need to be embedded in an organisation's culture. In other words, similar to the meta-analytic findings on diversity training, diversity management policies demonstrated a positive impact on staff, but the effectiveness of this type of intervention was limited by other factors.

Many of the interventions describing diversity management policies tended to focus on those that are relatively standard practice (such as employee affinity groups or networks, documenting an organisation's commitment to diversity in recruitment materials, having mandatory diversity training for new staff and managers, etc.). Conversely, Janssens and Zanon (2014) diverged from this standard practice to explore two relatively innovative and novel approaches to diversity management:

- valuing multiple knowledge, skills and competencies of a diverse personnel through a social recruitment policy, teamwork in multi-ethnic teams and competency-based job classification and evaluation
- adopting an intersectional lens, making it possible for all employees to bring their entire set of identities to work by encouraging multicultural practice (such as with regards to the intersection between ethnicity and religion and belief in particular), fostering a two-language workplace and implementing flexible work schedule policies.

The results of this case study analysis of a call centre in Belgium suggest that the approaches listed above not only fostered ethnic equality at the organisation more effectively than standard diversity management policies (as indicated by the equality it had achieved between ethnic majority and minority staff) but were also related to positive staff experiences (as indicated in semi-structured interviews with employees).

Impact of diversity management policies on the organisation

In contrast to the above research on the impact of diversity management policies on individuals, a handful of interventions explored how diversity management practices relate to organisational change (for example, Johnson et al., 2015) or organisational recognition (for example, Madera, 2013). While highlighting the importance of considering the broader impact of EDI interventions, these examples also demonstrate the difficulties associated with measuring impact at the level of an organisation. For example, Johnson et al. (2015) present a case study of a faculty implementing a suite of diversity management policies after conducting an equality audit where the success of the policies was measured by changes in the representation of women on senior committees within the faculty. The suite of initiatives included:

- adding gender equity as a key criterion when determining the membership of committees, alongside specific roles and relevant skills
- developing a register of women (and men) who were willing to serve on senior faculty committees
- monitoring gender representation of faculty committees every two years
- conducting faculty committee meetings at family-friendly times
- expanding the Faculty Staff Mentoring Programme.

While the representation of women on senior committees improved after the above policies were implemented, the faculty-level data presented could not be tied directly to the diversity initiatives. Missing information on other factors, such as the uptake of the Faculty Staff Mentoring Programme, and on staff members' experiences of the above policies limited the authors' conclusions about the effectiveness of this suite of diversity management policies (Johnson et al., 2015).

Similarly, although Madera (2013) presented a useful review of the individual diversity management policies employed by successful organisations (defined as being one of the 2010 Diversity Inc. Top 50 Companies for Diversity), the correlational manner in which the information was analysed limits any conclusions regarding the individual or combined impact of these policies on an organisation.

Instead, research that collects data at both the individual employee level and the organisation level is much better suited to investigating the impact of diversity management policies in the broader context. For example, Bieling et al. (2015) surveyed the human resources (HR) managers of 153 German companies and collated this information with information on their company websites about employee productivity, company size and the average age of managers and employees. Structural equation models revealed that companies described as adopting age diversity management strategies in the appraisal and compensation of staff (indicated by HR managers' endorsement of items such as 'In our company, young staff tend to be better assessed in staff assessments than old staff' or 'in our company, performance is the most important factor in determining pay') tended to have better ratings on employee welfare, which was in turn related to employee productivity and organisational performance. The results presented by Bieling et al. (2015) corroborate the links presented by Madera (2013) and Johnson et al. (2015), but the inclusion of additional staff-level data allowed Bieling et al. (2015) to uncover the mechanisms underpinning the effectiveness of diversity management policies in contributing to an organisation's success.

5.3.2 Family-friendly policies

The following set of interventions highlight policies that do not necessarily target underrepresented groups or specific protected characteristics but have an indirect impact on EDI goals such as improving the representation of women in senior positions. For instance, in collating examples from number of American companies, Bailyn (2011) describes the effectiveness of different types of flexible working policy designed to help make it easier for employees to integrate their work with their personal lives, but not at the expense of the effectiveness of their work. These policies included flexible working arrangements that:

- were available to all staff.
- considered employees' personal need for flexibility
- were decided on collectively instead of by the employee's supervisor alone.

The overarching goal of such policies was to help employees with their family responsibilities, which in turn was expected to help women progress to more senior roles within these organisations. However, these are presented as summaries of earlier empirical investigations and the results presented are unclear with regards to the overall effect of such policies. Nonetheless, based on these 'real world' examples, Bailyn (2011) questions the assumptions made by companies that define the ideal worker as one who is continuously at work and available, and concludes that organisations need to work with their employees in a collective manner to produce work-practice change.

Hegewisch and Gornick (2011) present a similarly limited review of the literature on parental leave policies, flexible and/or alternative working arrangements and childcare support to explore how these factors relate to the employment decisions of women, leave-taking and women's earnings. A strength of this review is that the authors explore these policies both across and within OECD countries. However, the variability in the results presented makes interpreting the overall effectiveness of such policies difficult without the data being synthesised in a more formal manner.

Finally, by combining national and survey data across all 28 EU members, Lomazzi, Israel and Crespi (2018) demonstrated that workplace arrangements which support the dual-earner or dual-caregiver family model (such as parental leave schemes, childcare provisions and flexitime) are tied to more egalitarian gender role attitudes. The authors discuss the importance of improving gender role attitudes in order to improve workplace culture and address institutional barriers to female staff progression and representation.

5.3.3 Positive action

Positive action, sometimes referred to as affirmative action, consists of a range of measures aimed at ending discrimination against that group

by redressing the effects of past discrimination. As discussed earlier, some jurisdictions will permit targeted actions (for example, quotas or specific schemes) but such interventions may not be considered welcome, appropriate or legal in other contexts (see Archibong et al., 2009).

Within the current sample, positive action interventions included the introduction of funding schemes with specific targets for the number of female applicants nominated, or only available to women, as well as initiatives aimed at improving the recruitment and retention of female students in higher education. According to Kalev et al. (2006), the implementation of quotas appears to be the most effective diversity strategy for increasing the representation of women in management. However, Bielby et al. (2013) suggest that this is also the least frequently adopted diversity management policy.

Within the current review, three organisations responding to the Call for Evidence described the implementation of such a framework, including the Max Planck Society's Lise Meitner Excellence Programme, Science Foundation Ireland's (SFI's) Starting Investigator Grant (SIRG) gender initiative, and the Swiss National Science Foundation's (SNF's) annual positive action competitive grant scheme, PRIMA. Initial time series analyses of the latter two interventions suggest that introducing a quota (SFI) or a scheme specific to women (SNF) have both been effective in increasing the number of nominations and applications from female researchers.

The effectiveness of such programmes is further supported by the case study analysis presented by Kilango, Qin, Nyoni and Senguo (2017), in which admission data from the university of Dar es Salaam was used to investigate whether the implementation of positive action initiatives improved the representation of women in Tanzanian higher education. The positive actions investigated included lower entry scores for female applicants, remedial pre-university programmes and financial assistance programmes for female students. Admission data revealed a general increase in the representation of female students at the University (for example, from 16% in the Faculty of Science in 1996 before the implementation of the lower cut-off for female applicants' entry scores to 27% in 2003 after implementing lower entry scores for female applicants). Further investigation of student outcomes showed that students enrolled in the remedial programmes performed well (for example, the best-performing student in physics in 2001-02 was a student who attended the remedial programme). There was also good uptake of scholarships by female students (for example, more than half of the students supported by the new scholarships were female). Taken together, the authors cite the above as evidence for the success and effectiveness of the university's affirmative actions.

The case study presented by Kilango et al. (2017) corroborates the initial results from SFI and SNF,

replicates previous research on the use of quotas to improve female representation on governing boards (such as Adriaanse and Schofield, 2014) and provides a useful example of how positive actions may impact gender inequality in a real-world context. However, three additional interventions within the current database query the effectiveness of positive actions and the use of targets or quotas to improve representation in greater detail by looking at how these relate to:

- **bias in recruitment:** using a between-groups experimental design, Beaurain and Masclet (2016) showed that gender quotas reduced bias in the ranking of female applicants for a post
- **women's own experiences within the workplace:** in their interviews with 30 women working within the US oil and gas industry, Williams et al. (2014) found that several participants did not endorse gender equality initiatives that set specific targets for hiring or promoting women as these imply that women in senior posts have obtained their current status due to their gender rather than professional competence and merit
- **resistance from middle management:** although the impact of the Horizon 2020 funded GENOVATE (2016 b) consortium is currently being evaluated, the source identified initial barriers to implementing positive action initiatives within European institutions, such as resistance from middle management and specific beliefs on gender roles and gender equality issues.

These two examples highlight how different methodologies and outcome measures can yield contrasting insights into the same intervention, with Beaurain and Masclet (2016) reporting a positive effect on reducing bias, and Williams et al. (2014) and GENOVATE (2016 b) showing how these programmes are perceived by the staff targeted by the positive action and those involved in implementation.

Summary

This section examined the effectiveness of diversity management strategies in general, as well as specific policies including those that are considered family-friendly and the introduction of positive action or targets.

While the meta-analysis presented by Webster et al. (2018) supports diversity management practices as an effective form of intervention, it also highlighted its limitations: these policies need to be embedded in an organisation's culture to be effective. This is an important factor for HEIs, commercial entities and research funders alike as all three types of UKRI members will have diversity management policies in place to some degree.

In contrast to the evidence supporting general diversity management policies, the effectiveness of family-friendly policies was unclear. While Bailyn (2011) and Hegewisch and Gornick (2011) are both useful reviews of these policies, the evidence reported leaves the overall effectiveness of such policies difficult to ascertain without additional research. Nonetheless, both reviews show that the impact of family-friendly policies will vary depending on context (for example, across OECD countries, or organisational cultures), and which outcome variable is considered.

Positive action was seen to be effective in improving the representation of women in funding award schemes and higher education, as well as reducing bias towards women in recruitment and hiring situations. However, this type of intervention can be less effective as it is seen as discouraging and even insulting by its target population. The lack of uptake with regard to quotas and targeted actions is thus not necessarily because this type of intervention is not effective, but instead because of the resistance to such initiatives from staff. While people's experiences of an EDI intervention are important to consider, the evidence for positive action from the Max Planck Society, SFI and SNF is of particular interest to UKRI as they support the implementation of positive action and targeted interventions into funding structures.

5.4 Career development programmes

Two types of intervention within the current sample related to improving staff experiences and career development: networks and affinity groups, and mentorship programmes.

5.4.1 Networks and affinity groups

While a number of the case studies and reviews within the current database (for example, Madera, 2013; McKinsey, 2016) describe the use of staff networks or affinity groups, only two interventions investigated the impact of these networks on staff experiences (Williams et al., 2014) and on potential staff's perceptions of an organisation (Gutierrez and Saint Clair, 2018). Despite their popularity, networks did not receive the same degree of empirical attention as other frequently adopted diversity initiatives such as diversity training and positive action.

The two empirical investigations of networks uncovered in the current research leave conclusions regarding their effectiveness unclear. With regard to networks and affinity groups, the women working in the US oil and gas industry had mixed views with many noting the value of the social and emotional support they receive from such groups, but also

that these groups may not be particularly useful for progressing their careers as they limit the networking opportunities with more senior members of staff who tend to be men (Williams et al., 2014).

In addition to this qualitative data, Gutierrez and Saint Clair (2018) used an experimental manipulation to explore whether the presence of an employee professional network that is either open to all staff or specific to a minority group (in this case, African American staff) influenced potential applicants' intentions to pursue a post at that organisation. Across the two experiments, the authors found that minority-specific employee networks were associated with reduced feelings of threat among majority group members (that is in contrast to previous research such as Dover et al., 2016) and increased the attractiveness of an organisation as a potential workplace.

What about networks of organisations?

The above research defines networks more narrowly, as applied within a single organisation, such as a women's network or network for people from minority ethnicities, within a department or a faculty. However, as an intervention, networks may also be applied more broadly, including at the sector level, or even at the national level (for example, SNF's National Centres of Competence in Research's (NCCRs) Molecular Ultrafast Science and Technology (MUST) network) and international level (for example, GENOVATE, 2016 a, 2016 b; de Madariaga, 2016).

Although many of these organisation-level networks are currently undergoing evaluation, the size and remit of the organisations involved in them, and the subsequent progress of their actions, is likely to be of considerable use to UKRI's own EDI policies and implementation.

genderSTE, a policy-driven, targeted network that consists of policy makers and experts on gender, science and technology. The network is funded by COST (European Cooperation in Science and Technology) and aims to "advance the state of the art in knowledge and policy implementation on gender, science, technology and environment".

The breadth of genderSTE's initiatives includes:

- addressing structural change within member institutions
- embedding gender in the content of their research
- introducing the gender dimension to avenues of industrial innovation.

Additional examples of European-wide networks and communities that have adopted a similar community approach include:

- GENOVATE
- Gender-NET
- the Horizon 2020 funded ACT project.

5.4.2 Mentorship programmes

Similar to the lack of research on the impact of employee networks, there were very few examples in which the effectiveness of employee mentorship (or sponsorship) programmes was investigated empirically. Where discussions of these did exist, they tended to be reviewed at the sector or national level or in combination with a suite of other initiatives (for example, Bauman, Howell and Villablanca, 2014; Chambers et al., 2017; Evans and Glover, 2012; GENOVATE, 2016 b, 2016 a; Johnson et al., 2015; Madera, 2013; Severina, Edabu and Kimani, 2016), with limited information regarding the methods of evaluation and the data used to support or refute the authors' conclusions.

Only two interventions presented mentorship programmes in isolation, one of which (Lewis, 2017) describes a peer mentorship programme for students with intellectual and developmental disabilities (IDDs) at the University of Rochester in great detail (referred to as Transition Opportunities at the University of Rochester (TOUR)), discussing the history of the programme, how it has progressed and been developed through an iterative process over the course of the last 20 years, and so on. However, the evidence for the programme's effectiveness is limited to quotes from interviews with students who have participated in it (as mentors and as mentees), and programme enrolment over time. While encouraging, it is unclear whether the benefits of this wide-scale programme extend beyond participating students' experiences to improve broader, contextual factors such as building an inclusive culture at the university.

Nonetheless, the qualitative interviews with female staff undertaken by Williams et al. (2014) echo the positive responses from students at the University of Rochester: women working in the US oil and gas industry were consistently positive in their reviews of formal mentoring programmes, especially when that mentorship took place within the first five years at a given organisation. However, their reviews of informal mentorship were less clear. Due to a lack of female representation in senior posts, women tended to be mentored by senior men in the industry. This had some benefits as male mentors had the 'knowhow' when it came to progressing within their given organisation and knew the right people to help achieve this, but the women interviewed also reported not feeling empowered by these relationships (for example, being "treated

like a daughter"). Thus, an important contribution of Williams et al.'s (2014) work is that it highlights how the effectiveness of mentoring programmes varies depending on the gender distribution of staff in a given industry and the gender composition of the mentor-mentee pair.

Summary

Overall, the effectiveness of interventions related to promoting career development amongst women and minority staff was unclear. Networks or affinity groups appeared to provide staff with valuable social and emotional support, but the connection between these and career progression was limited (as the networks tended not to include senior managers who are more likely to be men and from the majority group). Nonetheless, the presence of minority-specific networks helped reduce feelings of threat amongst majority staff and increase the attractiveness of an employer organisation.

Formal mentorship programmes were consistently viewed in a positive manner, especially when that relationship took place within the first five years of the mentee's career. The effectiveness of informal mentorship was unclear and varied depending on the gender composition of the mentor-mentee relationship.

Both types of intervention lacked empirical evidence demonstrating their effectiveness in isolation (from other types of EDI intervention) and using different evaluation methods (for example, both examples of mentorship programmes relied on qualitative data without corroborating evidence from experimental or longitudinal evaluations).

Although the research presented in this subsection was limited, the few sources that investigated the impact of a staff network or mentorship programme covered multiple characteristics including gender, ethnicity and race, and disability. At the individual level, the interventions reviewed here came from academic sources, but evaluations of larger-scale networks such as that involved in GENOVATE (2016) were more likely to be described outside of academic publications and on organisation websites.

5.5 Employer engagement and outreach

Two interventions described putting together a committee or an organisation to oversee and advise on EDI-related issues and initiatives: one implemented at the sector level in Norway (the Committee for Gender Balance and Diversity in Research, or KIF) and the other implemented at the faculty level within a university (Stepan-Norris and Kerrissey, 2013). These roles include monitoring

others' employment of gender diversity initiatives, providing support and championing the aims of gender equality interventions.

Submitted as part of the current study's Call for Evidence, the KIF committee (formed in 2004) aims to address the underrepresentation of women in senior positions within Norway's academic research sector and the underrepresentation of women in certain academic subjects. Described as playing the role of a 'watchdog', KIF has recently expanded its remit to incorporate other protected characteristics, with specific attention to race and ethnicity, as well as promoting and advocating for diversity in general. To accomplish these aims the KIF committee simultaneously provides research organisations with help in implementing their gender diversity initiatives, as well as speaking out on potential EDI challenges within institutions and national authorities. Recent achievements of the KIF committee include the following:

- three universities and one college now require all newly hired managers to demonstrate competence in equality and diversity upon recruitment
- improved knowledge of ethnic minority and immigrant-background people in Norwegian academia, including tackling national and sector monitoring challenges
- raised awareness around ethnic diversity in research through seminars and collecting data on the issue
- integrating the gender dimension (that is, adopting a gendered perspective in research) and improving female representation in Norway's participation in the EU's Horizon 2020 R&I programme.

The committee is funded by Norway's Ministry of Education and Research, adding clout to its regulatory position within the academic research section and among institutions. The KIF committee not only watches over institutions' current actions but also sets diversity goals for the Norwegian academic research sector and actively monitors progress towards these.

In many ways the Equity Advisers at the University of California, Irvine (Stepan-Norris and Kerrissey, 2015) mirror the responsibilities of the KIF committee but on a much smaller scale. As part of the university's ADVANCE programme aimed at promoting institutional transformation, Equity Advisers were tasked not only with monitoring and educating faculty members responsible for hiring decisions but also with holding departments accountable for those decisions. As senior members of staff, Equity Advisers were in a unique position to work closely with their faculty dean (thus having the power to influence decision making) as well as be involved in the Faculty's recruitment, promotion, mentoring programmes and award nominations. By statistically controlling for extraneous factors such as demographic issues, Stepan-Norris and Kerrissey

(2015) consider the improved representation of women and increased female hires in faculties that implemented the ADVANCE Equity Adviser programme, compared with faculties at the university that had not embraced the programme, as evidence for its effectiveness.

Summary

This subsection focused on the different ways in which employers can engage with other organisations and within their own institutions through the recruitment of EDI committees and advisers. Evidence from KIF (obtained through the Call for Evidence) and the University of California, Irvine (published in an academic source) suggests that the implementation of regulatory bodies that both oversee and advise on EDI initiatives can improve both awareness and female representation. However, future research is needed to examine the degree to which the effectiveness of this type of intervention extends to:

- other protected characteristics (outside of gender and ethnicity)
- contexts (such as within UK institutions and commercial entities).
- various degrees of implementation (for example, would an EDI committee or governance board have the same impact as the more active role of Equity Advisers in the ADVANCE programme at the University of California, Irvine?).

An important consideration for UKRI would thus be thinking about how they can fill this gap in research through leading by example (and embedding an evaluation of their own EDI committees or advisory groups into their EDI initiatives) and supporting membership organisations in doing the same.

5.6 Why were these interventions more or less effective?

Above we have reviewed the evidence for EDI interventions that have been more, or less, effective across a number of sources, contexts and (where possible) protected characteristics. While these examples address the yes or no component of the current research question, they lack insight into the 'why' element. This final section discusses reasons for the effectiveness and ineffectiveness of interventions, as noted in the sources, and any reasons suggested for why this might be the case. In this section, we look across the interventions reviewed to examine types of outcome reported and any reasons cited for successes and/or failures.

5.6.1 Outcomes reported

The evaluation framework extracted information regarding interventions' intended areas of impact (recruitment, career development, culture and wellbeing, etc.), intervention type and methodological approach in a quantitative manner. However, the variety of outcomes considered within the current database limited our ability to quantify this information in a meaningful way. In other words, there were too many different types of outcome explored across the 130 interventions analysed in this report to create categories that would be large enough for quantitative analysis.

As such, we extracted qualitative information regarding the outcome explored within each intervention and used a thematic approach to identify the main types of outcome being explored. The most commonly reported types of outcome were those related to individual participants' knowledge and attitudes, or to their behaviour and skills, such as their reactions to a programme. They included:

- an increased awareness and knowledge of EDI issues (for example, Bezrukova et al., 2016; Kalinoski et al., 2013)
- enhanced diversity behaviours (for example, Alhejji et al., 2016; Madera et al., 2013)
- improved skills to handle diversity issues (for example, Hennes et al., 2018).

These types of self-reported outcome suggest that the interventions described here have been effective in the short term and had a positive impact at the level of the individual. Interventions also considered a variety of outcomes that showed a broader effect on EDI in terms of people's experiences, achievements or progression, including those surrounding:

- intervention reach and engagement (such as whether an intervention increased awareness or engagement with EDI).
- process, policy and tool evaluation (for example, creating or piloting new processes or tools).
- impact on EDI for individuals or groups (such as career development, recruitment of target groups, or individual or team performance).

5.6.2 Why did they work?

As part of our evaluation framework, we extracted information regarding why interventions had been successful. This data elucidates the reasons why interventions proved to be effective, from the point of view of the author(s) of the publication.

Reasons for successes of particular interventions are necessarily highly specific to the intervention in question, for example relating to the use of a certain methodology during delivery of a project. However, some commonalities can be found between the interventions.

Most commonly cited reasons for success were:

- **collaboration:** across the organisation (for example, EDI professionals and senior management) and within senior management (for example, board members presenting a unified approach to support female board members' decision making power, Adriaanse and Schofield, 2014)
- **leadership:** senior management committed to intervention and diversity management strategies integral to their effective integration (for example, Heijes, 2011)
- **strategic alignment and drivers:** alignment with organisational or sector strategy and/or sector strategy or policy acting as a driver for the intervention
- **community:** interventions created positive relationships and networks for the individuals concerned
- **evidence:** use of evidence to design an intervention programme (for example, use of active forms of instruction), to raise awareness of EDI issues and to track progress and success of the intervention after implementation
- **project management and accountability:** having well-defined goals, expectations and roles
- **organisational culture:** embedding awareness and actions addressing EDI issues in organisational culture, rather than them being a bolt-on project or one-off.

5.6.3 Why did they not work?

During evaluation of eligible sources, we also extracted any information reported by author(s) on reasons why any aspect(s) of the interventions had not been successful.

Overall, none of the interventions reported that they had failed. However, many included reflections on challenges encountered or areas where the intervention had been less successful. This data illustrates why certain interventions proved not to be as effective as desired, from the point of view of the author(s).

Most commonly cited reasons for failure were:

- lack of clarity: in policy guidelines and how to implement them (such as positive action and use of quotas) and in definitions of protected characteristics, EDI issues and outcome measures
- reliance on the business case for diversity: failure to capture the complexity of EDI outcomes in an organisational context
- capacity and organisational constraints: individuals unable to enact new skills or attitudes within the current context due to workload, organisational culture or senior management
- reluctance: in discussing EDI issues (such as

female staff unwilling to raise issues for fear of being labelled as difficult) and EDI issues related to historical and structural factors, and thus unchangeable

- Inconsistencies: in the application of new processes or policies
- Small sample sizes: noted as a limitation by several authors
- Data: limitations surrounding missing data and inability to measure impact of interventions in isolation were cited as detrimental.

5.7 Conclusions and recommendations

This chapter has highlighted EDI interventions that have demonstrated impact within the international R&I context, with particularly strong corroborating evidence for the use of diversity training programmes and diversity management policies. There was also evidence that positive or affirmative action measures were valuable in improving the representation of women in funding award schemes and higher education, as well as reducing bias towards women in recruitment and hiring situations.

In contrast, the effectiveness of family-friendly policies, career development programmes and employer engagement and outreach through EDI committees and advisers was mixed. There was evidence of a positive impact in some areas (such as improving the experiences of and support for women and minority staff through networks and affinity groups) but not others (such as the lack of senior managers in such affinity groups limiting the utility of a network in improving career progression).

Overall, successful interventions tended to embed EDI awareness and initiatives in an organisational context, which was facilitated by commitment from senior management and involved collaboration – whether this be across teams within a single organisation or across multiple organisations. Common features of less effective interventions were:

- a lack of clarity around the policy and how to implement it
- a reliance on the business case for diversity
- methodological issues related to small sample sizes and data collection.

While this chapter highlights examples where interventions have been formally evaluated, and why they have or have not been effective, there were a number of sources relevant to the design and implementation of EDI interventions that were not described in detail.

These include studies investigating which factors underpin processes related to EDI, such as reducing bias in:

- the staff interview process (Parent, Weiser and McCourt, 2015)
- staff recruitment and post allocation (for

instance, Bohnet, Bazerman and van Geen, 2012; Nadler, Lowery, Grebinoski and Jones, 2014; Williams, 2018)

- selection processes for potential female and minority students (for example, Milkman, Akinola and Chugh, 2015).

Beyond reducing bias, our sample also included research on the factors that influence individuals' perception of:

- a specific group (such as individuals with mental health issues and depression, Hamann, Mendel, Reichhart, Rummel-Kluge and Kissling, 2016)
- an organisation and how attractive it is to work for (for example, Windscheid et al., 2016).

Finally, in their review, Nishii, Khattab, Shemla and Paluch (2018) note that HR strategy researchers have documented the power of mutually-reinforcing 'bundles' of actions which corresponds with the current sample of literature, where we saw a

number of publications reviewing the impact of a series of diversity management policies or EDI strategies within a single source.

While this bundling can improve overall effectiveness, as discussed by Webster et al. (2018) with regard to the effectiveness of LGBT-supportive policies, it can pose issues for evaluation. More specifically, bundles of EDI and diversity management policies tend to be implemented alongside one another, making it difficult to isolate and evaluate the impact of a single intervention. An additional problem is that organisations tend to use the same set of indicators to measure their success. We discuss these difficulties in the next chapter in our presentation of studies examining the impact of EDI interventions through the use of creative experimental manipulations and the application of systematic controls in their methodological approaches.

Table 5.1. Summary of recommendations for

what's working

Recommendations from this section	Policy makers	Funders	Employers	Research
Foster collaboration within and between organisations to improve issues surrounding limited resources, methodological issues, duplication of work and sharing of good practice.	✓	✓	✓	
Additional research needed to clarify the effectiveness of particular interventions, including family-friendly policies, mentorship programmes and employer engagement via EDI committees and advisers.			✓	✓
Embed EDI awareness and initiatives into organisational culture, ensuring alignment between an organisation's communicated messages and intentions, and actual practice.	✓	✓	✓	
Involve all levels of staff in the implementation of EDI interventions, with strong and demonstrable commitment from senior management and support for staff directly involved in overseeing and evaluating interventions.		✓	✓	
Make certain that any intervention is clearly defined (for example, clear agendas and outcomes for training programmes, detailed records of new policies and how to go about implementing them, etc.).	✓	✓	✓	✓

Chapter 6: Measuring Success



Which interventions have been proven effective, or less effective, and why?

6.1 Overview

This chapter considers the evaluation of interventions and how we can have confidence in the impact (or lack of impact) they report. In order to explore these themes, this chapter presents findings from the following questions in the evaluation framework:

- were outcomes measured and evaluated and are results reported?
- what type of evaluation data was captured (for example, quantitative, qualitative or both)?
- what types of methodology were used to evaluate the different types of intervention?
- where possible, can the Maryland Scientific Method Scale be used to describe the rigour of the evaluation or are there other ways to assess rigour that are more appropriate for EDI research in the R&I landscape?

Analysis of question responses, alongside other data from the framework, is used to present an account of methodological approaches to evaluate EDI interventions, the robustness of work taking place in the R&I sector, hierarchies of evidence and critical reflections, from the sources, on evaluation challenges.

6.2 Methodological hierarchies

To address the current research question regarding how the effectiveness of EDI interventions is measured, we present both an overview of the different evaluation methods used in the current sample and specific case studies that showcase how these methods can be applied effectively in the R&I landscape. However, this evokes a broader discussion of how we assess our confidence in a given source's methodological approach.

Traditional frameworks, such as the Maryland Scientific Method Scale, adopt a hierarchal approach that places simpler methods employing fewer statistical or experimental controls (such as a cross-sectional comparison of a control and an experimental group) on lower levels than those that exert a higher degree of control using randomisation techniques (in other words, a RCT).

A modified version of the Maryland Scientific Method Scale, adapted to the EDI context, was used in the current review to rate the scientific rigour of each evaluation and contribute to our overall rating of confidence in the interventions' effectiveness (see table 6.1 for a summary).

Table 6.1. Application of the Maryland Scientific Method Scale to current sample.

Level of the Maryland Scientific Method Scale	No. of interventions	%
1: Correlation (for example, departments with a female leader have more female staff).	29	22.3
2: Before and after assessment, with no control of conditions (for example, female staff in a department increased after the appointment of a female leader).	21	16.2
3: Before and after assessment, with experimental conditions (for example, female staff in a department increased after the appointment of a female leader, female staff in a department did not increase after the appointment of a male leader).	7	5.4
4: Before and after assessment, with multiple experimental conditions (for example, as with level 3 but with additional controls for gender culture in a department and individual backgrounds of staff).	11	8.5
5: RCT	4	3.1
Not applicable, unclassifiable or unclear	58	44.6
Total	130	100.0

Table 6.1 shows that almost half of the evaluations in the current review could not be classified on this scale, whether because there was insufficient information provided in the publication to clearly categorise its evaluation method, or because the intervention was assessed using non-experimental approaches (described in section 5.4.3). Of those interventions that were classifiable, only 22 (17.0%) exerted some form of experimental control (i.e. were labelled as level 3 and above), meaning that half of the Maryland Scientific Method Scale was not really applicable or useful in evaluating the rigour of the current sample.

This is a major limitation to the use of this hierarchy in the EDI context given the challenges with small sample sizes noted in earlier chapters.

Further to the difficulty of categorisation is the question of the equality impact of any such implied hierarchy of categorisation. A hierarchical valuing of different evidence sources has been conceptually critiqued for failing to take into account the context of how knowledge is produced and disseminated (for example, Nairn, 2012). For instance, certain academic disciplines may be more likely (or better resourced) to conduct large-scale experimental designs, whereas others focus on action learning. Moreover, certain types of data source may be privileged, for example quantitative data on presentation rather than qualitative interviews of experience.

A range of criticisms have been directed particularly at the view that RCTs are 'gold standard' for 'evidence-based policy' approaches: citing, for example, the benefits of 'storytelling' approaches (Cairney and Oliver 2017), 'realist' complex systems approaches (Pawson et al., 2005) or simply the challenge of different research methods reproducing wider societal inequalities (for example, Harding and Norberg 2005).

6.2.1 Alternative forms of evaluation and synthesis

While useful, experimental and meta-analytic approaches tend to reveal little about why an intervention is effective. As highlighted by Hansen (2014), the traditional hierarchy of evidence that prioritises meta-analyses and RCTs can and should be supplemented with other forms of analysis, mainly qualitative methods, to provide insight into the contextual factors and mechanisms underpinning variability in the results reported.

Although there were a number of limitations in measuring the effectiveness of EDI interventions and their application within the R&I landscape (discussed in section 6.4), it does not mean that our team could not assess an intervention's effectiveness. Instead, we needed to rely on other forms of narrative synthesis (see Popay et al., 2006 for an overview of ESRC guidance on narrative synthesis), such as triangulating evidence to assess the effectiveness of an intervention (as presented in chapter 5) or clustering evaluations around a common theme such as their methodology (as presented below). In

contrast to meta-analyses, which require evaluations' methodological approaches to be similar, these alternative forms of synthesis are strengthened by including studies that have adopted different approaches. For example, looking at the impact of quotas from different angles (such as qualitative interviews by Williams et al., 2014, and quantitative analysis of how these influence bias in recruitment by Beaurain and Masclet, 2016) allowed us to obtain a more complete and nuanced picture of the overall effectiveness of this type of intervention in chapter 5.

When it comes to EDI interventions, there are a number of advantages in considering their impact from multiple viewpoints to understand how this varies by:

- organisational context (such as large versus small institutions; research funders, commercial entities or higher education providers)
- individual experiences versus the organisational or structural change
- EDI area (for example, career-related outcomes such as recruitment and promotion, or leave policies; organisational culture and staff wellbeing; outreach and public engagement).

Given these advantages to corroborating multiple sources of evidence, we chose to highlight in this section:

- studies that employed conventional experimental methods, as these fulfil traditional interpretations of robustness and rigour
- sources that employed alternative methods in evaluating the impact of EDI interventions that were particularly useful in determining their overall effectiveness
- research that has been replicated or corroborates existing evidence.

6.3 Outcome measurement and reporting

One of the major challenges in investigating the impact or effectiveness of an EDI intervention is determining how to measure its intended outcomes: should impact be reflected in quantitative statistics such as Likert scale ratings on a survey, or promotion rates over time? Or should the outcomes of an EDI intervention take a closer look at the human experience and capture qualitative data from interviews, focus groups or journal entries?

There are a number of factors to consider when answering these questions, as which outcomes are measured, and how, will depend on:

- the nature of the intervention (for example, is it a change in policy or the introduction of a new training or mentorship programme?)
- the timeframe of the research (for example, would annual data capture the effect or would the policy changes take more than a year or two to come to fruition?).

Critics have also noted that the choice of 'measurement' of EDI intervention success may even at times perpetuate inequalities. For example, Berrey (2014) noted that in focusing on representation in leadership roles and those who are 'leadership material', many approaches to gender and 'race' diversity may actually perpetuate the hierarchy of social class.

Using the information yielded by the evaluation framework, this review considers how outcomes were measured across different contexts (for example, academic, grey or Call for Evidence sources), the various types of intervention and whether there are gaps in how data is measured and presented as evidence.

The interventions evaluated within the current sample of international sources tended to rely on quantitative data (40.0%) or a combination of qualitative and quantitative data (26.2%). An additional 18.5% considered qualitative data on its own and for 15.3% of interventions the type of data captured was unclear. There were striking differences in what type of data was presented across the different publication sources; for instance, 52.9% of the interventions reported through the Call for Evidence did not include sufficient information on their data to be classified as quantitative, qualitative or both, compared to only 9.2% of academic and 13.3% of grey publications. Interventions examined within academic sources tended to rely on quantitative methods (43.9% or 43 interventions out of 98 collected from this strand), whereas interventions evaluated within the grey literature tended to examine both quantitative and qualitative data (eight interventions out of 15 collected from this strand, or 53.3%). When data was presented within the Call for Evidence sources, it tended to be quantitative in nature (five interventions out of 17 collected from this strand, or 29.4%).

With regard to intervention type, 43.8% of strategy or policy change interventions did not have enough information presented to determine whether the data captured was quantitative, qualitative or both. This is almost half of the evaluations considering this type of EDI intervention in the current sample (seven out of 16 sources looking at strategy or policy change). In contrast, training interventions were likely to be evaluated using quantitative (47.1%) or combined quantitative and qualitative data (41.2%), and there was a similar reliance on quantitative data in sources evaluating more than one intervention type in one analysis.

There was also considerable variability in how the sources presented information regarding the measurement and evaluation of outcome variables. For example, one out of five academic sources were missing information on how their outcome variables were measured and evaluated (20 interventions out of 98 collected, or 20.4%), compared with three out of five within the grey sources (nine interventions out of 15 collected, or 60.0%) and three out of four within the Call for Evidence sources (13 interventions out of 17 collected, or 76.5%). Similar to the above results regarding the type of data collected, a considerable proportion of sources investigating the impact of an EDI intervention related to strategy or policy change did not present information on how their outcome variables were measured and evaluated (nine out of 16 sources looking at this type of intervention, or 56.3%). Conversely, the majority of sources investigating the other types of EDI intervention presented sufficient information on both the measurement and analysis of their outcomes; for example, 88.2% of training and 66.7% of mentorship evaluations satisfied these criteria in their reporting of evidence.

From the patterns present within the current sample of evaluations, it appears that a combination of both quantitative and qualitative data is preferred when publishing evidence in an academic or grey source. However, this mixed methodology tends to be selectively applied to interventions that lend themselves to experimental methods (such as training interventions) and may be more difficult to adopt (or present clearly) when discussing EDI interventions related to changes in organisational strategy or policy.

6.4 Evaluation methods

This subsection explores the types of evaluation method present in the current sample by looking at how they have been applied across different intervention types and EDI areas, as well as focusing in on the methods themselves, highlighting sources that have applied specific approaches in an appropriate and informative manner.

To develop a clearer idea about types of evaluation method and the contexts within which they were used, the framework presented a list of nine possible methods (and two additional categories for those that did not adopt or describe a method). Originally, these methods were presented in non-mutually exclusive categories so that we could see the overall frequencies of each methodological approach (see table 6.2) as well as which studies adopted multiple methods.

Table 6.2. Summary of evaluation methods within current sample

Type of methodology used to evaluate the intervention	No. of interventions
Within-groups design	6
Between-groups design	22
Time series analysis	6
Cross-sectional analysis	24
Case study or studies	17
Qualitative analysis of interviews	33
Qualitative analysis of focus groups	4
Ethnography or observation	4
Conceptual or not based on empirical evidence	5
Unsure	12
None	0

However, in order to gain insight into which methods were being used in different contexts, we created mutually exclusive categories for each method (presented in section 2.5.6) and compared them across (i) individual types of interventions, and (ii) the EDI areas listed in the evaluation framework (see table 6.3).

With regard to intervention type, a mixed-method approach (that is, combining more than one type of evaluation method) tended to be adopted in sources looking at more than one type of intervention in a single evaluation. This evaluation method was also relatively popular in sources investigating the effectiveness of a training or development programme, a mentorship or coaching programme, or a change in strategy or policy. The other two most frequently adopted evaluation methods were between-groups and cross-sectional analysis and qualitative analysis of interviews, both of which tended to be used across a variety of intervention types. A between-groups and cross-sectional approach was particularly suited to examining training programmes and learning or teaching resources.

A large proportion of the interventions evaluated in the current sample covered more than one EDI area and, not surprisingly, those that did investigate multiple EDI areas within a single source tended to adopt a variety of research methods, with a particular reliance on between-groups and cross-sectional approaches and interviews. Similarly, EDI interventions related to individuals' careers (such as recruitment, promotion or leave policies) were also evaluated

in a variety of ways, including within- and between-groups designs, case studies, interviews and combinations of these. In contrast, the effectiveness of interventions related to funding schemes (grant awards, scholarships etc.) was more likely to be evaluated through time series or document analysis, suggesting either that the data available regarding such interventions may be limiting the types of evaluation methods employed or that this EDI area in general requires additional evaluation that looks at effectiveness from another angle.

Taken together, the information presented in table 6.3 indicates not only which methods are most frequently adopted in EDI research but also where there is missing information about the different types of intervention and EDI areas.

6.4.1 Experimental manipulations and randomised designs

Experimental evaluation methods include those in which the researcher deliberately changes something (known as the independent variable) to observe the effect of that change on something else (known as the dependent or outcome variable). This might include methods that compare experiences within one group at two or more time points before and after the experimental manipulation (a within-group, time series or longitudinal study design) or between a treatment group (who were exposed to the intervention) and a control group (who were not exposed to the intervention) (a between-groups design). When the distribution of participants across both groups is randomised, this is known as a RCT.

Table 6.3. Evaluation methods by intervention type and EDI area

Intervention type	Within- groups / longitudinal	Between- groups/ cross- sectional	Time series analysis	Case study (or case studies)	Qualitative analysis of interviews	Document analysis	Other quantitative analysis	Mixed methods	Unknown
Training or development	0	6	0	0	2	1	2	3	1
Mentoring and coaching	0	0	0	0	1	0	0	2	0
Strategy or policy change	0	1	1	1	0	1	2	3	3
Awareness raising	1	1	0	0	0	0	0	0	0
Organisational review or assessment of EDI	0	0	0	0	1	0	0	1	0
Learning or teaching resources	1	4	0	0	0	0	0	0	0
Positive action	0	2	0	0	0	0	3	0	2
Networks, sponsors and champions	0	1	0	1	1	0	0	0	0
Family-friendly policies	0	1	0	0	0	0	0	0	0
Data collection and tools	0	0	0	0	1	0	2	0	1
Other	0	6	0	0	1	2	0	3	1
Multiple intervention types	2	8	3	7	5	1	3	11	6
Total	4	30	4	9	12	5	12	23	14
EDI area									
Careers (recruitment, promotion, leave policies etc.)	1	7	1	3	2	0	4	1	4
Culture and wellbeing (inclusion, experiences etc.)	0	2	0	0	1	1	1	2	3

Table 6.3 Evaluation methods by intervention type and EDI area.

Intervention type	Within- groups / longitudinal	Between- groups/ cross- sectional	Time series analysis	Case study (or case studies)	Qualitative analysis of interviews	Document analysis	Other quantitative analysis	Other qualitative analysis	Mixed methods	Unknown
Outreach and public engagement (community work, events etc.)	1	0	0	0	1	0	0	0	0	2
Data (equality monitoring, increasing disclosure etc.)	0	0	0	0	1	0	2	0	1	0
Funding (scholarships, grant awards etc.)	0	0	1	0	0	0	3	0	0	1
General policy, practice or governance	0	1	0	0	1	1	0	0	2	0
Other	0	9	1	2	1	0	0	0	2	0
Multiple EDI areas	2	11	1	4	5	3	2	1	9	4
Total	4	30	4	9	12	5	12	2	23	14

These approaches, particularly between-groups comparisons, were quite prominent in the current sample, likely because of the large proportion of studies coming from an academic source. As highlighted in the previous section, this type of design was particularly useful when looking at training or development programmes or learning and teaching resources as these types of intervention can be applied to one group (the treatment group) while being withheld from another (the control group).

However, whether participants can be randomly assigned to these two conditions depends on the context of the evaluation and the intervention. For example, in looking at the effectiveness of a diversity training programme, Jackson et al. (2014) were able to randomly assign staff to a control or experimental group because participants had enlisted in the research itself, allowing the researcher control over the experimental manipulation.

Jackson, Hillard and Schneider (2014). Using implicit bias training to improve attitudes toward women in STEM.	
Data collection strand	Academic
Aim	To evaluate the impact of diversity training on university staff's implicit associations and explicit attitudes toward women in STEM.
Method	The sample consisted of 234 STEM faculty staff from four diverse universities in the American Midwest. Random matched assignment (based on subject area and university) was used to create an experimental and a control group. A pre- and post-test design was used to evaluate change in participants' implicit associations (using the Go/No-Go Association task, a variant of the Implicit Association Task) and explicit endorsement of stereotypes (measured through a questionnaire). The content of the 30-minute diversity training presented to the experimental group included: national and local data on the representation of women in STEM; research on the effects of implicit bias on hiring, promotion and retention; and strategies for overcoming bias. Participants in the control condition attended a regularly scheduled department faculty meeting. A series of two (pre- and post-test) by two (control versus experimental group) by two (men versus women) analyses of variance (ANOVAs) were used to explore the influence of diversity training on implicit associations and explicit attitudes.
Results	Although the training improved personal implicit associations about women in STEM for men, the same improvements were not seen amongst women, who tended to have more positive implicit associations prior to the training. There were no comparable improvements in explicit endorsement of stereotypes about women in STEM as a result of the training, suggesting that the impact of diversity training is limited to men and implicit associations about women in STEM.
Why is this important?	Highlights how the employment of an experimental design and the inclusion of gender as a moderator can be used to examine the impact and limitations of an intervention's effectiveness.

In contrast, research focusing on the impact of diversity training that all staff or students need to complete may require creation of a different sort of control group (as seen in Madera et al., 2013), where they are still exposed to the intervention while participants in the experimental group receive an adapted version of it. This restriction regarding randomisation is particularly relevant

when withholding an intervention could potentially harm participants (for example, restricting access to a mentorship programme for some but not all participants could be detrimental to individuals' career progression), which may be the case more often than not when applying EDI interventions within a real-world context instead of in a controlled experiment.

Madera et al. (2013). Enhancing the effects of sexual orientation diversity training: the effects of setting goals and training mentors on attitudes and behaviours.

Data collection strand	Academic
Aim	To determine whether the short- and long-term impact of a diversity training programme with a specific module on LGBT awareness can be enhanced by including a goal-setting component.
Method	The study focused on a sample of 500 undergraduate students who were provided with a 1.5 hour diversity training session as part of their induction. Although all students participated in the diversity training, approximately half were randomly assigned to the adapted version of the training that included goal-setting and mentorship. The LGBT awareness portion of the diversity training involved senior students sharing stories describing the challenges they have faced while being at the university. In the experimental condition, this was followed by an activity where students developed a personalised contract and set goals to achieve these desired changes. Outcome measures included students' attitudes toward sexual orientation diversity (Attitudes Toward Lesbians and Gay Men Scale) and behaviours sexual orientation diversity supportive (assessed via ratings on items such as 'I have been to a gay or lesbian bar, social club, party or march' or 'I have laughed at a 'queer' joke'), both of which were measured at a three-month and an eight-month follow-up.
Results	The inclusion of goal-setting activities in the diversity training had a positive impact on students' supportive behaviours and attitudes towards LGBT individuals, compared to students who did not participate in this component of the training. Additional mediation analysis revealed that it was participants' supportive behaviours that underpinned the impact of goal-setting on improving their sexual orientation attitudes. In other words, the actions that students had to set for themselves helped generate longer-term changes in their attitudes towards others.
Why is this important?	The authors considered the long-term impact of the training session by including a three-month and an eight-month post-intervention follow-up with participants. While the design does not include pre-test assessments, it does employ an experimental manipulation and adopt a longitudinal design.

Hennes et al. (2018). Increasing the perceived malleability of gender bias using a modified Video Intervention for Diversity in STEM (VIDS).

Data collection strand	Academic
Aim	To examine whether the impact of a diversity training programme VIDS; Moss-Racusin, Sanzari, Caluori and Rabasco, 2018) is improved by the addition of a module (UNITE) that teaches participants that bias is malleable (that is, a growth-mindset approach) and provides tools to address bias.
Method	This research consists of two studies, the first with a convenience sample of adults from the general public (number: 343) and the second with STEM faculty working at an American university (number: 149). Experiment 1 adopted a between-participants design, with participants randomly assigned to one of four groups: an experimental group (who viewed the original VIDS training video); a control group (who viewed clips from existing science documentaries); a second experimental group (who viewed VIDS plus the UNITE module); or a second control group (who viewed the science documentary clips plus a growth-mindset module on improving employee engagement).

Method (continued)	The VIDS training programme included two narrative videos (stories illustrating the results of published papers on gender bias and stereotypes) and two expert interview videos (which described the same research but in a fact-based manner). The UNITE module used evidence and citations to present bias as malleable, as well as examples of individuals who have improved their biases and promoted equity in their fields. The module closed with concrete, evidence-based ‘tips’ for addressing bias. Pre- and post-test questionnaires assessed participants’: ■ awareness of gender bias in the workplace ■ own gender bias (Modern Sexism Scale) ■ ability to recognise bias ■ growth-mindset (Lay Theories of Racial Bias Scale) ■ self-efficacy (Individual Self-Efficacy Scale). To examine group differences in the above outcomes, a two (VIDS experimental versus first control group) by two (VIDS plus UNITE versus second control group) design. Experiment 2 replicated this experiment in a sample of STEM faculty.
Results	Both experiments replicated the impact of the VIDS intervention on participants’ awareness of bias and modern sexism and their ability to recognise bias, reported by Moss-Racusin et al. (2018). However, while VIDS had a positive influence on the above outcomes, participants in this group tended to adopt a fixed mindset about gender bias after viewing the VIDS videos. In contrast, the addition of the UNITE module buffered against this unintended effect, with participants in this condition displaying a growth-mindset towards, and increased self-efficacy in addressing, bias. Taken together these results suggest that the VIDS plus UNITE intervention may be most successful in improving attitudes without impairing efficacy in taking action.
Why is this important?	Demonstrates the importance of replicating results of previous research with different populations, as well as showing that this is a necessary first step when investigating whether the effectiveness of an intervention can be improved by the addition of a new component.

Strengths and limitations

In sum, there are three main challenges in adopting an experimental design to investigate the effectiveness of an EDI intervention:

- difficulty of randomly assigning participants to different conditions
- difficulty of applying this method to EDI interventions that are less open to manipulation, such as those related to organisational reviews or programmes that are open to all staff and/or students
- the fact that the design may not reflect how the intervention will present itself in the real-world context, where other psychological, social and environmental factors may improve or hinder its effectiveness.

However, with regard to the last challenge, a number of experimental studies have measured other factors (such as motivation to attend a diversity training programme, or sociodemographic factors such as current age and contract level when looking at promotion policies) that can be statistically taken into account when calculating the impact of the intervention on the outcome variable. Moreover, when applied in an appropriate manner, an experimental design provides clearer insight into the relation between an intervention and its outcomes, compared with other correlation-based approaches that do not exert the same degree of control over conditions.

Beaurain and Masclet (2016). Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence.

Data collection strand	Academic
Aim	To investigate the experimental impact of quota policies on gender discrimination in recruitment decisions.
Method	The experiment asked participants to consider a hypothetical situation in which they were asked to rank six applicants for a given post based on both their years and their field of education, their age and their gender, and select two to whom they would offer the job. Participants were randomly assigned to one of three experimental conditions: (i) the no quota group, where participants could select any two applicants of their choice; (ii) the low penalty group, where participants had to select two applicants to hire but one had to be a woman or their organisation would have to pay a small penalty; or (iii) the high penalty group, which had the same quota as the low penalty group except the penalty that the organisation would have to pay was significantly higher.
Results	Female applicants were ranked less favourably by participants in the no quota condition compared with participants in the low and high penalty conditions. Women were also significantly less likely to be selected for a post in the no quota condition compared to the high penalty condition, but not the low penalty condition, suggesting that there is a threshold up to which violation of an affirmative action initiative such as a quota for recruitment is acceptable.
Why is this important?	One of the few sources in the current database to employ an experimental design to investigate the impact of an intervention that was not a training programme.

6.4.2 Meta-analyses

The extensive literature review applied within the current research uncovered three valuable examples of meta-analyses (Bezrukova et al., 2016; Kalinoski et al., 2013; Webster et al., 2018), demonstrating that this rigorous approach can be applied within the EDI context. These meta-analytic reviews highlight the complex nature of this approach and provide EDI researchers, institutions and organisations within the R&I landscape with an example of what methods need to be put in place to make future meta-analysis possible. The results of these meta-analyses, presented in chapter 4, also highlighted another strength of meta-analytic reviews: when the number and variability of effect sizes included in a meta-analysis are sufficient, authors can explore what factors moderate the strength of the association between the independent variable (for example, diversity training in Bezrukova et al., 2016, and Kalinoski et al., 2013) and the outcome of interest (for example, attitude, motivation, behaviour and learning in Bezrukova et al., 2016, and Kalinoski et al., 2013). Essentially, moderation analysis within the context of EDI interventions would reveal which factors, including context, content, implementation differences and so on, might make that specific intervention more effective or less effective.

While not eligible within the current evaluation framework, the extensive literature review within the

current study also uncovered four additional meta-analyses containing useful information with regard to future EDI research and intervention development and evaluation:

- Bell, Villado, Lukasik, Belau and Briggs (2011): meta-analysis of 92 studies investigating the links between different types of team diversity (including diversity in team members' educational background and skills, as well as race and gender) and their overall performance (indicated by their efficiency, general performance and creativity or innovation); diversity in terms of individual team members' race and gender was negatively associated with team performance, suggesting that it is not just the presence of diversity that is required to increase performance; instead, diverse teams also need to embrace their differences and build a cohesive unit that works well together
- Byron and Post (2016): meta-analysis of 87 independent samples to examine whether and how female directors influence their organisations' socially responsible business practices and social reputation; results supported a positive link between female board representation and organisations' social performance, with this being particularly true in countries with higher gender parity.
- Eby et al. (2013): data from 173 samples (a

combined sample of 40,737 participants) showed that mentees (or protégés) who reported greater instrumental support and better relationship quality with their mentor were more likely to be satisfied with these relationships.

Strengths and limitations

The major advantage of meta-analyses is that they yield a clear estimate of an intervention's overall effect. In other words, they tell you whether the intervention had a significant effect on an outcome and, if so, how strong this effect was. However, although meta-analysis tends to be the gold standard for synthesising information across a number of sources, there are a number of limitations to this approach that make it particularly difficult to apply to EDI research.

Specifically, meta-analyses require a certain degree of consistency across studies if they are to be aggregable. This means that, within the EDI context, researchers would need to have:

- the same definitions and labels for protected characteristics
- comparable outcome measures (for example, Kalinoski et al., 2013, decided whether outcome measures were related to affective-, cognitive- or skills-based learning depending on the content of each study's measures to determine their relation to diversity training)
- adopted a quantitative approach that yields an effect size (or the statistics required to calculate an effect size).

Within the current sample of eligible studies, roughly one in three publications reported

an effect size and the majority of these were comparing different types of intervention or different types of outcome. According to traditional meta-analytic practice, meta-analytic techniques should not be applied to samples including fewer than seven independent effect sizes (Rosenthal, 1984). As such, not only were we unable to undertake any form of meta-analysis on the interventions and effect sizes present within the current database, but this limitation also mirrored one of the main challenges in investigating EDI within the R&I landscape – too few evaluations are adopting methodological approaches that go beyond looking at high-level data (staff recruitment and promotion rates, national labour force statistics, etc.) and taking a more rigorous approach to defining and measuring the outcomes of EDI interventions (see additional discussion by Evans and Glover, 2012).

6.4.3 Non-experimental methods and valuable alternatives

Non-experimental methods include a variety of methodological approaches, some quantitative in nature (such as surveys or questionnaires administered at a single time point or over time as in a longitudinal design, or a time series analysis) while others collect qualitative information (such as interviews, focus groups, document analysis and discourse analysis). The important distinction between these and the experimental methods described above is that non-experimental methods are missing the manipulation imposed by the researcher. As such, these methods are more susceptible to 'noise' in the data (such as differences in participants' background, experiences and motivations) that limits the interpretation of cause and effect in evaluating an intervention.

Heijes, C. (2011). Diversity management in the public sector: moving from hobbyism toward integration. An exploratory case study in the Netherlands.

Data collection strand	Academic
Aim	To describe in detail the development of diversity policy (with a particular focus on race) in a large public organisation in order to provide insight into the barriers faced along the way and which tools could be the most useful in this developmental process.
Method	The author adopted a case study approach to provide detailed, in-depth data that goes beyond what would be captured through questionnaires or surveys by also considering the organisation's context. The case study itself follows the development of diversity policy in a large organisation (over 30,000 employees) in the Netherlands over a period of two decades. Data was collected via interviews with managers and staff (number: 68) and participant observation (number: 120).

Results	The data showed a clear trajectory of development that started with a lack of diversity activity and attention, and progressed through a transitional phase in which diversity was described as a 'hobby' with the organisation paying some attention to diversity, with a good degree of enthusiasm from less senior staff, but missing clear policies and attention from senior managers. Finally, over time this hobbyism developed into a top-down system that, when combined with the enthusiasm of staff already present lower in the organisation, resulted in an abundance of ideas and a strong drive towards change and implementation, in a phase referred to as 'integration'.
Why is this important?	Provides insight into how qualitative data can be used to track diversity policy over time, and the factors that contribute to its development.

Science Foundation Ireland (SFI): Starting Investigator Grant (SIRG) gender initiative	
Data collection strand	Call for Evidence
Aim	To examine the effectiveness of a gender-based quota limiting the number of male researchers nominated for a specific early-career funding grant.
Method	SFI employed a time series approach to evaluate the impact of the SIRG initiative, which is an early-career funding award that is open to applicants who are three-to-eight years post-PhD without an academic position. Introduced in 2015, the initiative incentivises research bodies to nominate excellent female candidates by permitting a maximum of six (out of a possible 12) male candidates to the funding call (in other words, a quota is imposed on the gender composition of nominees to the scheme).
Results	Analysis of 2,778 applications from before and after the SIRG initiative was implemented showed an increase in the number of female applicants from 25% in 2013 to 47% in 2015, as well as an increase in female award holders from 27% in 2013 to 50% in 2015. The success rates for male and female applicants were very similar, both before and after the gender initiative was implemented. This data supports the fact that there were suitable female candidates available but that they were not being represented in the application pool.
Why is this important?	This example demonstrates how higher-level statistics such as application and success rates can be used to track an intervention's impact over time.

Strengths and limitations

Each non-experimental approach has its own advantages and disadvantages. For instance, a longitudinal study that follows the same group of participants over time may be able to look at which factors feed into outcomes such as career progression, performance and job satisfaction. However, longitudinal designs are difficult to employ as they are time-intensive and are subject to attrition. Moreover, the lack of experimental control in this approach means that answering

questions about whether a mentorship or training programme aided their progression depends on whether these interventions naturally occurred within the sample. However, when possible, a longitudinal design provides researchers with insight into trajectories, predictive relationships and the ability to statistically control for some of the 'noise'.

Qualitative approaches, such as the interviews conducted by Williams et al. (2014), also have their own set of advantages and disadvantages.

The main advantage of qualitative data is that it provides participants with the space to describe how an intervention affected them and thus provides insight into why an intervention is effective or ineffective. These methods are also well-suited to working with small sample sizes.

As the overarching aim of EDI interventions is to address structural barriers, the majority of equality and diversity initiatives are smaller in scale and aim to bring about change within a specific group of people, context or institution. When working with smaller samples or exploring the impact of a new intervention, these alternative qualitative evaluation methods may not only be more appropriate from a statistical point of view but also more informative about an initiative's effectiveness and efficiency.

6.5 Conclusions and recommendations

In sum, evaluations of EDI interventions were more effective when they:

- adopted multiple methods that allowed interventions to be viewed from different angles (in other words, a mixed-methods approach)

- considered a variety of outcomes or measured the outcome of interest in multiple ways (as seen in the meta-analyses by Kalinoski et al. and Bezrukova et al.)
- were able to examine the impact of an intervention both on its own as well as alongside other types of EDI intervention (in other words, the isolated impact of one intervention compared to a bundle of EDI interventions)
- included some degree of control, whether this be in the form of an experimental design (that is, a control versus an experimental group) or statistically controlling for extraneous variables related to the outcome measure.

In contrast, evaluations of EDI interventions were less successful or difficult to interpret when there was:

- a lack of structure or standard in reporting methodology and findings, with many sources not containing sufficient information to classify their approach and establish overall confidence in their findings
- only one type of evaluation method used (especially when this method was non-experimental) because of increased noise in the data and limited ability to isolate the effect of an intervention on outcome data collected.

Table 6.4. Summary of recommendations for measuring success.

Recommendations from this section	Policy makers	Funders	Employers	Research
Develop a clear framework or approach for reporting evaluations of EDI interventions, creating a centralised record or database of such evaluations (either within an institution, group of institutions or sector).	✓	✓	✓	
Encourage the use of qualitative data and mixed-method approaches to address limitations of quantitative methods (for example, issues surrounding small sample sizes, greater detail regarding how and why an intervention was effective, etc.).	✓	✓	✓	✓
Measure outcomes in multiple ways to gain a full picture of the intervention's impact.				✓
Consider other protected characteristics, sociodemographic factors and context when collecting data and examining the impact of an EDI intervention (as this will allow the influence of these to be taken into account statistically and the adoption of an intersectional lens).			✓	✓
Report effect sizes, or details regarding sample and descriptive statistics, wherever possible to enable future meta-analyses on the effectiveness of EDI interventions.	✓			✓

The background features a complex geometric design. A dark blue rectangle in the top-left corner contains the chapter title in white. The rest of the page is a mix of bright yellow and a blue grid pattern that recedes into the distance, creating a sense of depth and perspective.

Chapter 7: Enhancing data and disclosure

How can EDI data capture and disclosure rates in the international research and innovation landscape be improved?

7.1 Overview

Related to our discussion of methods and evaluation is the question of how data is used in EDI: to understand barriers and experiences, to provide comparison between sectors or organisations, and to measure impact and change. This chapter discusses data capture, disclosure and use in efforts to deliver EDI.

What is EDI data?

EDI 'data' is most commonly discussed in our sources as any individualised or aggregate data about identities or backgrounds of key populations (employees, grant applicants, postgraduate students, board membership and so on). In our review, 'equality monitoring' mainly focused on representation of different groups in different populations at any time, or tracking progression through processes, services or career pipelines. These could be across sectors or nations or within organisations (for example, two-yearly gender monitoring against targets in a large university faculty; Johnson et al., 2015). This may not answer all questions relating to inclusive experience and equitable relationships between those who are present in a population.

Our international sources also took place in varied legal and regulatory frameworks which impacted on how EDI data is:

- **categorised** (for example, there were national differences in recording ethnicity and race)
- **collected, stored and accessed** (often subject to data protection regulations)
- **drivers and the uses of the data:** reasons why the data is being collected for example, internal 'monitoring' (including benchmarking) or to directly support interventions (for example, prompting intervention, or use for quotas or as contextualised information in recruitment).

Some of these issues appeared in the discussion within our sources without detailed interventions or evaluation. However, themes arising throughout our study included:

- a desire for consistent, comparable datasets to understand and communicate EDI challenges (including benchmarking) and for understanding diversity and inclusion beyond basic population data
- increasing disclosure and ensuring meaningful categorisation of identity
- new ways to use EDI data to support interventions.

We discuss these themes further below.

7.2 Categorisation in the international context

Categorising identity characteristics or backgrounds (for example, different gender identities and different ethnic, national or migration backgrounds) can at the most basic level help organisations to:

- understand the diversity of their staff or stakeholders (presentation)
- identify barriers or gaps in progression or experience (through analysis of proportionality and representation at different stages of a process such as grant applications, or promotion)
- understand experience (for example, when responses to a staff or student survey are analysed with reference to equality data)
- compare any of the above with similar organisations or sectors.

Presentation of certain categories or subcategories may also serve to increase a sense of belonging in groups and individuals by 'recognising' their identities. Conversely, failure to present options which respondents could potentially have negative connotations. Changing cultural understandings of definitions and labels over time presents a key challenge to understanding long-term trends and measuring success across multiple cohorts (Pega, Reisner, Sell and Veale, 2017).

Our review highlighted the increased complexity of these issues in an international context. This has implications both for any comparative cross-national studies and for organisations working with international workforces or stakeholders.

Examples are discussed below.

7.2.1 Ethnicity, race, migration and national background

Ethnicity, 'race' and national and migration backgrounds had different uses and prevalence within our study. This included varying subcategories across organisations and jurisdictions, often due to different histories, cultures and legal understandings. Additionally some sources focused on migration history and/or language minorities (for example, Gill et al., 2018; KIF Call for Evidence submission) rather than primarily on ethnicity, colour or 'race'. Interestingly, KIF notes examples of a 'disconnect' between international diversity messaging around ethnicity and its own national context (Lindstad, 2017). Different nations are also clearly prioritising identities specific to their local context such as caste (SERB in India; see Metcalfe and Day, 2016), or indigenous or first-nation populations (CIHR in Canada, or Science in Australia Gender Equity (SAGE) in Australia).

This suggests a challenge in not only comparing different disciplinary fields cross-nationally but also tracking individual researchers or workers moving between nations or sectors (who could find their identity categorised inconsistently). We could theorise that the structural impact of racialised bias and inequality may then be harder to quantify and understand if relying solely on quantitative data. There are examples, however, of some comparative reports which have chosen high-level aggregation of 'white' and 'non-white' ethnic categories, for example as in McKinsey, 2016.

7.2.2 Gender, sex, gender diversity

Though more comparative studies and datasets appear in general in regard to 'gender equality' in R&I, we **note some areas of caution:**

- the binary nature of most indices (for example, ACE 2019; OECD 2017) (looking at 'men' and 'women' or 'male' and 'female') did not appear to recognise, record and communicate the experiences of other gender-diverse identities (such as non-binary or intersex people)
- understandings of how best to recognise the identities of trans people in equality monitoring whilst maintaining confidentiality, and relevant healthcare and research input, is a subject of much recent development (Pega et al., 2017).

- in some contexts, 'sex' and 'gender' appeared conflated, possibly through language or different conceptual understandings (for example, as to whether gender is binary), and it is possible that this could result in a mixed dataset; the implications of this could be that some datasets are including employees who are permitted to self-identify their gender, whereas an employee in a similar organisation who must record their identity based on formal documentation such as birth certificate; this distinction was not always clear within studies.

The implication of categorisation of identity characteristics relating to gender and sex identity is particularly relevant to **research outputs**. In Canada, CIHR has provided a range of practical toolkits and accessible communications to ensure better informed approaches to this work and reduce misunderstanding. In this sense **improving understandings of equality characteristics improves R&I outputs**.

Institute of Gender and Health (CIHR) resources for health equity	
Source	Call for Evidence, grey literature
Aim	To improve health equity through the integration of sex and gender in science. CIHR developed a range of resources to support researchers in their research approaches (including working with transgender research participants).
Method	While there has been no evaluation presented to date, this source offers comparative perspectives on whether or not to undertake action. However, the size of its impact on percentage increases in disclosure of certain characteristics was not quantified.
Results	Development of materials included online training modules such as 'integrating sex and gender in research for reviewers and applicants' and an online casebook to demonstrate the impact of integrating sex and gender in health research. Wider impact on health research is as yet unevaluated.
Why is this important?	Embedding awareness of equality data categorisation into research outputs and training. Considering the focus on clear communication and examples can be a useful way of supporting practice change beyond a new policy.

7.2.3 Disability

Our dataset included limited examples of evaluated interventions to support inclusion of disabled people. From a data perspective, some did focus on specific experience; for example, how training of line managers can reduce stigma around workers with depression (Hamann et al., 2016). Others,

however, did not clearly disaggregate different types of disability; for example, any differences between 'mental' or 'physical' disability (for instance, Gill, 2018) or different experiences of presentations of 'intellectual disability' (for instance, Judge and Gassett, 2015). We therefore suggest caution in comparative studies, greater clarification of terminology and tackling homogenisation in further

research. In terms of declaration or disclosure, we also know that, even within cultures, an individual may have similar experiences (such as marginalisation) but not identify as disabled: for example, not all Deaf/deaf or hearing-impaired people will identify as disabled (Woodcock, Rohan and Campbell, 2007). Evaluation and monitoring around disability may therefore need to consider both qualitative and quantitative methods to understand different identities and experiences, beyond either self-declaration or third party categorisation as 'disabled'. This could include perhaps the perceptions of others, broader identity categorisations for lived experience (perhaps providing a spectrum of identities, histories or understandings of health and wellbeing) and clarity as to when legal definitions (protected characteristics) are being used in the research or intervention.

7.2.4 Socio-economic background

Few of our sources contained evaluated interventions that were specifically related to socio-economic background. This may be a limitation of our search terminology. Where sources did discuss socio-economic background in some degree it was often alongside a range of other identity characteristics and framed in differing ways: for example, 'class' or 'poverty'.

Interestingly, though, a wide range of sources discussed financial support targeted at specific groups. While some of these could relate to value-signalling 'incentives', this could suggest an innate understanding of the relationships between financial resourcing and the ability to overcome

barriers, without expressing this need through an 'identity' or category of socio-economic status. Further research into the intersection of socio-economic status with the 'target' identities of these schemes (such as women, young researchers, parents) could be beneficial. Some of the examples in chapter 8 may prove useful here.

7.3 Encouraging disclosure

Related to the challenges of categorisation is that of disclosure or declaration. The reasons why someone may decide to share or withhold information about their identity characteristics are complex, as we have seen from some of the challenges around categorisation above. In addition, there is evidence in wider literature (Kim and Fredriksen-Goldsen, 2013; Rose, 2006; Stanley et al., 2011) of factors such as how and when questions are asked, data privacy and concerns about professional competency standards.

However, while many of our sources alluded to the need for 'increased EDI data', a specific focus on initiatives to disclosure practice was more limited than expected. This may reflect the wide variation of approaches to individual disclosure and monitoring categories internationally, as discussed above.

Where disclosure was discussed in detail, approaches focused on building trust and knowledge between the data-collecting organisation and individuals (data subjects), as well as technical considerations such as providing increased opportunities and methods for collection:

Gazmararian, Carreón, Olson and Lardy (2012). Exploring health plan perspectives in collecting and using data on race, ethnicity and language.	
Source	Academic
Aim	To increase disclosure and accuracy of ethnicity, race and language preferences through 'consumer education' of benefits of demographic data collection (utilising community and third party partners), staff training (including responding to concerns about disclosure) and collection at varied interaction points (beyond enrolment).
Method	Qualitative comparative interviews with organisations both collecting and not collecting this data. Limitations: effect size (such as percentage increases in disclosure of certain characteristics) not quantified. Strengths: comparative perspectives on whether or not to undertake action.
Results	Organisations collecting data: valued partnerships with NGOs and community groups within 'minority' communities and public education on benefits to the data subject from disclosure; and increased the number of interactions where informed discussions relating to disclosure could be prompted.

Why is this important?

Demonstrates the value of comparative organisational approaches to data collection. Considers drivers and restrictions relating to data collection (such as public education, resources, legal frameworks, data-sharing). Intra-organisation data sharing limitations and confidentiality are a key challenge.

For consideration:

Building trust: where and how is education and partnership useful in building trust in a non-consumer relationship (for example, employee data)? Organisational standards on data collection: self-disclosure was considered 'gold standard' but organisations alternatively stored proxy equality data derived from other sources (such as geocoding and surname identification software). It is possible that a multi-tiered approach to data sources could impact on drivers around direct disclosure and/or trust in the data collector. Organisations should consider their own legal and ethical frameworks around equality data collection and storage beyond self-disclosure (for example, in research data).

Where equality data is used for intersectional analysis, it should be noted that there are some indications that disclosure rates on **specific** categories, separate to engagement overall with equality monitoring, may vary between different groups; (for example, in one study sexual orientation disclosure varied by race and ethnicity in public health surveys (Kim and Fredriksen-Goldsen, 2013). Further research on patterns of non-disclosure across a wide range of characteristics in relation to different variables (individual background, institutional trusts) within R&I may be useful.

7.4 Improving datasets

As discussed, there are a number of challenges around collating 'global' data on EDI in research or innovation including different understandings of categorising identity characteristics. However, we make no assumption that the collection of certain types of identity data is unanimously supported (or permitted) in all jurisdictions: for example, we know that patent applications by ethnicity in the US is not systematically collected (Klingler-Vidra, 2019), and within European overarching legal provisions on equality there are still a wide variety of national (and federal) prohibitions around ethnicity, religion and migration or national background (Farkas, 2017). Data collection by declaration (or 'disclosure') by the data subject as opposed to third party identification may also vary (see Gazmarian et al, above).

However, our sources point to a general consensus that data collection can be advantageous to EDI work. For example, the creation of national-level datasets in Norway (KIF) and a range of demographic indicators in Sweden (see Klingler-Vidra, 2019) have been cited as a means of raising awareness of underrepresentation and therefore the need for public prioritisation of change.

7.4.1 Future approaches

In R&I, a recent literature review commissioned by the Wellcome Trust (Chambers et al., 2017) noted that EDI metrics in (mainly) clinical and biomedical

research were focused at individual level, rather than at aggregate or organisational level, and lacked appropriate measurements for multiple dimensions or intersecting dimensions of inequality. The need for more longitudinal datasets was also noted. A model which links 'diversity of the research workforce to research outputs' is discussed in detail, considering both supportive and obstructing factors for elements such as the research staff pipeline, retention and high-quality research projects on inequalities.

Examples of exploring new approaches to EDI data include proposals for 'composite metrics' within higher education (McLaughlin, McLaughlin and McLaughlin, 2015). A framework for an index of 'multidimensional diversity' within a higher education setting was proposed, building on 1,500 colleges and universities, based on the Simpson diversity index (which would calculate, if any two individuals were selected from an institution, the probability that they were of a different socio-economic background, for example). The aim was to inform policies which look beyond "the magnitude and proportion of race and ethnicity" and look at a wide range of variables (such as gender, if first-time study, age, ethnicity, full-/part-time) along with factors such as staff-student ratios, retention and graduation. The study noted that both the variables entering that model and the model results would need to be flexible enough to meet the varied 'missions, visions and goals of individual institutions. This reflects the challenges involved in cross-institutional and cross-sector comparisons: firstly, defining what element of 'EDI' is being examined and, secondly, establishing to what extent 'results' can and should be contextualised (whilst still giving meaning to the comparative exercise).

Even where data is collected, however, one source highlighted how **restrictions on data exchange and publication impeded the progress of collaborative EDI projects** between private organisations (Evans and Glover, 2012). This included concerns including concerns about

reputation and commercial sensitivity. It is recommended that data collection sharing agreements for such projects are agreed at the outset.

7.5 Data for decision making

Beyond 'monitoring' or encouraging understanding and progress through benchmarking or group and sector data sharing, there is evidence of EDI data being used to undertake specific interventions. Within our sources this was primarily in relation to how EDI data can inform affirmative action (that is, the data is used as context about an individual which informs steps to mitigate structural biases that may be in place). This has taken place in, for example, quota systems or contextual recruitment policies (see, for instance, Kilango et al., 2017). From wider academic literature we also know that beyond 'entry' decisions, data can also be used for performance interventions, such as support for mental health challenges or retention of underrepresented groups (see, for example, discussions by Agnihotri and Ott, 2014; Bennett, 2018; Prinsloo and Slade, 2017). However, the sources did not fully reflect the trend and it may be that further evaluation of such data-driven EDI initiatives in the R&I context is called for.

Specifically, a range of scholarship is being undertaken to address some of the legal and ethical challenges around the use of such data, including the challenge of algorithmic bias stemming in part from a lack of diversity within technology providers, and in part from unrepresentative datasets (see Ajunwa, 2018; Barocas and Selbst, 2016). This includes the creation of a number of centres for establishing ethical frameworks and development, such as the Data Transparency Lab (DTL) focused on international approaches to innovation, entrepreneurship academia and research, or the UK's Jisc ethical use of learner analytics. Ironically though, academics have pointed out that greater EDI data collection is required for the very reason of reducing algorithmic bias (Zliobaite, 2016).

7.6 Conclusions and recommendations

Our review found limited detailed studies of interventions surrounding disclosure in R&I, though gender-related initiatives in particular tended to be supportive of increasing the volume and consistency of gender-related datasets to help monitor progress.

A wide range of understandings and terminology exist around different equality issues and identities. This also has implications for:

- confidence in comparative datasets cross-nationally
- data disclosure and analysis when organisations are working with internationally diverse staff and stakeholders (ensuring understandings and interpretations, and considering adaptability to support the nuances of different identity constructions)
- research outputs: how EDI impacts on research methodology and analysis.

Some attempts to trial different metrics surrounding EDI have been discussed above; further work on the suitability of these tools for different R&I contexts would be valuable.

Data disclosure and use may be more effective when:

- terminology and categorisation are relevant to local contexts
- trust is built between individuals and organisations (potentially using trusted third parties)
- organisations provide a variety of opportunities to disclose
- there is clear communication on the aims of the data collection and there are limitations on transparency and sharing
- supporting specific targeted and affirmative action measures.

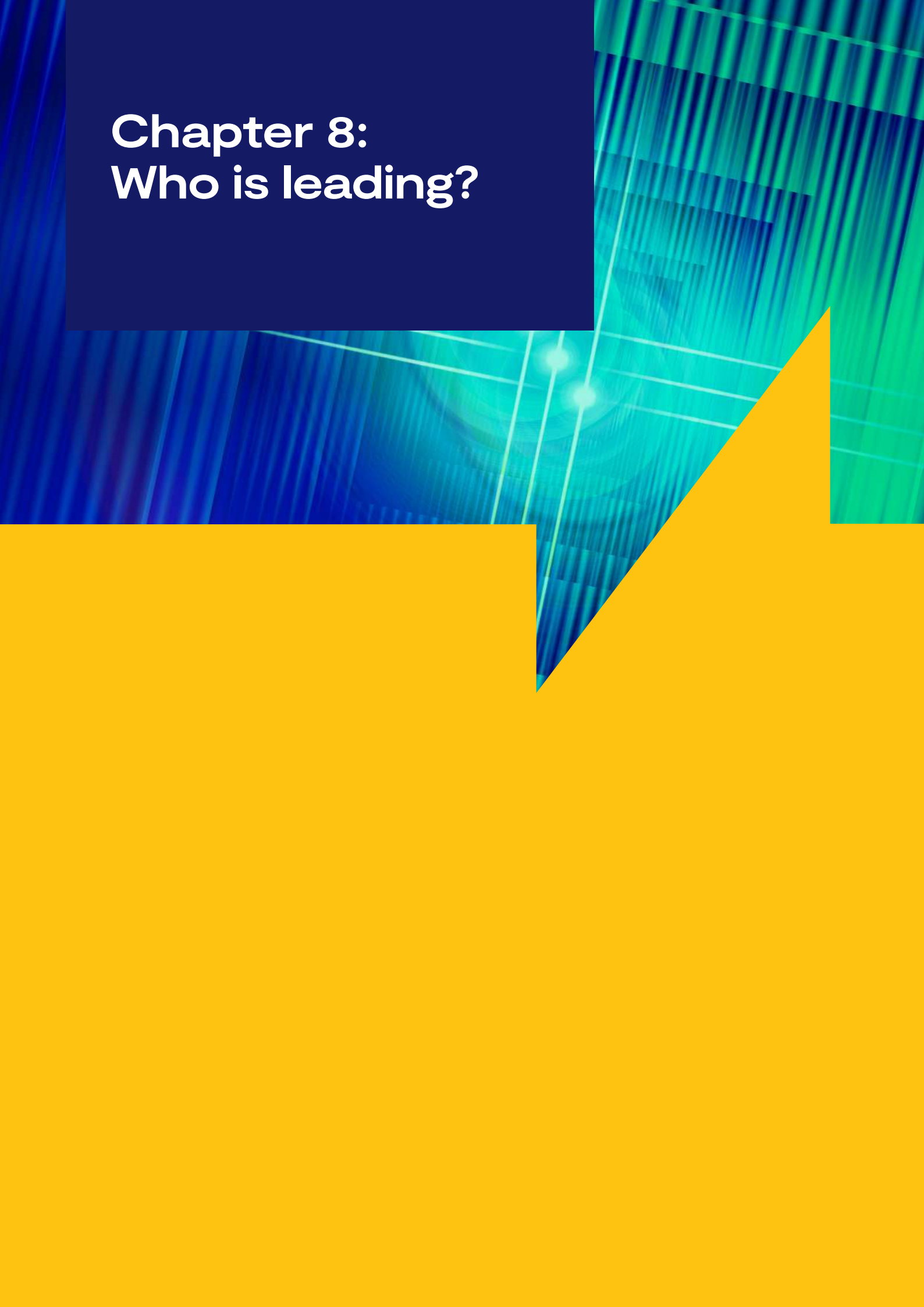
Data disclosure and use may be less effective when:

- comparative or cross-national datasets are unclear in terms of how categorisation is understood.

Table 7.1 Summary of recommendations for enhancing data and disclosure.

Recommendations from this section	Policy makers	Funders	Employers	Research
Reflect on the limitations and opportunities around identity categories and consult with individuals and groups on understanding and preferences. For internationally diverse workforces (or cross-national work), check understandings of both categories and reasons for collection.	✓	✓	✓	
Conducting further research on patterns of non-disclosure across a wide range of characteristics in relation to different variables (individual background, institutional trusts) within R&I may be useful.				✓
Model different approaches to EDI metrics and measurements (including intersectionality) and consider suitability for the R&I context.				✓
Develop key indices for the R&I field, sufficient to allow comparison and benchmarking but flexible enough to account for national and sector contexts and legislative prohibitions.	✓	✓	✓	✓
Increase outreach and researcher professional development on the importance of EDI data literacy on research outputs.	✓	✓	✓	✓

Chapter 8: Who is leading?

The background of the slide is a complex geometric composition. It features a dark blue rectangular area in the top-left corner containing the chapter title. The rest of the slide is composed of various shades of blue and yellow. A prominent yellow triangle points upwards from the bottom-right corner. Overlaid on these solid colors are several thin, white, perspective-distorted lines that create a sense of depth and movement, resembling a grid or architectural structure. The overall aesthetic is modern and technological.

How can EDI data capture and disclosure rates in the international research and innovation landscape be improved?

8.1 Overview

The review interpreted 'leading on EDI in the R&I context' in a broad manner. This was informed by early stakeholder conversations, as well as reflection on the challenges of 'measurement' of EDI progress (see chapter 7), and the limitations within the dataset of evaluation which determined effect sizes and impact (see chapter 6).

We chose to consider three broad indicators of leadership to enable identification of organisations leading in different, yet important, ways in the R&I context:

- **quantity:** organisations that appear more than once in our final dataset, indicating that they have delivered multiple EDI interventions and either published or returned information to us about them; this indicates that these organisations are likely to be prioritising, investing in and raising the profile of EDI in the R&I context
- **innovation:** organisations that have introduced EDI interventions which are innovative or trailblazing in the R&I context; these organisations can be seen as leading the way in developing and trialling new approaches or considering areas of EDI that often receive less attention
- **wider impact:** organisations that are influencing advancement of EDI within other organisations in the R&I context, for example through development of strategy, policy or interventions for other organisations; these organisations are encouraging, or sometimes enforcing, consideration of EDI by external organisations.

For this analysis, we focused on the sources included in the dataset that met our eligibility criteria and were analysed using our evaluation framework. We did not consider the sources highlighted under research question one.

8.2 Quantity

As noted, many of our sources reviewed multiple interventions (for example, different types of intervention within an organisational context). We separated publications that were looking at multiple interventions so that each intervention was represented on a single row, to examine which organisations and authors had contributed to the current database in more than one instance. In other words, while we included 109 sources (82 academic, 13 from grey literature sources and 14 responses to the Call for Evidence), together these covered 130 interventions.

As the majority of these stemmed from the

extensive search of academic literature, we first highlight the authors who contributed to more than one publication as well as those who investigated more than one intervention within a single publication, both of which are intended to signpost which areas of EDI are currently being investigated and by whom. Next, we turn the spotlight on organisations from the grey literature sources and the Call for Evidence that contributed to more than one source or delivered multiple EDI interventions, as these organisations can be viewed as prioritising EDI.

8.2.1 Academic sources

As the majority of sources were from the academic literature search, it was integral that we looked at the full list of authors for each publication to check which were involved in more than one source. There were only a handful of instances where authors contributed to more than one of the academic publications, with the majority being across a set of seven authors from five institutions in the US (Purdue University, Indiana University: Purdue University Indianapolis, Skidmore College, Yale University and University of Wisconsin-Madison) collectively piloting and examining the efficacy of a video-based diversity training intervention (Hennes et al., 2018; Moss-Racusin et al., 2016; Pietri et al., 2018; Pietri et al., 2017). Beyond these, an additional four authors contributed to two separate publications, one concerning gender equality initiatives in Tanzania (Kilango et al., 2017; Nyoni et al., 2017), another looking at diversity management in the customer service industry in the US (Madera, 2013; Madera et al., 2013), and two other authors co-authored two of the academic publications in the current database on the Workshop for Gender Equity Simulation (Shields, Zawadzki and Johnson, 2011; Zawadzki et al., 2012).

There were also nine academic sources that evaluated more than one intervention within a single publication, although these varied with regard to the rigour of their evaluation methodology and the amount of data or information presented as evidence. Moreover, some of these publications, such as Hegewisch and Gornick (2011), were reviews (rather than presenting an analysis of primary data). The authors highlighted above and publications noted in table 8.1 are presented in greater detail in chapter 4, with a focus on the interventions being evaluated, and in chapter 5 for their adopted methods of evaluation; however, we chose to present these here as well to serve as an additional signpost for future reviews and EDI research.

Table 8.1. Sources discussing or evaluating more than one intervention

Publication reference	Number of interventions evaluated
Williams, Kilanski and Muller (2014)	4
Hegewisch and Gornick (2011)	3
Phillips, Deiches, Morrison, Chan and Bezyak (2016)	3
Schiebinger and Schraudner (2011)	3
Seko, Kidd, Wiljer and McKenzie (2014)	3
Webster, Adams, Maranto, Sawyer and Thoroughgood (2018)	3
Bailyn (2011)	2
Eriksson-Zetterquist and Renemark (2016)	2
Nyoni, He and Yusuph (2017)	2

8.2.2 Leaders within the grey publications and Call for Evidence

Within the sources identified through the grey publication search and the Call for Evidence, five organisations were represented more than once

(see table 8.2). This sample of leaders consisted mostly of sector agencies and research funding organisations, with most being involved in both realms.

Table 8.2 Organisations represented more than once in the evaluated sources.

Name of organisation	Number of appearances
Canadian Institute for Health Research	4
Swiss National Science Foundation	4
Max Planck Society	3
Universities New Zealand Te Kāhui Amokura	3
German Research Foundation (DFG)	2

There are key limitations to selecting these organisations as ‘leading’: for example, our Call for Evidence relied on ‘snowballing’, had a limited timeframe and was issued in English, which may bias our sample towards larger resourced organisations, those more familiar with our work and context, and those ready to respond in English.

However, analysis does suggest useful approaches from organisations embedding EDI throughout multiple initiatives and activity:

Canadian Institute of Health Research	
Data collection strand	Call for Evidence
Aim	To embed gender equity and indigenous and first-nation equality into organisational practice, research and funding.
Activity	■ Implementation of a 'gender-based analysis' with regard to decision making within the workplace, the research action plan, and research funding. Includes training for staff on this approach. http://www.cihr-irsc.gc.ca/e/50970.html ■ Specialist centres: Institute of Indigenous Peoples' Health and Institute for Gender Research, supporting equity in research (and targeted support for underrepresented researchers). ■ Working with other national research bodies in Canada (through the Canadian Research Coordinating Committee or CRCC) to improve equitable access to research funding.
Why is this important?	Systematic approach i) embedded equity across research funding and policy making, ii) was supported by specialist institutes and ii) collaborated with other research bodies. Integration of EDI aims into research output and identity.

The Max Planck Society	
Data collection strand	Call for Evidence
Aim	To support gender equity and work-life balance.
Activity	■ Code of conduct regarding sexualised harassment (supported by line manager casework coaching). https://www.mpg.de/11961177/code-of-conduct-en.pdf ■ Gender-inclusive and gender-neutral language policy. ■ Range of practical and financial support for caring responsibilities (not limited to childcare) for employees and grant holders. ■ High-profile affirmative and positive action schemes for women researchers (such as the Lise Meitner Programme). https://www.mpg.de/11767653/lise-meitner-programme
Why is this important?	Breadth of activities supporting gender equality, steps to tackle sexualised harassment, and family or caring responsibilities as employer and research funder. Mixed approach: initiatives to normalise an inclusive daily working culture, and high-profile action targeted at underrepresented groups.

8.3 Innovative approaches

In this category, we focused on organisations which were:

- focusing on identity characteristics that often receive less attention
- using new or inventive approaches or tools.

Leadership through innovation in EDI can be subjective: what may be innovative for one sector or country may not be for another. Within the body of our sources, however, we note the following:

Swiss National Science Foundation	
Data collection strand	Call for Evidence
Aim	To support gender equity through setting expectations on joint childcare responsibilities; to respond to local contexts.
Description	‘Flexibility grants’ offered to researchers (men and women) to help balance professional and private lives by paying for childcare or to finance the salary of a support staff member so that the award-holder can reduce their working hours. The award is only available if (in dual partnerships) an equal contribution is made to childcare, to reduce gender disparity in dual-career couples. This is undertaken in the context of a limited and expensive national childcare infrastructure impacting on primarily women’s careers.
Method	Not yet evaluated.
Results	Not yet evaluated.
Why is this important?	Using policy to nudge changes to gender norms in childcare. Considering how the conditions of financial awards relating to EDI aim to ‘nudge’ behaviour change specific to national and sector contexts.

Science Foundation Ireland (SFI): Starting Investigator Grant (SIRG) gender initiative	
Data collection strand	Call for Evidence
Aim	To increase gender equity in grant applications (with implications for age) to facilitate the retention of excellent female researchers within academia.
Description	SFI noted that applications to a grant scheme were not proportionate to the gender breakdown of the PhD pipeline. A change was made to the nomination scheme for a grant scheme for early-career researchers to incentivise more female nominees: initially five applications per institution were permitted with no gender specifications. In two pilot schemes, a maximum of 12 nominations could be made, but with a cap of six male candidates. Upon submission to SFI, all applications are treated equally regardless of the gender of the applicant.
Method	Longitudinal quantitative data analysis. No control for other factors.
Results	Female applicants increased from 25% to 47% in the first two years of the pilot (with awardees rising from 27% to 50%). Similar success rates were achieved by men and women. A second-round pilot is in place.
Why is this important?	Increasing gender diversity of grant applications through research funding guidance. ■ Introducing scheme as piloting. ■ Acknowledging the resource requirements of a widening pool of nominees (cost deemed to be minimal in relation to impact achieved).

Beyond organisational activities, we acknowledge that a range of experimental interventions may also be of interest when adopted and evaluated. For example:

Bohnet, Bazerman and van Geen (2016): When performance trumps gender bias: joint vs. separate evaluation	
Data collection strand	Academic
Aim	To increase gender equity in grant applications (with implications for age) to facilitate the retention of excellent female researchers within academia.
Description	An 'evaluation nudge' whereby recruitment panels consider applicants under a joint evaluation of applicants, rather than focusing on individual applicant performance. Theoretically this decision making and information processing aimed to reduce stereotypes, particularly at more senior and managerial levels when this type of evaluation is less common.
Method	Experimental conditions (number: 554 'employers') with a control group, and consideration of gender differences in task performance as well as biases. An additional small control experiment with new information provided later at the decision making stage.
Results	Employers were significantly more likely to choose a higher-performing employee without gender stereotyping in joint evaluation methodology.
Why is this important?	Evidence-based approaches to diversifying recruitment and consideration of new ways of working to improve decision making in recruitment.

8.4 Wider impact

Finally, we present examples of organisations that have created wider impact on EDI in research and innovation through work at a national or regional level. These are generally EDI organisations, sector agencies or funding bodies and their actions serve to encourage and enable stakeholder institutions to 'deliver' EDI work. We found that the organisations that could be considered to fall into this category of leadership had delivered the following types of EDI work:

Integrating EDI into research and/or innovation policy or funding frameworks:

- **Science Foundation Ireland:** SIRG gender initiative
- **Max Planck Institute:** Lise Meitner Excellence Programme
- **Swiss National Science Foundation:** gender equality related funding grants.

Providing EDI recognition schemes to drive improvement within institutions:

- **SAGE:** a national (pilot) programme for gender equity in Australia utilising an Australian Athena SWAN Charter; the authors note that Athena SWAN is owned by Advance HE.

Providing EDI engagement, training and development to upskill other organisations on EDI:

- **National Science Foundation (US):** grants to support gender equity (discussed in, for example, O'Meara, 2019)
- **Wellcome commissioned research:** Chambers et al. (2017) - Review of diversity and inclusion literature and an evaluation of methodologies
- **CRCC:** collaborative approach to creating an action plan and priorities to strengthen EDI in research
- **KIF (Norway):** creation and work of independent advance committee; 'watchdog' for institutions and national authorities; provision of independent guidance to institutions on EDI efforts (for example, through restructuring)
- **Organisation for Women in Science in the Developing World:** national assessments on gender and STI (from 2012).

8.5 Conclusions and recommendations

By taking three examples of approaches to 'leadership' in EDI we aim to prompt reflection on how leadership in EDI is conceived. Each of these approaches taken individually have their limitations. For example:

- quantity of activity does not necessarily equate to impact, and 'visibility' of activity may bias larger resource-rich institutions. Multiple activities may also make evidence of causation and correlation challenging
- innovation without an evidence base may not always be best practice and can be highly context-specific
- influence and support for a broad spectrum of R&I organisations from one source may be seen as 'leading'. However, from the perspective of those organisations, synergy amongst

influencers (in terms of policy directions and grant requirements) may be more practical and resource-friendly.

However, we hope that, taken holistically, the breadth of 'leadership' can be seen. Additional areas of leadership to be explored in future could include:

- use of an evidence-based approach to policy and practice making (including linking academic research and practitioners)
- success against a range of sector indices on EDI (potentially context- and nation-specific)
- reputational rankings (potentially for underrepresented groups)
- EDI activity as a proportion of overall resource.

Table 8.3 Summary of recommendations for leading organisations.

Recommendations from this section	Policy makers	Funders	Employers	Research
Reflect on definitions of 'leadership' and future methods for recognising and rewarding innovation, commitment and collaborative practice.	✓			

Chapter 9: Conclusions



9.1 Conclusions

This review set out to develop greater understanding of EDI in R&I by identifying:

- organisations which have explored EDI challenges in the R&I context
- the approaches currently being utilised to tackle these challenges, and how these are evaluated
- how EDI data collection and disclosure can be enhanced
- organisations that are leading across these areas.

There were some common approaches towards EDI interventions such as training or mentoring (inclusive or targeted), policy change or use of targeted action through grant schemes or representation quotas. We have provided examples of a range of different activities and contexts. However, evaluation of these changes varied in rigour and consistency, with outstanding questions particularly on sustainability and long-term impact on both intended beneficiary groups and general attitudes and experiences of inclusive environments. Academic literature has sought to fill some of these gaps with meta-analyses and new approaches, but this remains a key priority for future evidence-based policy approaches.

Considering previous work, leading organisations and data challenges, it is clear that there are multiple ways of understanding EDI 'progress' or 'status', including a range of proposed metrics and frameworks to recognise change, and sharing of best practice and case studies. The lack of consensus can make comparison between organisations, sectors and nations challenging. However, contextualisation of challenges and identities remains an important factor for both understanding issues and designing and delivering effective interventions.

9.2 Gaps in the evidence base

Across all sections of this report, gender, and primarily women, was the focus of the sources and interventions identified. EDI in general was the next most commonly addressed characteristic, which may signal a holistic understanding of minority experiences, but this is often unclear and risks losing sight of particular challenges and the needs of certain groups (or intersecting structures of disadvantage). A focus on disability, caring responsibilities, LGBTQ and religious inclusion and age in the R&I context is a key suggestion for further research. Gaps in monitoring and benchmarking data for some identity characteristics (for legal or cultural reasons) is likely a contributory factor to current gaps. There was a stronger evidence base around research in academic settings, and a disciplinary focus on STEM. Many of our sources also focused on organisations influencing national or regional settings rather than on cross-national institutes or organisations.

Comparing the UK and international reviews

A separate review focusing on the UK context found a stronger source base in grey literature and responses

to a Call for Evidence than in academic sources specifically evaluating the UK R&I context; this may impact on understandings of evidence methodology as well as concepts of leading 'organisations'. Together, however, the two reviews reveal what works in R&I in different contexts and suggest enormous scope for transfer of effective practices.

9.3 Review limitations and avenues for future work

It is important to highlight that the review focused on existing literature and did not aim to survey or compare the EDI work of all organisations acting within the R&I landscape, nor did it assess the totality of their EDI work. The benefits and challenges of exploring this further have been discussed and are included in our recommendations. We could not include those organisations that had not published their EDI work or responded to our Call for Evidence. Additionally, the research primarily focused on the primary lens of EDI interventions (in practice, or experimental) rather than the primary lens of evaluating any one organisation (for example, an 'audit' of EDI strategies or issues). The latter was still an important factor in understanding 'what works' and 'who is leading', and the limited Call for Evidence provided some useful contextual perspectives on this. Throughout, our evaluation framework aimed to capture a sense of key 'drivers', enablers and limitations to effective implementation.

While attempts were made to mitigate biases towards the Global North, it is clear that our sources still present a dominance of North American, European and Australasian perspectives. This includes both the academic literature and organisations responding to our Call for Evidence or presenting detailed information on their websites in English. A more extensive review of any of the issues within this review may benefit from a longer timeframe to make use of global approaches through networks of researchers and illuminate further policies and procedures unpublished on websites (or not published in English).

The interdependence of multiple interventions is also often unclear (see sections 3.4.5 and 6.4). A range of intervention types within any one organisation or context may help address different aspects of structural and cultural discrimination or bias, but further work to increase the evidence base for different approaches is recommended. This is of particular interest in the R&I sector where any one environment (for example, a research unit) may also be subject to multiple external drivers and interventions (for example, from the host institution, research funders, policy makers, learned societies or external talent pipelines). A 'mapping' of EDI strategy across R&I may need further exploration to understand these relationships (sections 7.4, 7.5). Key areas for future research have been noted within the following recommendations and we hope that reflections on methods will add to wider literature.

Chapter 10: Recommendations

The background of the slide is a complex geometric pattern. It features a dark blue rectangular area at the top left containing the chapter title. The rest of the slide is a vibrant yellow, separated from the blue area by a diagonal line. The blue area also contains faint, lighter blue geometric patterns, including hexagons and dotted lines, which are part of the overall design.

A summary of the recommendations for actions and further research that have been made throughout this review is set out in the following table:

Recommendations from this section	Policy makers	Funders	Employers	Research
Organisational investigation of EDI in R&I				
Co-ordinate reviews to focus on key gaps on EDI challenges and interventions, to include: ■ disability, religious identity, language in minority and multilingual workplaces, gender and sexual diversity, and age; gender research also to ensure cross-national knowledge-sharing on R&I areas of research or job function where men are a minority ■ disciplines beyond STEM ■ R&I in the Global South ■ intersectional approaches to understanding EDI challenges and implementing change.	✓	✓		✓
Explore repositories for wider EDI literature in R&I (beyond gender), with consideration of accessibility and language options.	✓	✓	✓	
Consider catalyst and grant support for ongoing EDI research or knowledge exchange (including underrepresented communities).	✓	✓		
Interdisciplinary approaches and cross-disciplinary communication.				✓
Toolkit or support for researchers looking at comparative EDI across multiple organisation types in R&I.	✓	✓		✓
What works?				
Foster collaboration within and between organisations to improve issues surrounding limited resources, methodological issues, duplication of work and sharing of good practice.	✓	✓	✓	
Additional research needed to clarify the effectiveness of particular interventions, including family-friendly policies, mentorship programmes and employer engagement via EDI committees and advisers.			✓	✓
Embed EDI awareness and initiatives into organisational culture, ensuring alignment between an organisation's communicated messages and intentions, and actual practice.	✓	✓	✓	
Involve all levels of staff in the implementation of EDI interventions, with strong and demonstrable commitment from senior management and support for staff directly involved in overseeing and evaluating interventions.		✓	✓	
Make certain that the intervention is clearly defined (for example, clear agendas and outcomes for training programmes, detailed records of new policies and how to go about implementing them, etc.).	✓	✓	✓	✓

Recommendations from this section	Policy makers	Funders	Employers	Research
Measuring success				
Increase outreach and researcher professional development on the importance of EDI data literacy on research outputs.	✓	✓	✓	✓
Develop a clear framework or approach for reporting evaluations of EDI interventions, creating a centralised record or database of such evaluations (either within an institution, group of institutions or sector).	✓	✓	✓	
Encourage the use of qualitative data and mixed-method approaches to address limitations of quantitative methods (such as issues surrounding small sample sizes, greater detail regarding how and why an intervention was effective, etc.).	✓	✓	✓	✓
Measure outcomes in multiple ways to gain a full picture of the intervention's impact.				✓
Enhancing data and disclosure				
Consider other protected characteristics, sociodemographic factors and context when collecting data and examining the impact of an EDI intervention (as this will allow the influence of these to be taken into account statistically and the adoption of an intersectional lens).			✓	✓
Reflect on the limitations and opportunities around identity categories and consult with individuals and groups on understanding and preferences. For internationally diverse workforces (or cross-national work), check understandings of both categories and reasons for collection.	✓	✓	✓	
Conducting further research on patterns of non-disclosure across a wide range of characteristics in relation to different variables (individual background, institutional trusts) within R&I may be useful.				✓
Model different approaches to EDI metrics and measurements (including intersectionality) and consider suitability for the R&I context.				✓
Who is leading?				
Develop key indices for the R&I field sufficient to allow comparison and benchmarking but flexible enough to account for national and sector contexts and legislative prohibitions.	✓	✓	✓	✓
Reflect on definitions of 'leadership' and future methods for recognising and rewarding innovation, commitment and collaborative practice.	✓			

Bibliography

The background of the slide is a complex abstract design. It features a dark blue rectangular area in the top-left corner containing the title. The rest of the slide is composed of various shades of blue, including a grid pattern on the left and a series of radiating lines on the right. A large, solid yellow shape, resembling a stylized arrow or a corner cut, points from the bottom-left towards the center-right of the slide.

Bibliography

* Source identified through method but not evaluated with framework.

Background or contextual source not identified through method or evaluated.

References from this report are included below.

A full list of all sources which were included in the final dataset (that is, found eligible and evaluated) is set out in a public list at

https://www.zotero.org/groups/2326259/review_of_ed_i_interventions/items

Adriaanse, J., Schofield, T., (2014). *The impact of gender quotas on gender equality in sport governance. Journal of Sport Management*, 28(5), pp 485-497. Retrieved from bsu.

Agnihotri, L., Ott, A., (2014). *Building a student at-risk model: an end-to-end perspective from user to data scientist. Proceedings of the 7th International Conference on Educational Data Mining*. London, UK: International Educational Data Mining Society.

Ahmed, S., (2012). *On Being Included: Racism and Diversity in Institutional Life*. Duke University Press.

Ajunwa, I., *Age Discrimination by Platforms* (May 24, 2018). 40 Berkeley J. Emp. & Lab. L. (2019 forthcoming). Available at SSRN: <https://ssrn.com/abstract=3142979>

Alhejji, H., Garavan, T., Carbery, R., O'Brien, F., McGuire, D., (2016). *Diversity Training Programme Outcomes: A Systematic Review. Human Resource Development Quarterly*, 27(1), pp 95-149. <https://doi.org/10.1002/hrdq.21221>

American Council on Education (ACE), (2019). *Race and Ethnicity in Higher Education*. Retrieved from <https://www.equityinhighered.org/resources/report-downloads/>

Archibong, U., Eferakorho, J., Darr, A., Scally, A., Atkin, K., Baxter, C., Johnson, M.R.D., Bell, M., Waddington, L., Wladasch, K., Bedard, T., Oluyinka, A., Sahrps, P., Bradshaw, P., (2009). *Perceptions of the Impact of Positive Action in EU and non-EU Countries. International Journal of Diversity in Organisations, Communities and Nations*, 9 (5), pp.111-124.

Arday, J., Mirza, H. S., (2018). *Dismantling Race in Higher Education: Racism, Whiteness and Decolonising the Academy*. Switzerland: Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-60261-5>.

Bailyn, L., (2011). *Redesigning work for gender equity and work-personal life integration. Community, Work & Family*, 14(1), pp 97-112. <https://doi.org/10.1080/13668803.2010.532660>

Barocas, S., Selbst, A. D., (2016). *Big Data's Disparate Impact. California Law Review*. Available at SSRN: <https://ssrn.com/abstract=2477899> or <http://dx.doi.org/10.2139/ssrn.2477899>

Bauman, M. D., Howell, L. P., Villablanca, A. C., (2014). *The Women in Medicine and Health Science program: An innovative initiative to support female faculty at the University of California Davis School of Medicine. Academic Medicine*, 89(11), pp 1462-1466.

Beaurain, G., & Masclet, D., (2016). *Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence. European Economic Review*, 90, pp 350-362. <https://doi.org/10.1016/j.eurocorev.2016.04.009>

Behaghel, L., Caroli, E., Roger, M., (2014). *Age-biased technical and organizational change, training and employment prospects of older workers. Economica*, 81(322), pp 368-389. Retrieved from ecn.

Bell, S. T., Villado, A. J., Lukasik, M. A., Belau, L., Briggs, A. L., (2011). *Getting Specific about Demographic Diversity Variable and Team Performance Relationships: A Meta-Analysis. Journal of Management*, 37(3), pp 709-743. Retrieved from bsu.

Bennett, J. B., (2018). *Integral Organizational WellnessTM: An evidence-based model of socially inspired well-being. Journal of Applied Biobehavioral Research*, 23(4). <https://doi.org/10.1111/jabr.12136>

Berrey, E., (2014). *Breaking Glass Ceilings, Ignoring Dirty Floors: The Culture and Class Bias of Diversity Management. American Behavioral Scientist*, 58(2), pp 347-370. <https://doi.org/10.1177/0002764213503333>

Bezrukova, K., Spell, C. S., Perry, J. L., Jehn, K. A., (2016). *A meta-analytical integration of over 40 years of research on diversity training evaluation. Psychological Bulletin*, 142(11), pp 1227-1274. <https://doi.org/10.1037/bul0000067>

Bielby, WT., Krysan, M., Herring, C. (2013). *How Americans view workplace antidiscrimination interventions: Why we need a new conversation about race, gender, who wins, who loses, and what works*. Paper presented at Ford Foundation Research Workshop; New York, NY. 2013 Aug.

Bieling, G., Stock, R. M., Dorozalla, F., (2015). *Coping with demographic change in job markets: How age diversity management contributes to organisational performance. Zeitschrift Für Personalforschung*, 29(1), pp5-30. <https://doi.org/10.1177/239700221502900101>

Bohnet, I., van Green, A., Bazerman, M., (2015) *When Performance Trumps Gender Bias: Joint Versus Separate Evaluation. Management Science*, 62 (5), pp 1225-1234.

- Brunsma, D. L., Iyall Smith, K. E., Gran, B. K., (2015). *Expanding the human in human rights: toward a sociology of human rights*. <https://doi.org/10.4324/9781315634746>
- Byron, K., & Post, C., (2016). *Women on Boards of Directors and Corporate Social Performance: A Meta-Analysis*. *Corporate Governance: An International Review*, 24(4), pp 428-442. Retrieved from bsu.
- Canadian Institutes of Health Research and Institute of Gender and Health (Canada), (2012). *What a difference sex and gender make: a gender, sex and health research casebook*. Retrieved from <https://central.bac-lac.gc.ca/.item?id=MR21-164-2012-eng&op=pdf&app=Library>
- Cairney, P., Oliver, K., (2017). *Evidence-based policymaking is not like evidence-based medicine, so how far should you go to bridge the divide between evidence and policy?* *Health Research Policy and Systems*, 15, 1-11. DOI 10.1186/s12961-017-0192-x
- Carnes, M., Devine, P. G., Isaac, C., Manwell, L. B., Ford, C. E., Byars-Winston, A., Sheridan, J., (2012). *Promoting institutional change through bias literacy*. *Journal of Diversity in Higher Education*, 5(2), pp 63-77. <https://doi.org/10.1037/a0028128>
- Casey, C., Skibnes, R., & Pringle, J. K., (2011). *Gender equality and corporate governance: Policy strategies in Norway and New Zealand*. *Gender, Work and Organization*, 18(6), pp 613-630. <https://doi.org/10.1111/j.1468-0432.2010.00514.x>
- Chambers, D., Preston, L., Topakas, A., de Saille, S., Salway, S., Booth, A., Wilsdon, J., (2017). *Review of diversity and inclusion literature and an evaluation of methodologies and metrics relating to health research*. 116.
- Cite Black Women. (n.d.). Cite Black Women. Retrieved from Cite Black Women. website: <https://www.citeblackwomencollective.org/>
- Chronically Academic. (n.d.). Chronically academic. Retrieved from <https://chronicallyacademic.org/index.php/en/#>
- de Madariaga, I. S. (n.d.). *Advancing Gender in Research, Innovation & Sustainable Development*. 72.
- DFG, (2017). *Research-Oriented Standards on Gender Equality*. Retrieved 18 February 2019, from https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/chancengleichheit/forschungsorientierte_gleichstellungsstandards_2017_en.pdf
- Dover, T. L., Major, B., Kaiser, C. R., (2016). *Members of high-status groups are threatened by pro-diversity organizational messages*. *Journal of Experimental Social Psychology*, 62, pp 58-67. <http://dx.doi.org/10.1016/j.jesp.2015.10.006>
- Eby, L. T. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S., Evans, S. C., (2013). *An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protégé perceptions of mentoring*. *Psychological Bulletin*, 139(2), pp 441-476. <https://doi.org/10.1037/a0029279>
- European Commission, (2019). *She Figures 2019*. Luxembourg: Publications Office of the European Union. DOI: 10.2777/936. <https://publications.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en>
- Evans, C., & Glover, J., (2012). *Diversity management change projects: In need of alternative conceptual approaches?* *Journal of Technology Management and Innovation*, 7(3), pp 12-21. <https://doi.org/10.4067/S0718-27242012000300002>
- Farkas, L., (2017) *The Meaning of Racial or Ethnic Origin in EU Law: Between Stereotypes and Identities, European network of legal experts in gender equality and non-discrimination, European Commission 95*. Retrieved from GRC_Publications/Survey_Report_on_Equality_for_GRC_Vitae.pdf
- Milkman, K. L., Akinola, M., Chugh, D., (2015). *What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations*. *Journal of Applied Psychology*, 100(6), pp 1678-1712. <https://doi.org/10.1037/apl0000022>
- Moss-Racusin C. A., Pietri E., Hennes E. P., Dovidio J. F., Brescoll V. L., Roussos G., Handelsman J., (2018). *Reducing STEM gender bias with VIDS (Video interventions for diversity in STEM)*. *Journal of Experimental Psychology: Applied*, 24, pp 236-260.
- Moss-Racusin, C. A., Sanzari, C., Caluori, N., Rabasco, H., (2018). *Gender Bias Produces Gender Gaps in STEM Engagement*. *Sex Roles*, 79(11–12), 651–670. <https://doi.org/10.1007/s11199-018-0902-z>
- Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J., Handelsman, J., (2016). *A "scientific diversity" intervention to reduce gender bias in a sample of life scientists*. *CBE Life Sciences Education*, 15(3). <https://doi.org/10.1187/cbe.15-09-0187>
- Nadler, J. T., Lowery, M. R., Grebinoski, J., Jones, R. G., (2014). *Aversive discrimination in employment interviews: Reducing effects of sexual orientation bias with accountability*. *Psychology of Sexual Orientation and Gender Diversity*, 1(4), pp 480-488. <https://doi.org/10.1037/sgd0000079>
- Nairn, S., (2012). *A critical realist approach to knowledge: implications for evidence-based practice in and beyond nursing: Critical realism and nursing knowledge*. *Nursing Inquiry*, 19(1), pp6-17. <https://doi.org/10.1111/j.1440-1800.2011.00566.x>
- National Academies of Sciences, E., (2018). *Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine*. <https://doi.org/10.17226/24994>

- National Academies of Sciences, E., (2018). *Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine*. <https://doi.org/10.17226/24994>
- Neel R., Shapiro J. R., (2012). *Is racial bias malleable? Whites' lay theories of racial bias predict divergent strategies for interracial interactions*. *Journal of Personality and Social Psychology*, 103, pp 101-120. <https://dx.doi.org/10.1037/a0028237>
- Nishii, L. H., Khatatab, J., Shemla, M., Paluch, R. M., (2018). *A multi-level process model for understanding diversity practice effectiveness*. *Academy of Management Annals*, 12(1), pp 37-82. <https://doi.org/10.5465/annals.2016.0044>
- Nyoni, W. P., He, C., Yusuph, M. L., (2017). *Sustainable Interventions in Enhancing Gender Parity in Senior Leadership Positions in Higher Education in Tanzania*. *Journal of Education and Practice*, 8(13), 44–54. Retrieved from [eric](http://eric.ed.gov/?id=E1111111). (IISTE. No 1 Central, Hong Kong Island, Hong Kong SAR. Tel: +852-39485948; e-mail: JEP@iiste.org; Web site: <http://iiste.org/Journals/index.php/JEP>)
- OECD, (2017). *OECD Science, Technology and Industry Scoreboard 2017 | READ online*. Retrieved from https://read.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-scoreboard-2017_9789264268821-en
- Organisation for Women in Science for the Developing World. (n.d.). *Data Collection and Analysis | OWSD*. Retrieved from <https://owsd.net/resources/data-collection-and-analysis>
- Parent, M. C., Weiser, D. A., McCourt, A. (2015). "So what are you?": *Inappropriate interview questions for psychology doctoral and internship applicants*. *Training and Education in Professional Psychology*, 9(2), pp 136-143. <https://doi.org/10.1037/tep0000068>
- Pawson, R., Greenhalgh, T., Harvey, G., Walshe, K., (2005). *Realist review - A new method of systematic review designed for complex policy interventions*. *Journal of Health Services Research & Policy*, 10, S1:21-S1:34. DOI: 10.1258/1355819054308530.
- Pega, F., Reisner, S. L. ., 3, Sell, R. L. ., Veale, J. F., 6 (2017). *Transgender Health: New Zealand's Innovative Statistical Standard for Gender Identity*. *American Journal of Public Health*, 107(2), 217–221. Retrieved from eapublichealth.org.
- Peretz, H., Levi, A., Fried, Y., (2015). *Organizational diversity programs across cultures: effects on absenteeism, turnover, performance and innovation*. *International Journal of Human Resource Management*, 26(6), pp 875-903. <https://doi.org/10.1080/09585192.2014.991344>
- Phillips, B. N., Deiches, J., Morrison, B., Chan, F., Bezyak, J. L., (2016). *Disability diversity training in the workplace: systematic review and future directions*. *Journal of Occupational Rehabilitation*, 26(3), pp 264-275. <https://doi.org/10.1007/s10926-015-9612-3>
- Pietri, E. S., Hennes, E. P., Dovidio, J. F., Brescoll, V. L., Bailey, A. H., Moss-Racusin, C. A., Handelsman, J., (2018). *Addressing unintended consequences of gender diversity interventions on women's sense of belonging in STEM*. <https://doi.org/10.1007/s11199-018-0952-2>
- Pietri, E. S., Moss-Racusin, C. A., Dovidio, J. F., Guha, D., Roussos, G., Brescoll, V. L., Handelsman, J., (2017). *Using video to increase gender bias literacy toward women in science*. *Psychology of Women Quarterly*, 41(2), pp 175-196. <https://doi.org/10.1177/0361684316674721>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., et al., (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. Swindon: ESRC Methods Programme.
- Prinsloo, P., & Slade, S., (2017). *Big Data, Higher Education and Learning Analytics: Beyond Justice, Towards an Ethics of Care*. In B. Kei Daniel (Ed.), *Big Data and Learning Analytics in Higher Education: Current Theory and Practice* (pp. 109–124). https://doi.org/10.1007/978-3-319-06520-5_8
- Rose, C., (2006). *Do You Have a Disability - Yes or No? Or Is There a Better Way of Asking? Guidance on Disability Disclosure and Respecting Confidentiality*. Retrieved from <https://eric.ed.gov/?id=ED508509>
- Rosenthal, R., (1984). *Meta-Analytic Procedures for Social Research*. Thousand Oaks, CA: SAGE Publications, Inc. <https://dx.doi.org/10.4135/9781412984997>
- Saetermoe, C. L., Chavira, G., Khachikian, C. S., Boyns, D., Cabello, B., (2017). *Critical race theory as a bridge in science training: The California State University, Northridge BUILD PODER program*. *BMC Proceedings*, 11. <https://doi.org/10.1186/s12919-017-0089-2>
- Sang, K., (2017). *Disability and academic careers*. Retrieved from <https://migrantacademics.wordpress.com/2017/05/18/its-like-having-a-second-job-disability-and-academic-careers/>.
- Severina, W. N., Edabu, P., Kimani, C., (2016). *Analysis of influence of sponsorship career function of mentorship on women's leadership advancement in Kenyan universities*. *Journal of Education and Practice*, 7(28), pp 23-27. Retrieved from [eric](http://eric.ed.gov/?id=E1111111). (IISTE. No 1 Central, Hong Kong Island, Hong Kong SAR. Tel: +852-39485948; e-mail: JEP@iiste.org; Web site: <http://iiste.org/Journals/index.php/JEP>)
- Shields, S. A., Zawadzki, M. J., Johnson, R. N., (2011). *The impact of the Workshop Activity for Gender Equity Simulation in the Academy (WAGES–Academic) in demonstrating cumulative effects of gender bias*. *Journal of Diversity in Higher Education*, 4(2), pp 120-129. <https://doi.org/10.1037/a0022953>
- Stanley, N., Ridley, J., Harris, J., Manthorpe, J., (2011). *Disclosing disability in the context of professional regulation: a qualitative UK study*. *Disability & Society*, 26(1), pp 19-32. <https://doi.org/10.1080/09687599.2011.529663>

- Stratton, J., Canales, C., Armas, A., Miller, N., (2006). *Positive stereotyping: Influence tactic for prejudice reduction? Social Influence*, 1, 265–287. <http://dx.doi.org/10.1080/15534510601008882>
- Symonds, M. R. E., Gemmell, N. J., Braisher, T. L., Gorringer, K. L., Elgar, M. A., (2006). *Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance*. *PLOS ONE*, 1(1), e127. <https://doi.org/10.1371/journal.pone.0000127>
- Te Pokai Tara - Universities New Zealand. (2018). Te Kāhui Amokura. *Tauira Māori Initiatives – Sharing good practice in New Zealand universities*. Retrieved from <https://www.universitiesnz.ac.nz/sites/default/files/uni-nz/documents/TKA%20Case%20Studies%20-%20Good%20practice%20%282018%29.pdf>
- United Nations. *Goal 5 :: Sustainable Development Knowledge Platform*. Retrieved March 20, 2019, from Sustainable Development Goals Knowledge Platform website: <https://sustainabledevelopment.un.org/sdg5>
- Wang, M.-T., & Degol, J. L., (2017). *Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions*. *Educational Psychology Review*, 29(1), pp 119-140. <https://doi.org/10.1007/s10648-015-9355-x>
- Webber, K. L., & González Canché, M., (2015). *Not equal for all: Gender and race differences in salary for doctoral degree recipients*. *Research in Higher Education*, 56, pp 645-672.
- Webster, J. R., Adams, G. A., Maranto, C. L., Sawyer, K., Thoroughgood, C., (2018). *Workplace contextual supports for LGBT employees: A review, meta-analysis, and agenda for future research*. *Human Resource Management*, 57(1), pp 193-210. <https://doi.org/10.1002/hrm.21873>
- Williams, C. L., Kilanski, K., Muller, C., (2014). *Corporate diversity programs and gender inequality in the oil and gas industry*. *Work and Occupations*, 41(4), pp 440-476. <https://doi.org/10.1177/0730888414539172>
- Williams, D. R., (2018). *Stress and the Mental Health of Populations of Color: Advancing Our Understanding of Race-related Stressors*. *Journal of Health and Social Behavior*, 59(4), pp 466-485. <https://doi.org/10.1177/0022146518814251>
- Williams, J. B., (2018). *Accountability as a debiasing strategy: Testing the effect of racial diversity in employment committees*. *Iowa Law Review*, 103(4), pp 1593-1638. Retrieved from Scopus.
- Williams, K. W., O'Reilly, C.A., (1998). *Demography and Diversity in Organizations: A Review of 40 Years of Research*. *Research in Organizational Behavior*, 20, pp 77-140.
- Wilton, L. S., Good, J. J., Moss-Racusin, C. A., Sanchez, D. T., (2015). *Communicating more than diversity: The effect of institutional diversity statements on expectations and performance as a function of race and gender*. *Cultural Diversity and Ethnic Minority Psychology*, 21(3), pp 315-325. <https://doi.org/10.1037/a0037883>
- Windscheid, L., Bowes-Sperry, L., Kidder, D. L., Cheung, H. K., Morner, M., Lievens, F., (2016). *Actions speak louder than words: Outsiders' perceptions of diversity mixed messages*. *Journal of Applied Psychology*, 101(9), pp 1329-1341. <https://doi.org/10.1037/apl0000107>
- Woodcock, K., Rohan, M. J., Campbell, L., (2007). *Equitable representation of deaf people in mainstream academia: Why not? Higher Education*, 53(3), pp 359-379. <https://doi.org/10.1007/s10734-005-2428-x>
- Zawadzki, M. J., Danube, C. L., Shields, S. A., (2012). *How to talk about gender inequity in the workplace: Using WAGES as an experiential learning tool to reduce reactance and promote self-efficacy*. *Sex Roles*, 67(11–12), pp 605-616. <https://doi.org/10.1007/s11199-012-0181-z>
- Zliobaite I. Custers B., (2016). *Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models*. *Artificial Intelligence and Law*, 24, 183-201.

Glossary

This report uses several terms that are not in common usage or can possess different meanings in different sectors. For clarity, this report uses the following definitions, except when presenting data from sources where the language or terminology of the original author(s) is used:

- **Affirmative action:** see 'positive action'.
- **Disability, Disabled:** used as an overarching term to describe a range of long-term health conditions, impairments or physical or mental illness which impact on day to day life. Advance HE approaches disability primarily from a social model (that societal structures disables the individuals) but are aware that this approach has its limitations and different understandings (for example, some individuals who are deaf or hearing-impaired will identify as disabled, but others will not).
- **EDI:** an acronym for equality, diversity and inclusion. These concepts have different meanings and interpretations, often encompassing issues of representation, equity of access and opportunity, and active steps to ensure a sense of belonging.
- **Ethnicity:** see 'race'.
- **Gender, sex:** although the two words are often used interchangeably (and may not be different words in some languages), we understand these to have different meanings (with gender as a social rather than biological construct). Generally, Advance HE believes the word 'gender' is more inclusive than 'sex' for acknowledging a range of identities and experiences and is therefore used where appropriate throughout this review.
- **Identity characteristics:** a term used in this review to discuss different types of group or individual identities or backgrounds commonly considered in EDI. Its relation to 'protected characteristics' is discussed below. For the international review we have broadly conceived these as in the appendices; it is noted that cultural, legal and theoretical understandings of these groupings will vary widely in the global context, so any list used will have its limitations.
- **Innovation:** the creation of new products, services, and ways of doing business, the definition used by the WISE campaign for gender balance in science, technology and engineering.
- **Intervention:** used to refer to any new or changed activity (programme, training, policy, practice or way of working) with the aim of reducing differential access, experiences, progression or outcomes for those working or studying in or around the R&I sector. One source might include multiple examples of interventions.
- **Intersectional, intersectionality:** developed by Professor Kimberlé Crenshaw, a theory or approach that acknowledges the specific and compounding effects of oppression relating to multiple identities. Originally conceived as a 'lens' to analyse the effect of structural sexism and racism on the lives of black women, the approach has (not without challenge) been used to examine a range of different experiences.
- **Language minority:** referring to groups or individuals whose primary language is different to that used by the majority the majority of people in their organisation or environment (or the official language where this is different).
- **People of colour:** see 'race'.
- **Positive, targeted or affirmative action:** refers to a range of concepts where actions are targeted at a beneficiary group, aimed at ending discrimination against that group by redressing the effects of past discrimination. Legal definitions of these actions will vary (see discussion in section 5.3.3).
- **Protected characteristics:** identity or group characteristics which have specific legal protections against discrimination. There are nine identity characteristics covered under the UK's 2010 Equality Act. However, legal protections (and definitions) for different identities, backgrounds or needs will vary around the world and in different contexts (for example, in employment or services).
- **'Race':** where used this is primarily through its UK legal lens of referring to ethnicity, skin colour, ethnic or national origins, or nationality (including citizenship), although these factors may be referenced individually where appropriate. Advance HE approaches 'race' equality from the position that 'race' is a social construct and therefore has associated limitations and complex changing understandings. In discussing specific sources, wherever possible the source's chosen terminology will be reflected. A variety of terminology was used in the literature such as 'people or women of colour' as well as specific ethnic/national and racial group identified is also used; terminology used in specific reports has been mirrored where possible.
- **Source:** any document that provides information on EDI interventions and/or challenges. This might include a paper in an academic journal, an organisation's report, or a response to the project's Call for Evidence.
- **STEM(M):** an acronym for science, technology,

engineering, mathematics (and medicine).

- **UKRI:** United Kingdom Research and Innovation, which includes seven research councils, Research England and Innovate UK.
- **Women of colour:** see 'race'.

Appendix

The background of the page is an abstract geometric design. It features a repeating pattern of hexagons in various shades of blue and teal, creating a sense of depth and perspective. A large, solid yellow triangle is positioned on the right side, pointing towards the bottom right corner. The overall composition is modern and architectural.

A. Database search terms, restrictions and limitations

A.1 Search terms

Based on the feedback from the UK and international review Advisory Group, the research team used the following Boolean search terms to identify existing literature sources:

Equality OR equity OR diversity OR inclusi*
OR underrepresentation OR wellbeing OR
discriminat* OR prejudic* OR bias) AND (Age OR
Disab OR Sex OR gender OR men OR women OR
male OR female OR "gender reassignment" OR
trans OR Marriage OR civil partner OR Religio*
OR belief OR Pregnant* OR maternity OR Sexual
orientation OR LGB* OR sexual OR Race OR racial
OR Ethnic* OR socio-economic)

and

(project OR intervention OR initiative OR action
OR initiative OR programme OR policy OR "good
practice" OR "best practice" OR process* OR
plan* OR "action plan" OR strateg* OR monitoring
OR evaluation OR quota OR affirmative action
OR positive action OR mainstream* OR embed*
OR ethos OR mission OR "strategic plan" OR
"corporate plan" OR "impact assessment" OR
changing attitude* OR value* OR "professional
development" OR workshop OR disclosure OR
feedback OR career* OR recruit* OR promot*
OR employ* OR pay OR training OR engagement
OR represent* OR "role model" OR retain OR
retention OR progress* OR perform* OR develop*
OR training OR CPD OR mentor* OR leader* OR
talent OR pipeline OR rewards OR sponsor)

and

("research and innovation" OR knowledge OR
"research funding" OR "public engagement" OR
"research career" OR organisation OR innovat*
OR partnership OR "awareness raising" OR
resource* OR tool* OR campaign* OR "higher
education" OR *doctora*)

A.2 Restrictions

The following restrictions were added to the searches:

- terms must be contained in the title or abstract of the publication
- publication date must be between January 2011 and February 2019 (to capture the most recent research and to provide comparability with the UK review)
- publication must be available in English (originally or in translation)
- full-text of the publication must be accessible
- publication types or sources only include academic journals, journals, reports, trade publications, overviews, conference materials, books, government documents and reviews.

A.3 Limitations of data collection

Evidence-based recommendations

It became clear that many sources presented **evidence-based recommendations** rather than an **evaluation of an intervention**. For example, one source presented data from focus groups conducted with doctoral students who had experienced or were experiencing mental health issues. The source then identified work taking place within HEIs that focus group participants had described as being effective in helping with managing their mental health issues. In relation to this review, we included these types of sources as they might inform UKRI's future work. For an additional discussion of how an intervention was operationally defined within the current review, see appendix E.4.

Identity characteristics

The search terms returned many sources that referred to an identity characteristic but did not directly relate to EDI. This was particularly the case for 'age' and 'disability', which returned a large number of sources that related to healthcare and were not relevant to UKRI's work (for example, interventions in paediatric care).

Transferability to UKRI's work

The manual review of sources required researchers to assess whether work could relate to any area of UKRI's work. We were open to sources that documented interventions from outside the R&I sector which would still be relevant to UKRI (for example, as a large employer) but excluded sources that were not transferable to UKRI's work. This was particularly common with sources that discussed healthcare interventions.

Limitations of databases

It became apparent during the review of sources that a large number related to research in the fields of healthcare and primary-level and secondary-level education were from psychology or related

disciplines, and were conducted by researchers based in the US. These reflections on the potential limitations of EBSCO, OpenGrey and Scopus databases informed the other three strands of data collection.

English-only publications

Unfortunately given the timeframe and resources of the current review, we were only able to include publications available in English. This may have limited the number of sources from international journals, or within the grey literature search, from organisations that publish in other languages (Welsh, Chinese, Japanese etc.).

B. Organisational sources included from the targeted grey literature search or responding to the Call for Evidence

Some materials relating to organisational EDI activities appeared in multiple collection strands. For example, the desktop grey literature search and the Call for Evidence at times overlapped or complemented each other. These have been summarised together below. Additionally, some organisational activity was also present within academic literature (for example, evaluations of programmes which had been partially financed or supported by the US NSF).

- Canadian Institutes for Health Research (CIHR)
- Centre National de la Recherche Scientifique (CNRS) (France)
- DFG (German Research Foundation)
- Forbes
- FWF Strategy for Gender Equality and Diversity of Researchers (2019-2020) (Austria)
- GENDER-NET (Europe)
- GENOVATE (Europe)
- Global Research Council (GRC)
- International Human Rights Network of Academies and Scholarly Societies
- KIF Committee for Gender Balance and Diversity in Research (Norway)
- KIRAN Women Scientists Programme (India)
- Max Planck Institute (Germany)
- McKinsey
- OECD
- Research Promotion Foundation (Cyprus)
- Science Foundation Ireland (SFI) (Republic of Ireland)
- Science and Engineering Research Council (India)
- Science in Australia Gender Equity (SAGE)
- Swiss National Science Foundation (SNF)

- Swiss National Centres of Competence in Research (NCCRs) MUST (Molecular Ultrafast Science and Technology)

- Universities New Zealand

A wider range of organisational sources was reviewed but may have been excluded from the framework due to unavailability in English or in translation, or insufficient content to enable evaluation or analysis. Some organisational activity was also present within academic literature (for example, activity by the US National Science Foundation).

C. Call for Evidence

C.1 Data collection

The Call for Evidence was intended as a supplementary data collection method to help surface unpublished documents and encourage self-reporting, particularly of 'what doesn't work' or key learning from attempts at implementation of initiatives (practical, financial), and to provide opportunities for contextual reflection (policy drivers, scalability). It was not intended to provide 'representative' data of the extent or range of work in the sector. A larger systematic call with a wider timeframe could be a useful recommendation for future work.

Who was targeted?

The Call for Evidence was circulated using a snowballing method that harnessed:

- Advance HE and Advisory Group international contacts
- relevant Jisc mailing lists (such as riag@jiscmail.ac.uk)
- contacts identified from organisations within the grey literature search
- opportunities to advertise on the UKRI website
- UKRI international offices.

Those who received the Call for Evidence were encouraged to share it with others working in the sector to maximise its reach. On the advice of the Advisory Group, the email invitation that accompanied the call emphasised the benefits of the EDI review for institutions and the wider sector.

How was evidence collected?

Response form and methods to respond

Respondents were provided with a range of methods to respond, either attaching existing literature to an email (or providing URLs) or responding to structured questions via an online or electronic form, or telephone or Skype.

Questions were designed to be flexible so that respondents could share different types of intervention, and to be not too onerous to maximise participation and prompt institutions to return information that would help answer the project's research questions.

Accessibility requests (for example, for alternative formats) were welcomed and utilised.

Timeline

The timeline for the call was necessarily tight (around two weeks) due to the short duration of

this review and the allocation of time to collate and evaluate the total body of evidence. It was acknowledged that this short timeframe may have biased the sample towards organisations which were more familiar with our organisation or UKRI, and/or sufficiently resourced to respond quickly.

C.2 Call for Evidence form

Overview of the change or intervention

What? What was the challenge or problem being addressed? (For example, underrepresentation of a specific group of people, or improving experiences of minority groups)	
Why? Why did this work happen? What was the theory, rationale or driver behind the intervention? (For example, legal requirement, business needs)	
Who? Who was responsible for the intervention and/or provided resources or input?	
How? How often? How much (cost, scale)? (For example, describe any approaches or resources used for training, monitoring, changes to organisation or estate)	
Where? Location (geographical location of the intervention) and location within the organisation (eg Human Resources, senior leaders)	
Pilots, modifications and tailoring How often? Any adjustments, adaptations for over time or in different situations (planned or undertaken)?	

Additional information

Details of further information if published online (if applicable)	
--	--

Impact and measurement

Describe the time frame for the intervention (beginning and end)	
Was the effectiveness of these interventions measured?	
How was the effectiveness measured? (Tick all that apply)	☐ Survey/questionnaires ☐ Focus group/interviews ☐ Monitoring data ☐ Case studies ☐ Other quantitative method: ☐ Other qualitative method: ☐ Other:
Did you consider this measurement method effective?	Describe briefly:
Were there any unexpected or additional outcomes (positive or negative) beyond the original aim(s)?	
What element(s) of the intervention proved particularly useful or innovative?	
If the intervention successfully achieved its stated aims (fully or partially), why do you think this was?	
If the intervention was unsuccessful (fully or partially) why do you think this was?	
If you were to repeat the intervention, would you do anything differently?	
Any comments on whether this intervention could be scaled up or applied in a different context?	

About your organisation

Name of organisation	
Do you wish your organisation to remain anonymous in any final report?	☐ Yes: please provide a general descriptor for example "an international research funder" ☐ No
What best describes your organisation? (tick all that apply)	☐ Research institute (publicly or government funded) ☐ Research institute (private/for-profit) ☐ Higher Education institution/University ☐ Government or State body ☐ Research funding organisation ☐ Non-Governmental Organisation ☐ Charity ☐ Think tank ☐ Business ☐ Other (please describe or provide link):
Link to description of your organisation (optional)	

Which nation/region does your organisation operate in? (tick all that apply)	☐ UK ☐ England ☐ Northern Ireland ☐ Scotland
	☐ Wales ☐ Africa [specify] ☐ Americas [specify] ☐ Asia Pacific [specify] ☐ Europe [specify] ☐ Middle East/North Africa [specify] ☐ Other:

Impact and measurement

Name	
Role or Title	

D. Working understanding of identity characteristics within the international context

The following was used internally when reviewing inclusion criteria and also communicated with the Call for Evidence guidance with explanatory text.

Characteristic	Alternative terms which may be encountered or related themes (indicative)
Age	Young, early career, mid-career, retirement, old, elderly, senior, ‘second chance’, mature
Disability	Disabled, impairment, accessibility, enabling, mental health, illness, long-term illness, physical estates, neurodiversity, learning difference
Sex, gender, gender identity	men, women, male, intersex, female, girls, boys, feminism, misogyny, patriarchy
Trans or non-binary identity or history	Transgender, non-binary, intersex, gender queer, gender diversity, gender reassignment or affirmation, transphobia
Marital status	Marriage, civil partnership, status, union, marital status, married, same sex partner*, spouse, husband, wife
Religious or philosophical belief or heritage	Faith, religious practice or observance, secularism, atheism, humanism, sectarianism, spirituality, religious harassment (for example, islamophobia, antisemitism)
Pregnancy, breastfeeding, childcare	Maternity, paternity, working mother or father, parental leave or pay, adoption, parent, childcare, childbearing, returner
Race, ethnicity, nationality	Heritage, citizenship, white, BME, BAME, majority ethnicity, minority ethnicity, person of colo[u]r, POC, national identity, indigenous background, traveller background, first nations, background, migrant or migration background, caste, Global North, Global South
Sexual orientation/LGB+	Lesbian, gay, bisexual, queer, asexual, same-sex relationship, homosexual, heterosexual, sexual diversity, homophobia, biphobia
Socio-economic status	Low-income, parental occupation, ‘class’ background, social mobility poverty, economic background
Language	Minority languages, official language, speakers of other languages, multilingual, dialect, sign language, communication

E. Data refinement

E.1 Inclusion and exclusion criteria

To limit the scope of the review and most effectively answer the project's research questions, the research team sought input from the Advisory Group and established inclusion and exclusion criteria (see table 2.2). These criteria were applied to the academic and grey literature database search, the targeted grey literature search and responses to the Call for Evidence.

Where appropriate, sources which were ultimately excluded but which could inform future analysis or interpretation (for example, providing a national context, or discussing a theoretical approach) were retained and used to inform final analysis.

Table E1.

Included	Excluded	Rationale (summary)
Published on or after 1 January 2011.	Published before 1 January 2011.	Seeking the most current thinking and interventions. Significant earlier work likely to be cited. Parity with scope of the UK review.
Reasonable degree of certainty of validity (for example, from an 'organisational' email address or URL) and with permissions from the relevant organisation.	Sources from personal email addresses, blogs, journalism, 'exposés', legal cases or reports (such as from employment tribunals) or for-profit consultancies or trainers.	Confidence in source origins in the short timeframe. Focusing on evidence-based reporting as priority, with room for organisational self-reporting.
Discusses at least one identity characteristic from a list analogous to the UK protected characteristics, or socio-economic status, or 'diversity' or 'equality' in general.	Does not discuss a characteristic from the list.	The UK Equality Act 2010 range of characteristics is broad so provided a useful starting point. However, strict equivalency may be unsuitable; this list ensures a wide range of characteristics. Example: 'caste', 'indigenous' identities, minority languages considered in scope. Mindful of comparability and relevance to the UK context.
Includes some measurement of outcomes and is: An empirical evaluation of an EDI intervention. A review, meta-analysis or gap analysis of EDI interventions.	Descriptive sources that: Evidence the existence or experience of EDI challenges without reference to interventions or actions taken in response. Are first-person experiential accounts. Provide information about interventions without clear outcomes.	Selected because the primary research questions for review focus on understanding what activities are taking place, how these are measured and what is proving successful and less successful.
Took place within an organisation involved in research or innovation or translatable to the funding, practice or communication, or employment of R&I.	Took place in a context not relatable to R&I.	Relevant to scope of the research.
Any personal data is anonymised, aggregated or given with consent.	Contains personal data without consent to share or appropriate levels of data protection.	Data protection.

Available in English.	Not available in English.	Due to time and resource constraints of the project. Possible mitigations (including using summary reports of non-English materials) discussed.
Discusses interventions conducted outside of the UK (or inside and outside the UK).	Discusses interventions conducted only within the UK.	Scope of the review.

E.2 Final sample

The table below summarises the total number of eligible sources and interventions across the three strands of data collection.

Strand	No. of eligible sources	No. of interventions
Academic and grey literature database search	81	98
Targeted grey literature search	13	15
Call for Evidence	14	17

E.3 Reliability

To ensure that the eligibility criteria had been applied in a similar manner across the four main researchers on the current team, a subsample of 10% of all identified sources was double-coded by a fifth researcher who was blind to which sources had been labelled as eligible by the research team.

To estimate the reliability of the individual eligibility criteria, we compared the proportion of sources that were included or excluded by the research team to those included or excluded by the fifth researcher (summarised in the table below).

Criterion	Research team	Fifth researcher
Duplicate	11.1	11.2
Eligible - international	9.5	9.3
Eligible - UK	1.4	1.9
No access	9.0	10.0
No PC/empirical combined	64.1	63.2
Total	100.0	100.0

E.4 What is an intervention?

In order to apply inclusion and exclusion criteria, the research team had to agree on a common definition of the term 'intervention'. As noted, sources that focused on theoretical or conceptual approaches, discussions or persuasive essays were excluded.

Many sources adopted a grounded approach that explored an EDI-related phenomenon (for example, the factors that make black women more likely to join a mentorship scheme) and presented

possible reasons (length of time working for the organisation, support from line manager etc.). As these examples did not evaluate a specific intervention introduced to address a challenge, they were excluded.

In other sources, EDI-related phenomena were discussed but the outcome variable was not a protected characteristic. As an example, compare these two sources:

Excluded: the impact of an ethnically diverse senior leadership team on an organisation's overall productivity.

In the above example, the excluded source includes an independent variable related to EDI (that is, an ethnically diverse senior leadership team), but the dependent variable was not related to a protected characteristic or EDI (that is, it looked at productivity instead of improving the representation or career development of staff from ethnically diverse backgrounds). Both studies are related and important, but the focus of our review sought to examine the latter.

F. Evaluation framework

F.1 Design

The research team designed a framework that was flexible enough to evaluate different types of source but universal enough so that subsequent analysis was meaningful and able to tell a coherent story. As far as possible, discrete response options were presented to improve the quality of quantitative analysis. The framework also had to capture information about sources and single or multiple interventions contained within each source. The framework allowed a maximum of five discrete interventions to be shared per source.

To facilitate the gathering of evaluation data, the framework was hosted on Survey Monkey. This enabled researchers across the team to simultaneously input data.

The framework required researchers to describe the intentions of interventions, the challenges they intended to address and their relevance to UKRI's work, and to assess the robustness of evaluation methods and the successes or failures of interventions. The framework provided space to input data on the level of confidence that the intervention was responsible for the stated outcomes, for example using the Maryland Scientific Method Scale, as well as an intervention's reach (number of people, areas of work) and the extent of its impact (individual or institutional change). The framework's flexibility also presented opportunities to report on interventions that lacked a rigorous evidence base of impact but suggested exciting potential, as well as interventions that had limited or unexpected outcomes. All sections of the framework allowed for free-text responses to ensure no meaningful information was lost during the evaluation process.

The framework underwent testing, which involved the evaluation of two sources (one academic, one grey) to help identify questions that were missing from the framework, areas of overlap, questions that did not make sense and the refinement of response options. Results from this testing led to the revision of some framework questions, including adding a question to clarify whether the source evaluated an intervention, or was a review or meta-analysis, or if the source presented evidence for examples of EDI best practice without necessarily including primary data collection and/or analysis. For example, if a source described an EDI policy or initiative alongside unpublished evidence (for example, an organisation built a new university programme for recruiting women into software engineering that doubled the number of female software engineer interns), we opted to expand the evaluation framework and include the source in subsequent analysis.

F.2 Evaluation framework

Source title:	
Geographic focus:	☐ UK ☐ England ☐ Northern Ireland ☐ Wales ☐ Scotland ☐ International (please specify countries/regions)
Data collection stream:	☐ Academic (eg peer-reviewed journal article) ☐ Grey (eg organisation report) ☐ Call for Evidence response ☐ Athena SWAN application (UK only) ☐ Other (please specify):

Characteristics covered:	☐ Gender/sex ☐ Disability (including mental health) ☐ Trans identity (gender reassignment) ☐ Marriage and civil partnership ☐ Pregnancy and maternity ☐ Race (ethnicity or nationality) ☐ Religion and belief ☐ Age ☐ Sexual orientation ☐ Unspecified/general EDI ☐ Other (please specify):
Does the source refer to socioeconomic status?	☐ No ☐ Unclear ☐ Yes (please specify measure eg. income, post code, parent education etc.)
Does the source explicitly apply an intersectional lens?	☐ No ☐ Unclear ☐ Yes
What area(s) of EDI work does the source focus on?	☐ Careers (recruitment, promotion, leave policies etc.) ☐ Culture and wellbeing (inclusion, experiences etc.) ☐ Outreach and public engagement (community work, events etc.) ☐ Data (equality monitoring, increasing disclosure etc.) ☐ Funding (scholarships, grant awards etc.) ☐ Other (please specify):
How would you describe the source?	☐ Evaluation of intervention ☐ Evidence-based recommendations (ie focus group findings) or contextual information to better understand EDI interventions (ie research that could inform future interventions) ☐ Review of multiple interventions ☐ None of the above/source should be excluded
Where was the research/intervention(s) developed/designed?	☐ Higher education institution ☐ Research institute ☐ State ministry or government agency ☐ Non-governmental organisation ☐ Commercial entity ☐ Learned society ☐ Research funding organisation ☐ Unsure ☐ Other (please specify):

Where was the research/intervention(s) intended to impact?	☐ Higher education institution ☐ Research institute ☐ State ministry or government agency ☐ Non-governmental organisation ☐ Commercial entity ☐ Learned society ☐ Research funding organisation ☐ Unsure ☐ Other (please specify):
Did the research/intervention(s) involve a partnership of multiple organisations?	☐ No ☐ Unsure ☐ Yes (please specify):
What is the sector/discipline focus of the source?	☐ HE/research/STEM ☐ Business/management/leadership ☐ Education/teaching/learning ☐ Healthcare ☐ Creative arts ☐ Charity/community/public ☐ Other (please specify):
What area(s) of UKRI's work could the source relate to?	☐ Research funder ☐ Employer ☐ Research and innovation policy ☐ Public engagement/outreach ☐ None ☐ Other (please specify):
<i>This part of the evaluation asks about individual interventions. If the source includes multiple interventions (for example, a meta analysis or review), you will be invited to complete this page up to a total of four times. If the source features more than four, please provide information on interventions with the most available data.</i>	
Date/date range when research/intervention was undertaken (leave blank if unknown):	
Size of organisation where research/intervention was undertaken?	☐ Small (under 50 people) ☐ Medium (50-250 people) ☐ Large (over 250 people) ☐ Unsure ☐ Other (please specify):

What type of intervention is discussed? Or what is the focus of the research?	☐ Training/development ☐ Mentoring/coaching ☐ Strategy/policy change ☐ Awareness raising ☐ Organisational review/assessment of EDI ☐ Learning resources/tools ☐ Outreach ☐ Unsure ☐ Other (please specify):
Briefly describe the research/intervention:	
What did the intervention intend to change? If research, how could this inform future interventions?	
What type of methodology was used to evaluate the intervention?	☐ Within-groups design ☐ Between-groups design ☐ Time series analysis ☐ Cross-sectional analysis ☐ Case study/ies ☐ Qualitative analysis of interviews ☐ Qualitative analysis of focus groups ☐ Ethnography/observation ☐ Conceptual/not based upon empirical evidence ☐ Unsure ☐ None ☐ Other (please specify):
Briefly describe the research/intervention:	
What did the intervention intend to change? If research, how could this inform future interventions?	
What type of methodology was used to evaluate the intervention?	☐ Within-groups design ☐ Between-groups design ☐ Time series analysis ☐ Cross-sectional analysis ☐ Case study/ies ☐ Qualitative analysis of interviews ☐ Qualitative analysis of focus groups ☐ Ethnography/observation ☐ Conceptual/not based upon empirical evidence ☐ Unsure ☐ None ☐ Other (please specify):
Information on the intervention (if known):	

Target sample:	
Sample size:	
Control variables:	
Number of people involved in design/delivery:	
Location of intervention within organisation (eg senior leadership, HR):	
Financial cost of intervention:	
How are the EDI challenges the research/ intervention intended to address understood?	☐ individual (eg confidence building, individual adjustments) ☐ Structural (eg quotas) ☐ Both ☐ Unsure ☐ Other (please specify):
What data was captured from the intervention?	☐ Quantitative ☐ Qualitative ☐ Unsure ☐ Other (please specify):
If possible, assess the intervention using the Maryland Scientific Method Scale:	☐ Level 1: Correlation (eg departments with a female leader have more female staff) ☐ Level 2: Before and after assessment, with no control of conditions (eg female staff in a department increased after the appointment of a female leader) ☐ Level 3: Before and after assessment, with experimental conditions (eg female staff in a department increased after the appointment of a female leader, female staff in a department did not increase after the appointment of a male leader) ☐ Level 4: Before and after assessment, with multiple experimental conditions (eg as with level 3 but with additional controls for gender culture in department and individuals backgrounds of staff) ☐ Level 5: Randomised control trial ☐ Unsure/Not applicable
Please use this space to provide further information on any assessment of 'robustness':	
Were outcomes of the intervention measured/ evaluated and the results reported?	☐ Outcomes measured/evaluated but not reported ☐ Outcomes measured/evaluated and reported ☐ Outcomes neither measured/evaluated nor reported
If measured, please note the method(s) used:	☐ Self-reported ☐ Impact evaluation ☐ Other (please specify):

If reported, please provide information on outcomes:	
<i>This final page asks you to again think of the source holistically, rather than individual interventions noted within the source. This information might be found in a concluding section on reflections or recommendations.</i>	
Does the source present reasons for success?	☐ No ☐ Yes (please provide information):
Does the source present reasons for failure?	☐ No ☐ Yes (please provide information):
Does the source present recommendations or suggestions for future work?	☐ No ☐ Yes (please provide information):
Does the source report EDI challenges that lack current interventions?	☐ No ☐ Unsure ☐ Yes (please list challenges identified):
Any other comments:	

F.3 Reliability

Evaluation framework variable	Percent agreement (%)
Source (for example, academic paper, grey literature, Call for Evidence)	94.1
Protected characteristic(s) addressed*	95.9
Application of intersectional lens	82.4
Area of EDI (careers, culture and wellbeing, outreach and public engagement, etc.)*	87.1
Location of intervention development (HEI, commercial entity, government organisation, etc.)	91.7
Location of intervention impact (HEI, commercial entity, government organisation, etc.)	83.3
Sector or discipline (HE, STEMM, business, education, arts, etc.)	83.3
Relevance to UKRI (e.g. research funder, employer, R&I policy, public engagement, etc.)	81.3
Intervention type (training and development, mentoring or coaching, strategy or policy change, etc.)*	81.0
Type of method employed (between groups, within groups, interviews, focus groups, etc.)*	88.2
Type of data (quantitative, qualitative, mixed, etc.)	75.0
Maryland Scientific Method Scale	50.0

* indicates variables with multiple categories, and as such the percent agreement presented is an average.

Acknowledgments

The authors thank all those who have supported this review with feedback, questions and constructive conversations including: our Review Advisory Group (below); UKRI staff (UK and international offices); the UKRI EDI External Advisory Group; and respondents to our Call for Evidence, which included:

- Canadian Institutes of Health Research (CIHR)
- Committee for Gender Balance and Diversity in Research, Norway (KIF)
- German Research Foundation (DFG)
- Max Planck Society

- National Centre of Competence in Research Molecular Ultrafast Science & Technology (NCCR-MUST)
- Research Promotion Foundation (Cyprus)
- Science Foundation Ireland (SFI)
- Science in Australia Gender Equity (SAGE)
- Swiss National Science Foundation (NSF)

Additional thanks to sector colleagues unable to formally join the Advisory Group but who gave their time to discuss contextual issues and approaches, provided useful materials and/or supported the dissemination of our Call for Evidence.

Advisory Group members

Name	Organisation
Shaun Holmes	British Council
Louis Stupple-Harris	British Science Association
Rachel Handforth	Careers Research and Advisory Centre, Vitae
Gregory Crouch	Equality and Human Rights Commission
Lindsey Crosswell	European Bioinformatics Institute
Matthew Guest	GuildHE
Rochelle Fritch	Science Foundation Ireland
Anna Bradshaw	The British Academy
Karen Salt	UKRI Equality Advisory Group, University of Nottingham
Rachael Gooberman-Hill	University of Bristol
Sheila Riddell	University of Edinburgh

The international review was undertaken by Jess Moody and Dr Amanda Aldercotte at Advance HE in close collaboration with the UK review, led by Dr Kevin Guyan with Freya Douglas Oloyede. Both reviews received feedback and guidance from Dr

Pauline Hanesworth, Dr Joan O'Mahony and Gary Loke at Advance HE and from associate Ashlee Christofferson. Alice Greenslade provided project support.

About Advance HE

Advance HE was formed in March 2018 from a merger of the Equality Challenge Unit (ECU), the Leadership Foundation for Higher Education and the Higher Education Academy. We have over ten years' experience supporting institutions and research institutes to remove barriers to progression and success for all staff and students. We provide a central source of expertise, advice, research and leadership on equality and diversity that drives forward change and transforms organisational culture in teaching, learning, research and knowledge exchange. We are:

- A specialist body: Advance HE has substantive practical experience, expertise, and insight with relation to equality and diversity and underrepresentation pertaining to staff and students at every level and in every function of the HE and research sector.
- A focus on identifying, sharing and evidencing impactful practices: Identifying and recognising more systemic solutions to barriers to EDI is the focus of our gender and race charters work for HE institutions and research institutes. Our work in

understanding, identifying and embedding impactful practice is informed by our overarching knowledge of activity in the sector and our understanding of the latest innovative interventions through discrete projects and relationships.

- Not for profit: As a registered educational charity, all funds are directly reinvested back into strengthening Advance HE's mission for the benefit of stakeholders.

Advance HE is a company limited by guarantee registered in England and Wales no. 04931031. Registered as a charity in England and Wales no. 1101607. Registered as a charity in Scotland no. SC043946. Advance HE words and logo should not be used without our permission. VAT registered no. GB 152 1219 50.

Contact us

+44 03300 416201
enquiries@advance-he.ac.uk
www.advance-he.ac.uk
@AdvanceHE

About UK Research and Innovation

Big challenges demand big thinkers- those who can unlock the answers and further our understanding of the important issues of our time. Our work encompasses everything from the physical, biological and social sciences, to innovation, engineering, medicine, the environment and the cultural impact of the arts and humanities. In all of these areas, our role is to bring together the people who can innovate and change the world for the better.

We work with the government to invest over £7 billion a year in research and innovation by partnering with academia and industry to make the impossible, possible.

Through the UK's nine leading academic and industrial funding councils, we create knowledge with impact.

Promoting equality, diversity and inclusion is at the heart of UK Research and Innovation's (UKRI's) vision.

We believe that equality, diversity and inclusion – of people and ideas – is integral to excellence in research and innovation, letting us access the best talent and nurture great ideas. We therefore embed equality, diversity and inclusion at all levels and in all that we do, both as an organisation and as a funder.

UKRI works to ensure that every employee is treated with dignity and respect. We will not accept bullying and harassment in any form and we are developing an action and engagement plan to ensure the UK research environment is safe, supportive and effective.

Contact us

01793 444000
equality@ukri.org
www.ukri.org
@UKRI_News

Registered office
UK Research and Innovation
Polaris House
Swindon
SN2 1FL

London office
UK Research and Innovation
58 Victoria Embankment
London
EC4Y 0DS

Bristol office
UK Research and Innovation
Research England
Nicholson House
Lime Kiln Close
Stoke Gifford
Bristol
BS34 8SR



Accessibility

To request copies of this report in large print or in a different format, please contact the Equality, Diversity and Inclusion team at UKRI:

01793 444000

equality@ukri.org

UK Research and Innovation

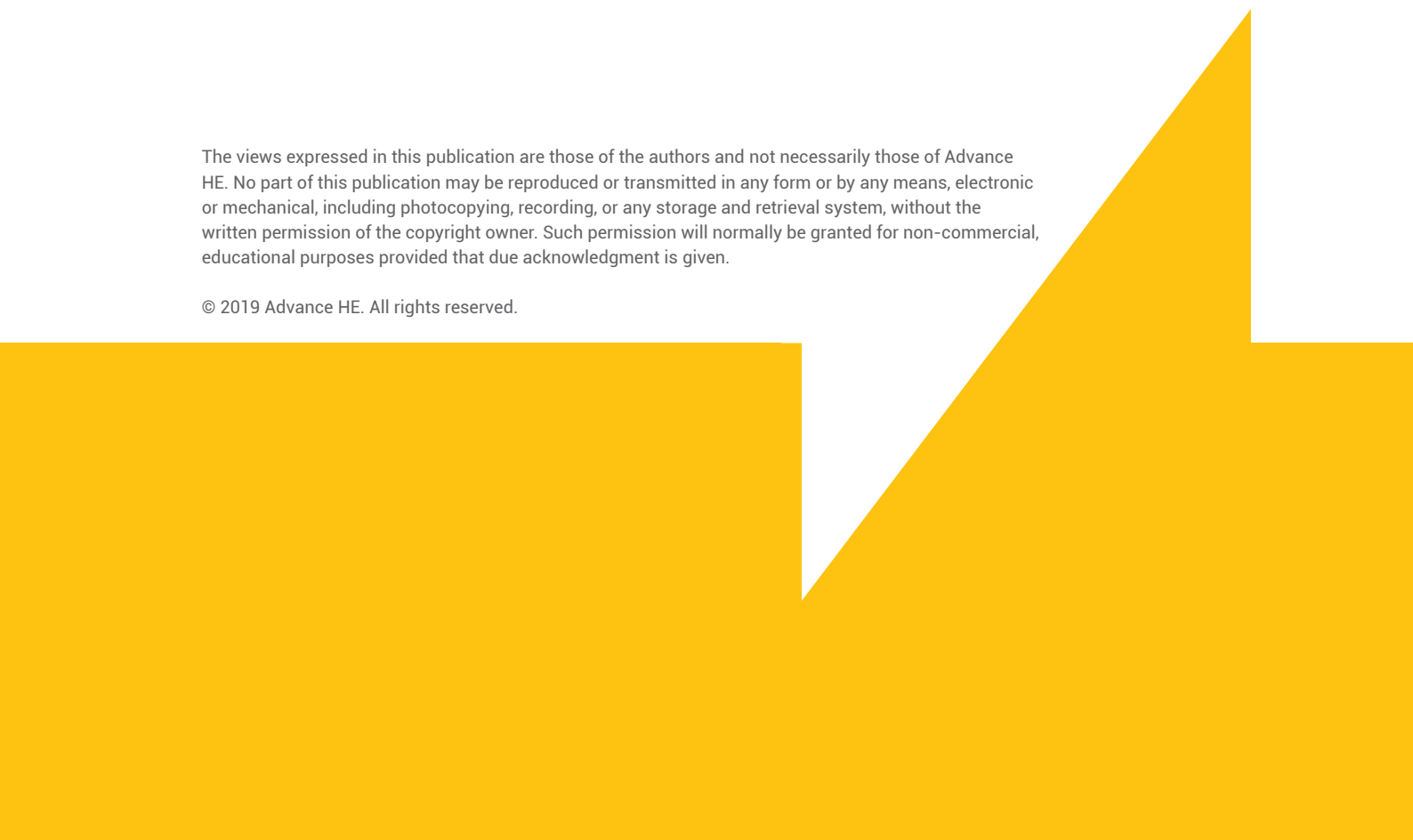
Polaris House

Swindon

SN2 1FL

The views expressed in this publication are those of the authors and not necessarily those of Advance HE. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any storage and retrieval system, without the written permission of the copyright owner. Such permission will normally be granted for non-commercial, educational purposes provided that due acknowledgment is given.

© 2019 Advance HE. All rights reserved.

A large yellow geometric shape, consisting of a triangle and a rectangle, is positioned in the bottom right corner of the page. Another yellow rectangular shape is located at the bottom left, partially overlapping the white space.