ETHICAL EVIDENCE AND POLICYMAKING

Interdisciplinary and International Research

EDITED BY RON IPHOFEN
AND DÓNAL O' MATHÚNA

PRO-RES

ETHICAL EVIDENCE AND POLICYMAKING

Interdisciplinary and International Research

Edited by
Ron Iphofen and Dónal O'Mathúna

PRO-RES

First published in Great Britain in 2022 by

Policy Press, an imprint of
Bristol University Press
University of Bristol
1–9 Old Park Hill
Bristol
BS2 8BB
UK
t: +44 (0)117 374 6645
e: bup-info@bristol.ac.uk

Details of international sales and distribution partners are available at
policy.bristoluniversitypress.co.uk

Editorial selection and matter © the editors; individual chapters © their respective authors

The digital PDF and EPUB versions of this title are available Open Access and distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits reproduction and distribution for non-commercial use without further permission provided the original work is attributed.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 978-1-4473-6395-8 hardcover
ISBN 978-1-4473-6396-5 OA EPUB
ISBN 978-1-4473-6397-2 OA PDF

The right of Ron Iphofen and Dónal O'Mathúna to be identified as editors of this work has been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved: no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without the prior permission of Bristol University Press.

Every reasonable effort has been made to obtain permission to reproduce copyrighted material. If, however, anyone knows of an oversight, please contact the publisher.

The statements and opinions contained within this publication are solely those of the editors and contributors and not of the University of Bristol or Bristol University Press. The University of Bristol and Bristol University Press disclaim responsibility for any injury to persons or property resulting from any material published in this publication.

Bristol University Press and Policy Press work to counter discrimination on grounds of gender, race, disability, age and sexuality.

Cover design: Nicky Borowiec
Front cover image: iStock/Anastasiia Guseva
Bristol University Press and Policy Press use environmentally responsible print partners.
Printed and bound in Great Britain by CPI Group (UK) Ltd, Croydon, CR0 4YY

FSC
www.fsc.org
MIX
Paper from responsible sources
FSC® C013604

Contents

PART III Case studies

List of figures, tables and boxes

Figures

Tables

Boxes

Notes on contributors

Nchafack Almighty is a communication specialist and multidisciplinary public health PhD candidate on the Scaling-up Packages of Interventions for Cardiovascular disease prevention in selected sites in Europe and sub-Saharan Africa (SPICES) project at Nottingham Trent University, UK. Her doctoral research explores the understandings of heart health within various communities in Nottingham and their influence on knowledge translation for cardiovascular disease prevention. She has experience in corporate communication, community mobilisation, social marketing and communication strategies for family planning methods, post-abortion care and integrated community case management of child killer diseases in Cameroon and community engagement in the UK.

Sibyl Anthierens is a social scientist and associate professor at the Department of Family Medicine and Population Health and co-chair of the University Centre of Qualitative Health Research Antwerp (QUALUA), Belgium. She focuses particularly on evaluating the impact of evidence in primary care through implementation science. She works on how evidence-based treatment and management strategies are used or not used in everyday practice by patients and professionals and how this new evidence can be used to improve practice. Her research predominantly uses qualitative methods (process evaluation, general descriptive qualitative research and grounded theory). She supervises several undergraduate and PhD students.

Matias Barberis Rami is a researcher at the European Future Innovation System Centre, Belgium. He has more than ten years' experience in the fields of policies for disaster risk reduction, resilience and sustainability, socio-economic impact assessment, and research policy. He has expertise in qualitative and participative methodologies and speaks several languages. He has contributed to several research projects in Latin America and Europe on urban regeneration, citizen participation, migration strategies, and crisis and resilience management.

Hilde Bastiaens is a primary healthcare practitioner, researcher and Associate Professor at the Department of Family Medicine and Population Health Care of the University of Antwerp, Belgium. She is co-chair of the University Centre of Qualitative Health Research Antwerp and currently coordinates the Horizon 2020 project SPICES (Scaling-up Packages of Interventions for Cardiovascular disease prevention in selected sites in Europe and sub-Saharan Africa). Her research is mostly oriented towards implementation science in primary care in high-, low- and middle-income countries. Her research focuses on person-centred care, connected health, prevention, health behaviour change and the relation between the natural living environment and (mental) health(care).

Sarah Broadley is part of the Author and Community Engagement team at Emerald Publishing, UK, creating campaigns that bring researchers together around themes with the potential for real-world impact. She has worked as a marketer in academic publishing for more than 15 years. As part of Emerald's Policy Impact project, she has worked with academics and policymakers to understand their different aims and needs and find ways that publishers can help to bridge the gap between them.

Sally Burns is an educator, trainer, author and researcher living in France. She has worked in education for almost 40 years, teaching, managing, designing curriculum, consulting and facilitating training in primary, secondary and tertiary sectors, in private, state and international establishments in Oman, UK and Sudan. Sally co-authored *Values and Visions: Engaging Students, Refreshing Teachers* (The Values and Visions Foundation, 2019) and is a certified *V&V* trainer. She contributed chapters to Warwick University's Educational series: *Values Education and Cultural Diversity* (2000) and *The Omani ELT Symphony: Maintaining Linguistic and Socio-Cultural Equilibrium* (2012), and has published numerous education articles.

Helen Busby is an independent research ethics advisor. After studying medical anthropology and sociology at the Universities of Brunel and Nottingham, UK, she worked as a researcher at Nottingham. Having a particular interest in health, she was the Principal Investigator for projects about blood donation, stem cell banking and frameworks of altruism, and author of publications in the fields of sociology, ethics and law. Now focused on research ethics, she works with leading international organisations to improve ethical oversight of research.

Costas A. Charitidis is Professor, School of Chemical Engineering of the National Technical University of Athens (NTUA), Greece, and Director, Department of Materials Science and Engineering. He has extensive international R&D experience, authoring 196 articles, 14 book chapters and 2 books. He does ethics evaluations of European Commission funding proposals and coordinates the NTUA Advisory Ethics Committee. He helped draft Research Ethics and Research Integrity guidelines which are in the Greek Educational system legal framework. He coordinates Ethical Aspects in Research and Technology for Human network (EARTHnet) and is a member of the European Network of Research Integrity Offices (ENRIO).

Caroline Gans-Combe is Lead Senior Researcher with INSEEC (Institut des hautes études économiques et Commerciale) business school, France, where she works on innovation and value creation new paradigms. She has published more than 40 articles related to the modelling of value creation through the adoption by diverse institutions of ethical practices specifically with regard to the valorisation of collected data. She chaired a working group dedicated to best practices in the use of data in research (data protection and privacy ethical guidelines) and was appointed to the Expert Group Assessing the Investment Potential of small and mid-size enterprises (SMEs).

Fosca Giannotti is Full Professor at Scuola Normale Superiore, Pisa, Italy. She is a pioneering scientist in mobility data mining, social network analysis and privacy-preserving data mining. She leads the Pisa KDD Lab – Knowledge Discovery and Data Mining Laboratory, a joint research initiative of the University of Pisa and ISTI-CNR, founded in 1994 as one of the earliest research labs on data mining. Her research focuses on social mining from big data: smart cities, human dynamics, social and economic networks, ethics and trust, and diffusion of innovations. She is author of more than 300 papers. She has coordinated tens of European projects and industrial collaborations. Fosca is the former coordinator of SoBigData, the European research infrastructure on Big Data Analytics and Social Mining, an ecosystem of ten cutting-edge European research centres providing an open platform for interdisciplinary data science and data-driven innovation. She recently received a prestigious ERC Advanced Grant entitled, Science and technology for the eXplanation of AI decision-making (XAI).

Linda Gibson is Professor of Public Health in the Institute of Health & Allied Professions (IHAP) at Nottingham Trent University, UK. Linda has over 20 years' experience in health promotion and working with local communities in the UK and internationally, particularly in East Africa. Linda's research focus is on health systems strengthening in low- and middle-income countries, community health workforce in primary care in Uganda, non-communicable diseases, antimicrobial stewardship and patient safety. She is Principal Investigator for the UK Nottingham Scaling-up Packages of Interventions for Cardiovascular disease prevention in selected sites in Europe and sub-Saharan Africa (SPICES) project.

Hazel Goodes is Commissioning Lead for Journals at Emerald Publishing, managing a portfolio of academic journals in the fields of public policy, economics and tourism. Hazel has worked in publishing since 2007, and has experience of publishing academic journals and commissioning books across different subject areas. As co-lead of Emerald's Policy Impact project, Hazel has worked with a team of colleagues at Emerald to bring academic research to policymakers, culminating in Emerald publishing policy briefs for the first time.

Ron Iphofen is an independent consultant based in France with international recognition for expertise on research ethics and professional standards in research. He has served in the Association for Research Ethics UK, and advised on research ethics for the Scottish Executive, UK Government Social Research, UK Research Integrity Office, ANR (French Research Funding agency), and many others. He is widely published, with his most relevant edited volumes including *Ethical Issues in Covert, Security and Surveillance Research* (with Dónal O'Mathúna, Emerald, 2021), *Handbook of Research Ethics and Scientific Integrity* (Springer Nature, 2020) and *SAGE Handbook of Qualitative Research Ethics* (SAGE, 2018).

Helen Kara has been an independent researcher since 1999 and an independent scholar since 2011. She writes about research methods and ethics, and teaches doctoral students and staff at higher education institutions worldwide. Her books include *Creative Research Methods: A Practical Guide*, 2nd edition (Policy Press, 2020), *Research Ethics in the Real World* (Policy Press, 2018) and *Qualitative Research for Quantitative Researchers* (SAGE, 2022).

Panagiotis Kavouras is a physicist with a MSc and a PhD in Materials Science and Technology. His research interests focus on the development of experimental protocols for the measurement of physical properties at the nanoscale. He is the author of 46 peer-reviewed papers, a book chapter, a textbook and many participations in scientific conferences. He participates in four H2020 *Science with and for Society* (Swafs) and two Horizon Europe projects. He is a member of the European Network of Research Integrity Offices (ENRIO), the European Network of Ombuds in Higher Education (ENOHE), and he is the Greek correspondent for the SUPER-MoRRI project.

Gabi Lombardo is Director of the European Alliance for Social Sciences and Humanities (EASSH), the advocacy and science policy organisation for social sciences and humanities in Europe. She is an expert in both higher education and global research policy, and has extensive high-level experience operating at the interface of strategy, science policy, research support and funding. As Director of EASSH, she advocates for a strong evidence-based approach to policymaking, and the inclusion of researchers in science policy development for strategic and broad-based research funding. In 2018, she received the Young Academy of Europe Annual Prize.

Kristi Lõuk is Junior Research Fellow in Ethics at the Centre for Ethics at the University of Tartu, Estonia. Her research interests include research ethics, research integrity and medical ethics. She has degrees in philosophy from the University of Tartu and is currently completing her PhD, which focuses on issues related to research ethics and research integrity. She has been involved in numerous European Commission–funded research projects. Her work is related to the Centre for Excellence in Estonian Studies. Since 2016 she has been a vice chair of the Research Ethics Committee of the University of Tartu.

Nicolas Lux is currently Programme Manager at the Brussels office of Stiftung Wissenschaft und Politik (SWP), the German Institute for International and Security Affairs. At SWP, his research has covered the state of European integration and European defence policy. He has studied the role of foreign policy think tanks that are facing elite and science scepticism and political polarisation. He has contributed to a comprehensive study on think tanks by researching the Brussels policy advice landscape. He has presented at many conferences and chaired debates on the ongoing discussion process between think tank scholars regarding their role in challenging times.

Vassilis Markakis is a physicist (Aristotle University of Thessaloniki, AUTH) with a master's degree in nanosciences and nanotechnology (AUTH). His scientific work is in the fields of ethics in nanotechnology research and in dual use research, while he has participated in four Horizon 2020 funded projects related to research ethics and research integrity at the School of Chemical Engineering of the National Technical University of Athens (NTUA), Greece. In the past he has dealt with the dispersion of minerals into polymeric matrixes, as well with the study of AKT protein expression in healthy and cancer cells.

Katharina Miller is a change agent using legal tools for ethics and integrity in innovation and technology, currently working for the H2020 project Path2Integrity. She is also a European Commission Reviewer and Ethics Expert, Advisory Board member of the Berkeley Center on Comparative Equality and Anti-Discrimination Law and President of the European Women Lawyers Association.

Tebogo Maria Mothiba is Executive Dean in the Faculty of Health Sciences at the University of Limpopo in South Africa, having served there as a Research Professor from 2018 to 2020. Her research interests are in Health Service Management and Non-Communicable Diseases. She has worked as Co-Principal Investigator on projects such as the Scaling-up Packages of Interventions for Cardiovascular disease prevention in selected sites in Europe and sub-Saharan Africa (SPICES) project at the University of Limpopo. She has participated in several funded research projects on strengthening health systems for maternal and child health, and other chronic diseases.

Dónal O'Mathúna is Associate Professor, College of Nursing and Center for Bioethics, The Ohio State University, USA. His research focuses on systematic reviewing and on ethical issues in humanitarian settings, disasters and pandemics. He has led funded research projects and contributed to ethics initiatives with international agencies. He has authored numerous peer-reviewed articles and co-edited *Ethical Issues in Covert, Security and Surveillance Research* (with Ron Iphofen, Emerald, 2021), *Ethics and Law for Chemical, Biological, Radiological, Nuclear & Explosive Crises* (Springer Nature, 2019), *Disasters: Core Concepts and Ethical Theories* (Springer, 2018) and *Disaster Bioethics: Normative Issues When Nothing is Normal* (Springer, 2014).

Francesca Pratesi received MS and PhD degrees in computer science from Pisa University, Italy, in 2013 and 2017, respectively. She is a researcher at the Information Science and Technology Institute of the National Research Council in Pisa since December 2021 and a member of the Knowledge Discovery and Data Mining Lab, a joint research group with the Computer Science Department of the University of Pisa. Her research interests include data mining, data privacy and privacy risk assessment, mainly in spatio-temporal data. Recently, she has broadened her interest, moving towards the Ethics-by-Design paradigm and Trustworthy AI.

Ian Roderick is the director of the Schumacher Institute for Sustainable Systems in Bristol, UK. His career began at the Building Research Establishment looking at complex project management. He moved into business planning and became head of strategic forecasting at a large international company. He co-founded a successful software business, moved to venture capital investment in the software industry, and then concentrated on systems thinking applied to environmental and social justice. He is a co-founder of a charity, The Converging World, that installs energy systems and plants trees in India.

Alan Simpson was a Member of Parliament (MP) in the UK for 18 years and since then has concentrated on being a climate policy advisor, writer and campaigner. Formerly Sustainable Economics Advisor to the UK Labour Party's Shadow Chancellor, his work focuses on the transformative changes needed if we are to live securely within the climate roller coaster we have created.

Paul Spicker is Emeritus Professor, Robert Gordon University, UK, where he was Grampian Chair of Public Policy (2001–15) and Director of its Centre for Public Policy and Management. He is now a writer and commentator on social policy. His published work includes 22 books, several shorter works and almost 100 academic papers. His early research focused on housing and welfare rights, and since then has included studies of benefit delivery systems, care of old people, psychiatric patients, housing management and local anti-poverty strategy. He is a consultant on social welfare in practice, and has consulted for local, national and international agencies.

Eleni Spyrakou is a researcher at the School of Chemical Engineering of the National Technical University of Athens (NTUA), Greece. She has degrees in philosophy, with her PhD focused on Political Philosophy. Her main research interests include applied ethics (bioethics, environmental ethics, business ethics, research ethics and integrity), the ethics–politics relation and its epistemological problems, and the methodology/ies of social sciences. She participates in four *Science with and for Society* (SwafS) Horizon 2020 projects. She is a translator from English into Greek of textbooks on emerging literacy and its applications to Education, and a forthcoming volume on the philosophy of digitality. She has presented at national and international conferences.

Margit Sutrop is Professor of Practical Philosophy and a founding Director of the Centre for Ethics at the University of Tartu, Estonia. She was elected a Member of the Estonian Parliament in January 2021. She is a member of Academia Europaea and an independent ethics expert of the European Commission. Her research interests include moral and political philosophy, philosophy of mind, bioethics, research ethics and research integrity, AI ethics, aesthetics and philosophy of education. She is also a member of the international advisory board of the Centre for Ethics at the University of Tübingen. She has published over 100 peer-reviewed articles and edited 15 books.

Roberto Trasarti is a member of the Information Science and Technology Institute 'A. Faedo' of the National Research Council (ISTI-CNR), Pisa, Italy. He is also a member of the Knowledge Discovery and Delivery Laboratory. His PhD thesis was entitled 'Mastering the Spatio-Temporal Knowledge Discovery Process'. His research interests include data mining, spatio-temporal data analysis, artificial intelligence and automatic reasoning.

Nina J. Zugic is an innovative education and learning specialist, currently working as a Principal Consultant and European Commission Expert Evaluator, Reviewer and Research Ethics Expert. She is the Horizon 2020–funded MARINA Project Principal Consultant and Ethics Expert, and is on the Ethics Advisory Board of the NAUTILOS Project.

Fabian Zuleeg is Chief Executive of the European Policy Centre (EPC), and is also its Chief Economist. He has been an economic analyst in academia, and public and private sectors. His interests include the political economy of European integration, particularly the UK–EU relationship and Brexit. He also works on the future of think tanks in Europe. He works closely with decision-makers in European institutions, with EPC members and partners and the wider Brussels community. He regularly comments on current EU political and economic issues in the media, and has published widely on European integration and European economic and social policies.

Introduction: Ethical evidence and policymaking

Ron Iphofen and Dónal O'Mathúna

Scientists, science advisors and journalists have an obligation to ensure clarity and openness in debates and discussions about the kinds of issues that attract policy decisions, especially when they involve controversy, uncertainty or entrenched viewpoints. Public trust in and understanding of science is enhanced when policies appear with transparent discussions about the evidence on which they are based, acknowledgement of areas where evidence is lacking, uncertain or contradictory, and the limitations to their conclusions. Mutual trust among citizens, and between citizens and the state, is essential to a healthy functioning democracy. The more transparent the science and the policymaking, the less likely that biased or fake reports will influence the public and policymakers. This book aims to provide clear and practical ways to achieve truly evidence-informed policymaking.

The chapters in this collection were inspired by the EU-funded PRO-RES project which produced a guidance framework that can help to deliver responsible research and innovation (RRI). PRO-RES is a Horizon 2020 project coordinated by the European Science Foundation, involving 14 different partners across Europe.[1] The chapter authors here were drawn from consortium partners, stakeholders to the project and affiliates who support the project's aims. All seek to encourage policymakers and their advisors to seek evidence for their decisions from research that has been conducted ethically and with integrity.

The PRO-RES project was conceived in response to a concern that too many well-funded and well-conducted projects in the fields of research ethics and scientific integrity become 'shelf-bound' on completion and fail to impact where it matters most – on policymakers and their advisors. As one of a series of open access publications of the project (see Iphofen and O'Mathúna, 2021), this volume needs to be, in the first instance, something that can be placed in the hands of policymakers to give them practical and efficient ways to respond to the issues addressed. In this sense, it has a pragmatic aim – influencing where it counts. We also intend this work to have another applied benefit – for research ethics reviewers tasked to appraise proposals and help ensure that they are ethically sound. The issues addressed here have proven to be some of the more challenging facing policymakers and, indeed, research reviewers. Ethics is always about dilemmas – making difficult choices, and they must always be made in light of the most rigorous information. Ill-informed choices lead to poorly chosen decisions. Such decisions have consequences for individuals, communities and societies. It is only when the integrity of research is carefully pursued that the evidence produced can be assured of its value. Finally we also hope this collection will be used to train researchers in good practice in the production of research evidence. If they learn the lessons to effectively deliver ethical evidence[2] early on, it will help avoid mistakes and heighten the impact of their subsequent work. If we can affect both ends of the policy 'supply chain' we hope to benefit science, society and the public at large.

This volume will challenge policymakers and their advisors to find ways of ensuring that the research they seek, and the evidence they use and apply, was generated ethically. In that sense, we intend for it to challenge current policymakers and their advisors who may feel themselves unobliged to disclose the rationales for their decisions, especially if they are not supported by the evidence (see Sedley, 2016). At the same time, policymakers face many challenges in their work, which have been taken into account by the authors of the chapters. The choices made to prioritise the issues addressed by policymakers, and how they select the relevant evidence, may be influenced by ethical concerns, but also, inevitably, by cultural, socio-economic and political values. Policies will often have good and bad outcomes for different groups of people, and the underlying values have to be balanced by policymakers. Such issues have become both transparent and controversial during COVID-19, as exemplified by the refrain that 'the cure should not be worse than the illness it seeks to address'. It is therefore crucial that policymakers articulate why they based their policies on one approach to an issue when competing claims are calling for different policies. In many cases the reasons can be traced to certain evidence being viewed as more convincing. Such decisions should be made transparently and on grounds that can be demonstrated to prioritise the public good.

The challenges faced by policymakers as they grapple with values, ethics and cultures – their own and others' – will be examined specifically and in detail in Chapters 11 and 12, with their foci on COVID-19 and other crises. But with the contemporary prominence of a pandemic all of these issues will be addressed in other chapters that cover different, but equally compelling problems. Part I of this volume provides an overview of the issues that are vital to ethics and integrity in science. These relate to how global the underlying values and principles can be assumed to be, whether a Westernised hegemony continues to apply, or if and how cultural variations are inevitable and must be managed. Some consistency in standards seems necessary to facilitate international and interdisciplinary work and to encourage the trust that is vital for researchers to collaborate and for the general public to have faith in their products. Part II seeks to apply these issues to a selection of topics of immediate contemporary relevance but which undoubtedly will concern us for many years to come. These include the two just mentioned related to COVID-19, and other areas including technological developments like genetic modifications (GM) and artificial intelligence (AI), and research involving children or public policy. Part III examines particular case studies on policymaking and evidence, including ones on Black Lives Matter and think tanks focused on foreign policy.

We cannot overemphasise the importance of getting the balance right between the professional autonomy of researchers and any regulation or interference by the state. While the danger of political ideology 'directing' research is ever-present, the pervading historical example of what can happen when things go wrong remains the Holocaust (see for example Ciesielska et al, 2020). Stacey Gallin makes the point:

> The problem is not that the system does not work to a point, but it lacks the most important element of the moral formation of the professional (medical) researchers themselves, which would install a robust inner moral compass giving one the ability to withstand political and ideological pressure and forces – and even the pure self-interest. (Gallin, 2020: 11)

It is that 'robust inner moral compass' that is essential to researchers sustaining their integrity in the face of political pressures. It is for this reason that policymaking and policymakers themselves must be open to investigative research that may bring challenges to conventional principles of research ethics. In his chapters here, Paul Spicker (Chapters 2 and 7 and in many other publications, see for example Spicker, 2007) outlines how ethical policy research needs to be guided primarily by public accountability and societal welfare – principles that may require challenges to consenting processes and confidentiality. To conduct ethical research that results in sound evidence for policymaking requires those challenges to be met head on.

Beyond academia: from RPOs to EGOs

Our focus here is not just on 'social' policy – it is about *all* policymaking that draws upon evidence from research. The PRO-RES project addressed issues that affect all non-medical sciences – from environment to engineering, from public health to public relations, from the humanities to nanotechnology, and much more. Given the range of evidence employed in policymaking it became obvious that we had to adopt a very broad definition of 'research'; one that included all forms of data gathering that could be or was intended to supply evidence for policymaking. As a result the agencies gathering the useful data might include academic researchers in Research Performing Organisations (RPOs), but also in think tanks (Chapters 6 and 17), lobbying agencies, public relations consultants, advocacy agencies, civil society organisations, early adopters/influencers (bloggers and so on) and conventional news media. These criteria would not rule out novice researchers, citizen scientists and members of the public. All these 'agencies' could be regarded as 'Evidence Generating Organisations' (EGOs) – able to gather evidence that could be used to make policies or applied in a range of socially useful ways. There is no explicit requirement for only experienced researchers to be treated as 'legitimate'. The key is to be transparent about exactly who the researcher/agency is and who they are working for – even if it is for themselves. It is to be expected that researcher CVs/resumés would be made available together with any agency track records, details about the RPO/EGO/Agency's background and its main funding sources – which could be large corporations with heavily vested commercial interests or crowdfunding schemes in which the interests might be more diverse. Mission statements or adherence to codes and guidelines and/or professional association memberships would be appropriate here. A key question for the evidence-gathering agency would be: how does the EGO fund itself? Does it have a diversity of funding or is it dependent on a particular stakeholder and with what contractual commitments?

Ethics and method

Evidence can't always offer confirmation of the validity of decisions taken – but it certainly can nuance or 'qualify' existing policy views. Research designs vary widely and can be subject to a wide range of perspectives on their value, so policymakers often have to assess their worth according to the available criteria (Chapter 12). We make no implicit judgement of the 'ethical quality' of the variety of methods that can be employed. What matters is, again, the transparency of those conducting the research, and their offering of clear justifications and rationale for any methods used. Thus covert research, deception, community/societal engagement, social engineering and so on are not to be regarded as inherently unethical – the judgement of whether they are or not might depend upon the context in which they are used and, whether a policymaker/advisor can make a valid case that the evidence derived from a particular method is morally justifiable both in terms of its source and its application (another PRO-RES volume examines some of these methods in detail: Iphofen and O'Mathúna, 2021). Of course there is always the risk that evidence claimed as morally justifiable in one sphere is used unjustifiably for another purpose – the 'dual use' problem.

Neither is there any implication that only primary research is of evidential value – all forms of secondary data analysis can be subjected to these questions: from meta-analyses of controlled experimental studies to simple frequency counts of questionnaire responses. The 'validity' of primary research data depends upon the rigour of the research design and its accurate execution; the validity of most forms of secondary data analysis depends upon access to and availability of raw source data and the rigour of the secondary analyses. Even documentary or archival analyses are to be tested against accurate use of source materials according to the standards of the methods chosen.

Caring for 'subjects'

Whatever research method or research design is adopted, the researchers' relationship with their subjects – human or otherwise – is key to the ethical appropriateness of findings. These elements concern the relationship between researcher and researched, and how the researcher treats the researched. The subjects/objects/participants could have been humans, animals, organisms or parts of such, material objects, ecosystems, organisations, communities, societies and so on – or any combination of the aforementioned. Thus research by economists might be a study of banking 'systems' without reference to bankers per se. Research enquiries related to public health might be concerned with the public – the society, the community – and not individual members of that 'collectivity'. Researcher welfare issues are likely to arise out of their relationships with the subjects/objects of study, so researcher health and welfare needs to be considered and any forms of reflective practice they adopt encouraged and disclosed. Once more these questions are not just related to primary research, nor simply to research with humans or live animals – they apply equally to any form of secondary research or data gathering and to material objects or places. Thus, for example, volcanologists are unlikely to be able to cause undue harm to the objects of their study, but are likely to put themselves at risk when engaging with the primary objects of their attention. On the other hand, if they adopted some physical engagements with volcanoes (such as planting explosives to observe effects), the possibility of harm to other aspects of the ecosystem and communities has to be envisaged. At the same time, a piece of research might be completely geological, but if the methods (like collecting samples) have strong cultural meanings for local communities, those factors need to be addressed and discussed as part of the ethical evaluation of the research (Fraser, 2019).

Research motives

Motive and intent are always foundational ethical issues. They go to why the research was conducted in the first place and what outcomes were hoped for and by whom. It is important to look at both the research agents' motives and those of the funder or commissioner. What did the funder hope for in commissioning the research and what might the researcher seek to achieve by competing to win the 'call'? Both were surely keen to achieve some form of impact – either to add to the sum total of human knowledge and/or to ensure benefits to individuals, communities or societies. Hopefully, the motivations were not merely to advance their own esteem in the eyes of the public or advance their careers, though such outcomes remain part of the big picture.

Impacts could be environmental, social, psychological, political and so on, and hence the question of who commissioned and funded the research/enquiry is doubly important – it is key to full transparency. Outcomes must be related back to intent. And, by the same token, all stakeholders must be aware of and allow transparency on unintended or unanticipated outcomes. The many complex variables involved in societal policymaking indicate just how limited the control of the significant variables can be sometimes. Again, for such reasons, the open and honest evaluation of outcomes becomes a key element in the ethical use of evidence. It is vital to know how the evidence coming from the research was implemented and if it 'worked'.

Research ethics review

Many steps in ethics approval and/or appraisal processes assess the quality of and risks (ethical, economic and political) involved in research projects. In some countries and some institutions these processes are absent or severely under-resourced. The increase in multinational, interdisciplinary approaches to research places a responsibility on research institutions to ensure that some formal reviews are conducted (Chapter 11). Reviewing standards and standard operating procedures are increasingly shared internationally and across institutions leading to an expectation that certain research principles and values are becoming accepted universally. In addition, novel citizen science evaluation methods are emerging, such as crowd reviewing. It may be difficult for every form of RPO or EGO to secure independent assessment for the ethics of their work. Increasingly organisations do strive to establish their own in-house system with a degree of independence provided by some external (lay and professional) memberships. No 'best' or single way exists for doing this, but transparency is key – clarifying if any form of assessment of quality and ethics is done prior to the commencement of research and/or enquiry and how it is conducted.

The European Commission's own system of research ethics appraisal has grown and developed throughout the Commission 'framework' research funding programmes and is well understood by experienced researchers and evaluators alike. Together with a series of guidance notes and support documents, the system offers some confidence for both the general public and policymakers that the research being funded will be safe, protecting those who participate and those who conduct it, and offer robust outcomes that can benefit science and society. It does offer something of a 'model' of good practice for how such appraisals should be conducted (Kinderlerer and Schroeder, 2016).

We have no intention of spending too much time on the ethics review process in this volume. The perception by some researchers that ethics review is overly obstructive, lacks understanding about specific methodologies and, on occasion, is undemocratic is well covered – even excessively so – elsewhere (see Iphofen, 2009: Ch. 13; Iphofen, 2017). However, there are certainly lessons for reviewers, for those who construct research ethics appraisal systems and for those who submit to them, to be found in this collection. We would challenge the notion that ethics review is 'risk averse' and therefore obstructs innovation and research progress. Rather it is the aim of this volume and the PRO-RES project itself to raise 'risk awareness' and that might mean encouraging innovative research that contains risks and a culture of risk assessment that recognises the inherent nature of risk in all scientific progress. This will mean that reviewers cannot be assumed to always agree on what the risks may be and whether they are worth taking, but achieving consensus in ethics review relies upon a collective responsibility, the growth of shared understanding between researchers and reviewers, and a commitment to best practices (see van den Hoonaard and Hamilton, 2016).

Context

Nevertheless, most ethical judgements rely upon a full understanding of the context in which the action under consideration occurred. This requires a comprehensive understanding of place and time: geographical, institutional, organisational *place* and diurnal, annual, chronological and historical *time*. Thus when research proposals are being assessed the reviewers must take into account the wide contextual variations between a laboratory site, urban settings entailing risk and threats, libraries containing archives, and high- and low-resource countries. Laboratories can vary in licensing levels, while field sites vary in the kinds of permissions required. Historical archival research varies considerably in terms of ethical risk from the study of more contemporary documentation but engaging in historical enquiry may still entail risks to the present in terms of societal or communal stigmatisation and/or reputation. For example, knowledge of how and why a particular organisation was established may 'taint' its current reputation. Users of research emanating from such diverse contexts must take into account the range of settings in which the researchers' findings emerged. This means that any easy transposition of policy from one setting to another, without taking into account the full contextual factors, is not possible. Clearly researchers have a responsibility to ensure that policymakers and their advisors do not assume the evidence generated in one sphere is necessarily applicable elsewhere.

Hence the importance of careful dissemination of research findings. Findings can be disseminated in a range of different ways – in academic publications, peer-reviewed scholarly publications, in-house technical reports, commissioned reports, independent white papers, official policy documents, policy briefings, participant feedback, social media, news media and so on. What is done with such 'outcomes' links back to the original 'why?' question, or what was hoped for or intended for the research. The researchers might not be in a position to directly apply the findings, but they might be better able to guide and assist those who can – the policymakers. Publishers can play an important role in helping researchers 'translate' their findings into forms and formats which are better suited for policymakers. Chapter 18 provides a case study from one academic publisher on how they are assisting in this area. The methods discussed there are seen by many publishers as a way they can assist researchers fulfil their ethical responsibility to get their results into the hands of those who can use them, especially policymakers. Think tanks can play a similar role here in helping researchers present their findings in policy briefs that will be more likely to be used, which simultaneously places ethical responsibilities on think tanks to offer policy advice with integrity (explored in Chapter 17). Another case study (Chapter 16) examines systemic racism in education policy and shows what can happen when existing evidence has not been used to influence policy change and when the impact of policies themselves is not evaluated.

On the other hand, a decision might be made to withhold publication of findings on the grounds of risk of harm – and justifications for such an action would have to be clear and strong. In this sense, researchers and evidence analysts have to think through what policymakers might do with their findings for their own personal or political gain. The problem here is that those conducting the research may not be best placed to manage the media outputs – but that does not mean that they should not try. The ethics of research drives through to what happens to the final message.

The theme

Under European Union (EU) regulations, the PRO-RES project was a Coordinating and Support Action, so its main aim was to help progress science advice; it was not about knowledge production. The idea was to bring together the existing knowledge, analysts and knowledge producers to find ways of bringing the best advice on ethical evidence to the door of the policymakers.

While the European Commission's funding call for the PRO-RES project specifically sought to emulate the Oviedo Convention and the Helsinki Framework in medical fields, the PRO-RES Consortium acknowledged the debt owed to the countless existing codes and guidelines in other fields. Many international organisations have continued to review and update their ethics codes over many years and PRO-RES has taken account of those developments. For example, the Council for International Organizations of Medical Sciences (CIOMS; established 1949) updated their 2002 Code in 2016 and ensured it was related to 'health' and not just 'medicine' – in recognition of the need to understand the societal and cultural contexts in which medical and clinical research is conducted. Subsequent Declarations of Helsinki for the World Medical Association (WMA; originally 1964 and updated most recently in 2013) are well known, but it is clear that the Declaration 'is addressed primarily to physicians. The WMA encourages others who are involved in medical research involving human subjects to adopt these principles' (WMA, 2013: Preamble, para 2). And indeed there are considerable overlaps in the values, principles and standards across all disciplines and these should hold in other research fields. Given the existence of all this preceding work, the PRO-RES project did not attempt to reinvent the wheel nor ignore the valuable work done in producing the diverse codes and guidelines for ethical research. The project delivered a resource listing as many of these codes and guidelines as could be found and drawing out those values, principles and standards that are held in common.

This current volume is part of the overall endeavour to bring advice on ethical evidence to policymakers. Each chapter has been written to attract a readership that includes policymakers, their advisors, funders, publishers and researchers (including think tank analysts and report writers). To avoid overly abstract or academic approaches, we sought discussion, with some chapters drawing upon experience and others focused on problem solving. This has resulted in a diversity of writing styles that we hope appeals to our wide range of stakeholders. Thus this book is not aimed primarily at a scholarly audience, which might be interested only in research ethics and research methods.

Recognising the time limitations on policymakers and their advisors, each chapter offers an 'executive summary' set of recommendations to facilitate an effective focus on the specifics while addressing the substantive focus of each chapter. This summary reflects the chapter structure we proposed to our authors. First we asked them to outline 'the problem' as they envisaged the topic they were covering. Then we asked them to discuss the status of the 'evidence' explaining or describing the problem or issue of concern. We wanted to ask, can the evidence and its source be trusted? What makes evidence 'reliable'? Does it matter if evidence derives from government reports, funded academic research, investigative journalism or anything else?

How can debates over evidence be addressed ethically? Finally, we asked them to attempt some assessment of what can be done about 'the problem'. What could be considered feasible in an economically, culturally, politically acceptable sense? What obstacles/hurdles could hinder progress? Given those identified limitations, what *should* be done – or at least attempted? In other words: what advice about managing the problem could they give to policymakers?

The target

In part this volume will be a guide for policymakers and their advisors. But it cannot be prescriptive. It must also take account of the challenges policymakers face as their own values and culture interact with those of others as they address issues of ethics and evidence. Some polemical elements are necessary to engage the many dilemmas about policies in these fields that must be addressed and thought through – that is what policymakers understand: debate and disagreement. It is our view that our main target audience will not tolerate heavily academic or even scholarly works. They seek solutions, evidence to back up the decisions that they take, and issues being discussed in a forthright manner.

While we hope that these chapters will have global relevance, the funding source for the PRO-RES project and this volume has, of necessity, led to more of a European focus. And this volume alone cannot achieve the aim of ensuring policymakers seek to use only ethical evidence. This volume is only part of a much larger effort intended to achieve that aim – other open access publications (including Iphofen and O'Mathúna, 2021), articles in the professional and more popular press, a website, workshops, one-to-one engagements and, crucially, the behind-the-scenes networking of the consortium partners and the declared stakeholders to the project.[3] As an element in the overall action it is necessarily selective as to topics and approach.

Part I of this collection addresses the 'umbrella' concepts for ethical evidence under which specific topics can be approached. The issues here involve whether or not there are, can be or ought to be global principles and standards that can be agreed regardless of country or nation. Part of that problem is the suggestion of the hegemony of Western European values, which reeks of postcolonialism in some spheres or even a kind of cultural arrogance. This poses problems for the practicalities of the ethics review of international and/or interdisciplinary research proposals and the forms of guidance that might be seen as acceptable across and between cultures and countries. Ultimately this can affect issues of mutual trust – between scientists and the public and, equally importantly, between scientists and those seeking evidence to support policymaking.

Part II addresses some key substantive topics that were proposed in the call to which the PRO-RES project responded, and which have been stimulated by the project and by contemporary events. These topic areas remain of contemporary concern and interest. They include ethical decision-making for agri-industry/food policies (Chapter 8), ethical issues in AI and robotics (Chapters 9 and 10), research ethics in disasters and crises (Chapters 11 and 12), and ethical policies in finance and economics (Chapter 13). We were interested in the ethical concerns over public policy in general (Chapter 7) and it was particularly timely that this enabled us to include discussion about ethically conducting research into and dealing with pandemics.

The issues raised in these different topic areas cannot be considered as discrete concerns. When confronting the evidence related to a pandemic, for instance, sources of data as well as their management become vital; how we understand 'the environment' and how we manage 'crises' underpins varying perspectives on societal 'collapse'; research involving children can occur in any of these areas (Chapter 14); the mutual trust already referred to is undermined by deception and lack of transparency; and all efforts at comprehensive solutions to a national and/or global crisis will be frustrated in the face of structured or systemic social inequality.

Hence it was important that one of the case studies in Part III address concerns about evidence and policy with implications for systemic racism (Chapter 16). The cases studies in this part conclude the volume by examining specific ways that ethical evidence can be incorporated into good policymaking.

The chapters are written by authors coming from a variety of disciplines. That in itself does not make this an interdisciplinary volume. However, many of the chapters come from those who have been actively involved in the PRO-RES project, which is an interdisciplinary and international consortium. These chapters are informed by the discussions and debates that have occurred between members of different disciplines, both online and in meetings and workshops. Furthermore, most chapters were made available for review by all other contributors to the volume so that their discipline-specific input can inform the final version of each chapter. In this way, the final book is the result of a truly interdisciplinary and international collaboration.

Notes

[1] PRO-RES is a European Commission–funded project aiming to *PROmote ethics and integrity in non-medical RESearch* by building a supported guidance framework for all non-medical sciences and humanities disciplines adopting social science methodologies (https://prores-project.eu/). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788352.

[2] By 'ethical evidence' we mean evidence that has been collected, compiled, analysed and applied ethically and with integrity.

[3] https://prores-project.eu/

References

Ciesielska, M., Gajewski, P. and Bałuk-Ulewiczowa, T. (eds) (2020) *Medical Review Auschwitz: Medicine Behind the Barbed Wire – Conference Proceedings 2019*, Kraków: Polish Institute for Evidence Based Medicine

Fraser, B. (2019) 'Scientists sample Earth's highest glacier', *Nature*, 573: 171–2.

Gallin, S. (2020) 'The Holocaust as a critical point in the development of medical and research ethics', in M. Ciesielska, P. Gajewski and T. Bałuk-Ulewiczowa (eds) *Medical Review Auschwitz: Medicine Behind the Barbed Wire – Conference Proceedings 2019*, Kraków: Polish Institute for Evidence Based Medicine, pp 17–26.

Iphofen, R. (2009) *Ethical Decision Making in Social Research: A Practical Guide*, London: Palgrave Macmillan

Iphofen, R. (ed) (2017) *Finding Common Ground: Consensus in Research Ethics Across the Social Sciences*, Advances in Research Ethics and Integrity, Volume 1, Bingley, UK: Emerald Publishing.

Iphofen, R. and O'Mathúna, D. (eds) (2021) *Ethical Issues in Covert, Security and Surveillance Research*, Advances in Research Ethics and Integrity, Volume 8, Bingley, UK: Emerald Publishing.

Kinderlerer, J. and Schroeder, D. (2016) *Assessment of the Ethics Appraisal Process of Horizon 2020*, Brussels: European Commission.

Sedley, S. (2016) *Missing Evidence: An Inquiry into the Delayed Publication of Government Commissioned Research*, London: Sense About Science.

Spicker, P. (2007) 'The ethics of policy research', *Evidence & Policy*, 3(1): 99–118.

WMA (World Medical Association) Declaration of Helsinki (2013) *Ethical Principles for Medical Research Involving Human Subjects*, available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 24 January 2022].

van den Hoonaard, W.C. and Hamilton, A. (eds) (2016) *The Ethics Rupture: Exploring Alternatives to Formal Research-Ethics Review*, Toronto: University of Toronto Press.

PART I

Fundamentals of evidence and ethical dimensions in policymaking

1

Ethical research in a global context: a dynamic tension between universal values, principles and contextual applications

Margit Sutrop and Kristi Lõuk

Introduction

At the level of research policy there is a growing consensus on what counts as ‘ethical research’ across Europe, North America and in other countries that can be seen as influenced by ‘Western’ cultural perspectives. Standards set by these countries are regarded as requirements for all who wish to engage in research seen as ‘acceptable’ by their research institutions, funding organisations and publishers. Since international research projects are becoming more and more common and, besides Europe and North America, China, India and South Korea are also currently making huge investments in research and the development of new technologies, agreement is sorely needed on values and principles of ethical research. The growing interest in organising international congresses such as World Conference on Research Integrity shows that globalisation of ethics in research is a popular trend.

The challenge to the globalisation of research ethics is how to solve the tension between universal values and principles on the one hand and their contextual applications on the other. Focusing too much on universal values and principles without sensitivity to different cultural contexts, research fields and specific contexts of application may run the risk of trying to hit too many birds with the same stone. Debates in research ethics show that there are various understandings of values which are sometimes equated with valuable objects, principles or virtues. In the philosophical literature there is no agreement on whether ethical research consists of virtuous behaviour, acting upon principles or a mixture of both. It is also debatable whether there are universal principles at all, and if so, what these might be. Pluralism also prevails in the field of research integrity: professional ethics codes express different values, principles and standards. A further complication is that even if codes and regulations make use of the same words, these can be interpreted differently, and their relationship to each other (their hierarchy) depends on the ethical framework employed by those applying the principles to specific issues and research contexts.

The problem is that the regulation of ethics in research does not seem to take all of this diversity into account. While research ethics committees and review boards have standardised procedures for handling ethical issues in research projects, and research integrity offices have adopted similar policies to handle scientific misconduct, researchers themselves have raised several concerns. There have been complaints that the whole framework of values and principles is Eurocentric (Jordan and Gray, 2013; Zhang, 2017; de Albuquerque Rocha and Vasconcelos, 2019; Dahal, 2020; Grant and Gazdula, 2020) and based on the perspective of biomedical sciences; the allegation is that it does not fit with non-medical sciences (for example, Haggerty, 2004; Schrag, 2010, 2011, and see PRO-RES project); that the ethics review system is too formal and boils down to a tick-the-right-box approach. A recent study of the ethics review of the grant proposals in the Commission Horizon 2020 showed that ethics experts reviewing the projects identified twice as many ethical issues compared to applicants across funding schemes (Buljan et al, 2021). Furthermore, the rapid development of research integrity systems for handling misconduct treats all scientists as potential 'criminals' who must prove that they are not in violation of any norms and regulations. Thus, instead of securing trust between scientists and society and among scientists themselves, the growing importance of ethics in science is perceived by scientists as a sign of mistrust and bureaucratisation.

There is an evident need to bring ethical and legal norms of research more in line with scientists' own values. To secure mutual trust and the acceptance of norms and standards for ethical research, these should be made understandable to scientists and aligned with their own values. If there is a gap between norms and values, there is a danger that scientists will try to ignore or violate the norms, and this opens the door to misconduct. If researchers follow the norms only because they must do so or because they fear sanctions, they lack integrity. Researchers' integrity means 'wholeness' – harmony between one's values and actions.

In this chapter we will suggest two ways to avoid this problem. First, we argue that the development of norms and regulations should avoid 'imperialism' and keep in mind the dynamic tension between universal values and their contextual applications. Second, we should invest more in education which, through developing researchers' reflexivity and sensitivity to different contexts, can lead to an agreement on universal values.

The search for universal values and principles in research ethics and research integrity

The search for universal values and principles has been pursued by two different fields of study: research ethics and research integrity. While research ethics has been an established field since the Second World War, research integrity is a newer endeavour, beginning in the 1980s. If we look back at the rise of *research ethics*, we see that it is largely a response to malicious use of scientific developments such as nuclear weapons or harmful research endeavours using prisoners, psychiatric patients or other vulnerable groups without their consent (for example, the Nazi medical experiments, the Imperial Japanese experiments in China, the Tuskegee syphilis study, the Milgram experiments, the Tearoom Trade study; see Emanuel et al, 2011; Briggle and Mitcham, 2012; Israel, 2015; Shamoo and Resnik, 2015).

Also, the field of *research integrity* has its origin in scandals where researchers have been caught at fabrication or falsification of data, or plagiarism (for instance, the Baltimore affair, the debate over cold fusion, the Hwang case; see Resnik, 1998, 2009; Shamoo and Resnik, 2015). On the basis of such negative experiences, moral values have been articulated that need protection and principles have been formulated that scientists must follow in their work.

Although research ethics is not a coherent field, it can be claimed that its central focus is to protect research subjects (human and animal). For human subjects this means respect for their autonomy, keeping participation voluntary and obtaining free and informed consent. Another important element of research ethics is the protection of research subjects' privacy and the confidentiality of the data processed in the research project. In addition, research ethics maintains that risks should be managed and research should be beneficial, either to specific participants or, in the future, to all who find themselves in similar situations. Research ethics review/appraisal systems were established to ensure these conditions. At present journals and/or funders of research projects may request approval by a research ethics committee (REC) or institutional review board (IRB).

Research integrity has emerged as a response to various misconduct cases concerning fraud, falsification and plagiarism. However, integrity in research should not be regarded narrowly as misconduct (including questionable research practices), but should rather be treated more broadly, encompassing researchers' ideal behaviour when conducting and organising their work. Sarah R. Davies has suggested that: 'Research integrity as an issue might also be placed in a much wider frame, one that is concerned with the soft governance of science through codes of conduct, ELSI activities, or ethical norms' (Davies 2019: 1237). David Shaw makes a distinction between *scientific* integrity and *research* integrity, arguing that the first of these includes 'not only the research endeavour itself but also the knowledge produced by previous research' (Shaw, 2019: 1086).

For a long time the fields of research ethics and research integrity have existed independently, but recently more and more arguments have been advanced for looking at them in tandem as ethics and integrity have many overlapping issues (Iphofen, 2020: 17). The European Commission has initiated and funded projects to investigate research integrity and scientific integrity together (for example, EnTIRE and VIRT2UE). According to Robert Braun, Tine Ravn and Elisabeth Frankus (2020) the difference between research ethics and scientific integrity can be seen as ethics committees doing their job in advance, while research integrity offices handle (possible) misconduct cases retrospectively. While research ethics focuses more on the planning and design of a project, its specific research outline, research integrity primarily stresses conduct and implementation. Therefore both research ethics and research integrity offices are needed for research to be ethical: depending on the direction in which the research is headed, one or the other may be foregrounded.

What counts as ethical research?

Good science has two simultaneous dimensions: high research quality and ethically correct science. The basis of ethics in research is a collection of values, principles and virtues that scientists must uphold in their work. In broad terms one can describe the requirements as follows: at all stages of research, scientists must observe certain values and principles, which together constitute good scientific practice. When planning research, it is essential that they work in conjunction with the values, principles, standards and legal norms of research ethics, and obtain the requisite permits, agreements and coordination with their ethics committee. When applying for research grants, publishing results and giving expert opinions, scientists must report financial or other conflicts of interest that might negatively affect the trustworthiness of their work. In conducting research it is important that they employ appropriate research methods, use critical analysis when drawing conclusions, and publicise data fully and objectively. The scientist neither falsifies data, supplements at will, nor invents data. When studying human subjects and collecting their personal data, the researcher must obtain informed consent of the research subject and guarantee that the agreement be conscious and voluntary. The scientist must honour the freedom of choice, autonomy, human dignity and privacy of those participating in the research, to ensure the protection of the subjects' well-being and abstain from harming them. When publishing research results it is important that any and all authors be listed, but that only those be included who meet the agreed-upon criteria for authorship. The scientist deems it important that the results of their research find a socially valuable application.

Where do these values and principles come from?

What is the normative basis on which these requirements stand? They stem from the goals of science and our expectations of members of the scientific community. Science is both a social institution (social practice) and a profession (vocation) (Resnik, 1998: 35). As a social institution science is a social arrangement that encompasses activities in which different parties (research institutes, funding agencies, publishers, ethics committees) fulfil their distinctive roles, often acting cooperatively for the achievement of common purposes (across disciplines, cultures and national boundaries). Indeed, science has many and various goals: epistemological goals such as knowledge, trust and explanation, as well as practical ones, such as prediction, power and control (Resnik, 2007: 51). Science is also the basis of the world picture, supporting both human self-understanding and comprehension of the surrounding biological, cultural, social and technological environment. It is part of culture and it creates a foundation for education.

To accomplish these goals, science must follow epistemological as well as moral and social norms that provide guidelines for the good conduct of research. Science is based on trust and cooperation: researchers must be able to trust one another; their results must be trustworthy, and society must be able to trust scientists. Unethical actions destroy trust, and thus it is important for researchers to act honestly, be objective, assume a caring attitude toward their subjects and the environment, respect the autonomy and privacy of human subjects, be fair and responsible, and prevent possible harmful consequences from their work and its applications.

A certain tension prevails in the mutual relationships between scientists and society. Scientists expect appropriate conditions be created and that they be given the means to conduct research; however, society increasingly wants to have a say in the content and financing of research, to participate in decisions as to how much and according to what criteria research funding is distributed. Researchers consider it important to be free to choose their own research topics; then again, society may expect them to direct their investigation towards rapidly solving problems facing humanity or toward financial gain. Vaccines are a good example: the ground for their creation was decades of basic research, which could be applied immediately when a pandemic hits. Unfortunately, the importance of basic science remains unclear to many people (including policymakers) since they do not understand why one should study the nature and mechanisms of phenomena without overt and proximate applications.

Science is not only an institution, but also a profession or vocation with its own values and vocational standards, which are inscribed in ethics codes. Representatives of the profession are entrusted with the means to fulfil their tasks and societally important interests; they are given many rights along with associated responsibilities. On the one hand representatives of the profession are presumed to have specialised skills in their field; on the other, they are expected to pursue certain values and principles of professionalism that they must express in all that they do. The correspondence between a scientist's values and behaviour is referred to as researcher's integrity. Scientists have various duties that can sometimes engender conflicts of expectation. A scientist can simultaneously be a university teacher, supervisor of graduate students, leader of a research group; hold administrative responsibilities, provide expertise in court or in a social debate; they can be an investor or the owner of a business. It is said that such conflicts are a danger that may lead to moral violations (Werhane and Doering, 1997: 174). For example, a successful scientist who receives many invitations to conferences may have scant time for thesis supervision and for leading their research group. The accumulation of duties and the time bind may lead to carelessness in checking scientific data, or even to the pursuit of self-plagiarism. In order that conflicts among duties not lead to unethical behaviour, the scientist must become conscious of this danger and try to keep their duties in balance; if possible, reducing their burden.

An important role in securing ethical research is also played by organisations where scientists work. Today research is carried out not only in relation to specific organisations, but in view of other stakeholders, such as think tanks, civil society organisations, lobbying agencies, public relations companies, bloggers, influencers and so on. This means that not only professional researchers are doing research, but the concept of citizen science must also be kept in mind (Iphofen and O'Mathúna, 2021). Most scientists work in research institutions that provide the conditions for carrying out research and the education of the next generation of researchers. As such, research institutions have their own goals, organisational values and standards that must be followed by all of the scientists they employ. The responsibility for creating the culture of integrity at the institutional level must be carried by the leaders of the institutions (for example, Forsberg et al, 2018; Boulter, 2020). Research is financed by foundations in the public sector, businesses or industry (including the military industry), all of which fulfil their goals and organisational values, and who can set requirements and stipulations that scientists must meet upon acceptance of financing. Also, both scientists and research institutions must take the responsibility of refusing financing that is at odds with the ethical values of society, of science, the specific research institution or the scientist's own ethical values.

Disagreements on the conceptual level of ethics

Behind this seeming consensus on what counts as ethical research and who is an ethical researcher, disagreements and contradicting positions are rife. There is no shared understanding of the nature of values, or how they differ from principles and virtues. Also, it is debated whether it is more important to develop *proper* virtues or to have the *right* values or principles. Depending on which normative ethics theory one prefers, one either focuses on virtues or principles. If one is a supporter of virtue ethics, one emphasises the moral character or virtues, whereas if one finds principle-based ethics to be a correct approach, one focuses on rules that are central to deontological ethics and rule-utilitarianism.

The strengths of the principle-based approach are seen as their usefulness in providing guidelines for activity in specific situations. However, critics argue that principles are often articulated negatively, they are too formal and they lack the internal impetus to follow them. The strength of the virtue-based approach is that the person who has developed proper virtues not only has an understanding of what is morally good or right but is also habituated to act accordingly. If one has already developed proper character traits one can be relied upon to act appropriately. The weakness of this approach is that virtue ethics does not give clear guidance on what to do if in a given situation one has to choose which virtue is more important than another. In Aristotle's account of virtue ethics one needs special knowledge or understanding – practical wisdom (*phronesis*) to know what one has to do in a given situation. David Resnik (2012) claims that virtue-based and principle-based approaches complement each other: while virtue-based approaches can be useful for leadership, mentoring and instruction, principle-based approaches might be preferred for policy development and enforcement.

A similar conclusion was reached by normative analysis of research integrity and misconduct carried out in the EU-funded PRINTEGER project (Promoting Integrity as an Integral Dimension of Excellence in Research), showing that principles and virtues highlight important aspects of moral life and should therefore be viewed as complementary. In the project deliverable we argued: 'While it is important to live up to principles, this also requires the cultivation of character and proper motivation. While principles serve as good guides to cultivate professional virtues, being virtuous helps to make sure that principles are followed in a reflective and wholehearted manner' (Meriste et al, 2016: 37). Even many deontologists today agree that this kind of knowledge is needed since correct application of action-guiding rules requires situational appreciation – the capacity to recognise what is required in a particular situation.

What is the difference between virtues, values, norms and principles?

Virtue is an old term that goes back to antiquity where virtue ethics originates. Virtues were understood as character traits that enabled the possessor to live a good life. For ancient philosophers a good life was equated with a morally good life. In contemporary virtue ethics, virtues are seen as character traits or habitual dispositions that lead to morally good actions. A virtuous person is a morally good, admirable person who acts and feels well. Virtues can be seen as a link between value and action as they contain an understanding of what is morally good and at the same time express the readiness to act accordingly.

In opposition to virtue, the word *value* is a relatively new scholarly term which 'originated in the late 17th century economy and migrated through philosophy in the 19th century to the social sciences and to ordinary language in the 20th century' (Meisch et al, 2011). In its most general meaning, value is that which is deemed worthy, that which is worth having, getting or doing, or that possesses some associated property or properties. Evaluative and normative dimensions of value are interlinked: for something to be good means nothing more than that there are reasons to respond to it in a favourable manner (Orsi, 2015: 11). Since values are objects or criteria of valuations, they give orientation, but they do not themselves lead to actions.

Norms, in contrast to values, provide instruction to specific groups of people. According to Hans Joas (2000: 14–15) norms are restrictive, while values are attractive. Values give us an idea of what we want, while norms tell us what to do in a specific context. With the help of a deontic operator norms give different guidance for action: norms may exclude action or allow action, forbid action, or express the right or duty to do something. A norm is a rule that can be of various types: epistemic, conventional, legal, social or moral. Moral norms are binding with respect to the group to whom they are addressed. Ethics in research formulates moral norms that provide guidance to scientists and other stakeholders involved in the research process. Moral norms can be formulated as standards, regulations or policies.

Principles are a specific category of norms, usually very high or the highest ranking. In contemporary practical ethics, the word 'principle' seems to have a very loose meaning: all sorts of rules, obligations and rights are called 'principles'. In the broad sense, principles are general norms of conduct that describe obligations, permissible actions or ideals of action. Principles can also be seen as points of departure for the development of specific norms of conduct (Beauchamp and Childress, 2009: 2).

The analysis and comparison of different concepts showed that the borderline between values, virtues, principles and norms is not so clear-cut. Virtues can be articulated simultaneously as values or principles. So, for example, honesty is a virtue (or character trait), a principle of action and also a value (something worth having). Values provide orientation but do not demand a specific course of action; they attract us and motivate our action. This difference between values and norms has been explained very well by Simon Meisch and Thomas Potthast: 'A value is not a rule in the sense of a practical, context-sensitive instruction that specifies situations, actions, agents, nor does it include deontic operators. The gap between value and action can be and has to be bridged by norms or virtues' (Meisch and Potthast, 2010: 14).

Due to the multiplicity of values, the question arises of interrelationships among them, that is, whether values can be (in)comparable and (in)commensurable. Value (in)commensurability entails two further issues: how moral values should be related to each other and what one should do in cases where they conflict (Viganò and Lombardi Vallauri, 2020: 87). For monists, there is one ultimate value; a single scale enables the agent to compare conflicting *ought*s and reduce different values to one ultimate value. So, for example, most utilitarian and Kantian theories are monist and identify, respectively, either happiness/utility or respect as the ultimate value. According to pluralists such an ultimate value does not exist. They argue that many values are characterised by incommensurability or incomparability. Some pluralists argue that it is possible to arrange values according to their relative importance, but the bases of prioritising them, arranging them in a hierarchy, are also multiple. So, as distinct from the monism that provides just a single scale, for pluralists there can be a multiplicity of measures used carrying out the evaluations. Since practical life requires making choices and solving value conflicts, this makes monism attractive. It has been argued that pluralism typically appeals to those who want to explain why moral conflicts arise, whereas monism appeals to those who seek to organise the precepts of morality without moral remainders (Brunning, 2019). This can also explain why policymakers tend to prefer monism.

Pluralism in research ethics and research integrity

Pluralism in research ethics

The most influential attempt to defend foundational pluralism in research ethics has been made by Tom L. Beauchamp and James F. Childress who, in their seminal book, *Principles of Biomedical Ethics* (1979), showed that moral principles are plural at the level of choice and that there is no single principle that subsumes all others. With the formulation of mid-term principles they wanted to avoid disagreements between ethical theories. Beauchamp has provided the following explanation of what they mean by principles: 'A principle is a regulative guideline stating conditions of the permissibility, obligatoriness, rightness, or aspirational quality of an action falling within the scope of the principle. If principles are adequately expressed, more particular or specific moral rules and judgements are supported by, though not deduced from, the principles' (Beauchamp, 2010a: 154).

The account of Beauchamp and Childress has been called 'principlism' because they grouped all important moral requirements under four general principles: (1) autonomy (a principle of respect for the decision-making capacities of autonomous persons); (2) non-maleficence (a principle of avoiding the causation of harm to others); (3) beneficence (a group of principles for providing benefits and balancing benefits against risks and costs); and (4) justice (a group of principles for distributing benefits, risks and costs fairly). These principles are firm obligations and can be set aside or be compromised only if they come into conflict. As none of these principles is overriding and they can be ranked differently depending on the specific context, they leave considerable room for judgement about individual cases and policies, as well as for negotiation and compromise (Beauchamp, 2010a: 154). These principles are drawn from common morality, which is understood as the set of universal norms shared by all reasonable persons who are committed to morality. *Principles of Biomedical Ethics* has become the 'bible' of bioethics and is currently in its eighth edition.

In parallel to the writing of *Principles of Biomedical Ethics*, Beauchamp was also involved in the writing of the Belmont Report, prepared by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979), which was meant to provide a moral framework for research with human subjects. Beauchamp has commented on the drafting of the *Belmont Report* that it is wrong to think that their book was based on the document, since by the time he joined the Commission, he had already drafted substantial parts of *Principles* (Beauchamp, 2010b: 8–10). He has also pointed to substantial differences in the approach taken by the Commission and the authors of the book. Although at first sight it seems that the difference can be summed up in that the Belmont Report came up with only three principles (respect for persons, beneficence and justice) and the book *Principles* is built on four principles (respect for autonomy, non-maleficence, beneficence and justice), in reality the *Principles* and *Belmont* developed substantially different moral visions.

This process of translating the general principles into action-guiding requirements made it clear that even if one agrees that persons have to be treated with respect, one can require this for several different reasons. Initially the National Commission, which had been formed in the aftermath of public outrage over the Tuskegee syphilis experiments and other questionable uses of humans in research, suggested the principle of respect for persons in order to avoid harm. For Beauchamp and Childress it was important to distinguish between three principles: respect for autonomy, non-maleficence and beneficence. In *Principles* the principle of respect for autonomy is related to the need to respect the decision-making capacity of humans, and not simply to protect them from abuse. In the book, the protection of those who lack decision-making capacity is covered by the principle of non-maleficence.

Years later, the German philosopher Theda Rehbock criticised Beauchamp and Childress for reducing autonomy to 'respect for autonomous choices' and argued that the principle of respect for autonomy should not be reduced to decision-making capacity (that is, 'competence') (Rehbock, 2011: 524).

Rather, autonomy has to be understood in a much broader sense as respect for the person's *will* (Rehbock, 2011: 526). Rehbock explained that autonomy in the Kantian sense of personal autonomy is unconditional and universal – one has to respect everyone's will, not only that of those able to give informed consent. Also, in Rehbock's view autonomy cannot be overridden even in the case of moral conflict.

The principlism of Beauchamp and Childress has been criticised by several other scholars. The most severe criticism has been provided by K. Danner Clouser and Bernard Gert, who were of opinion that the principles do not give enough normative guidance and that the approach in general lacks theoretical unity (Clouser and Gert, 1990).

Some other authors have doubted that principles identified by Beauchamp and Childress are universal. Basic ethical principles in European bioethics and biolaw were analysed in the framework of the BIOMED II project. The project identified four ethical principles: autonomy, dignity, integrity and vulnerability (Rendtorff, 2002). It has been argued that in view of different fields of biomedicine, the four principles should be applied in different ways (Rendtorff et al, 2000). Matti Häyry (2003) has argued that dignity, precaution and solidarity are not universal but reflect best the European ethos. Also, these values can be interpreted in various ways. For example, depending on whether one believes that dignity refers to God, reason, genes, sentient beings or some special social status, it may mean different things and lead to different norms (Häyry, 2004). Robert Baker has pointed out that the concept of solidarity is a European one as it stems from the cultural tradition of the French Revolution and 'the concept of *solidarité* has no resonance' in North America and much of Asia (Baker, 2005: 194).

In relation to scientists' professional conduct, a significantly longer list of principles has been proposed by Adil E. Shamoo and David B. Resnik, who outline 15 principles: honesty, objectivity, carefulness, credit, openness, confidentiality, collegial respect, honouring of intellectual property, freedom, protection of experimental animals, protection of human subjects, stewardship, respect for the law, professional responsibility and social responsibility (Shamoo and Resnik, 2015: 18–19).

Plurality of ethical frameworks

There is not only a plurality of principles but also of ethical frameworks. Liberalism stresses individual rights whereas communitarianism stresses the common good. The basic difference between liberal and communitarian ethical frameworks lies in the fact that liberalism claims that the individual is more important than the society in which they live, whereas communitarianism regards society as more important than the individual. Liberal individualism, with its conceptual base of autonomy, dignity and privacy, enjoyed a long period of dominance in research ethics, but it has increasingly come under attack from ideologies promoting a more salient role for concepts of solidarity, community and public interest and one speaks about the 'communitarian turn' in bioethics. Since the turn of the millennium there have also been attempts to balance individual rights against the common good.

However, mainstream research ethics continues to hold the line of privileging the rights of the individual over the interests of science and society, and reinforcing such privilege in ethical regulations (Sutrop, 2011a, 2011b). In keeping with the WMA (1964/2013) Declaration of Helsinki and the Oviedo Convention (Council of Europe, 1997), as well as the UNESCO Universal Declaration on Bioethics and Human Rights it has been stated: 'The interests and the welfare of the individual should have priority over the sole interest of science or society' (UNESCO, 2005). However, it is not clear if this position is indeed shared by all cultures where society is ranked more highly than the individual. Although many leading ethicists (Macklin, 2005; ten Have, 2005, 2006; Andorno, 2007) have acknowledged the importance of the UNESCO Declaration, this document has also provoked quite extensive critical debate (Baker, 2005; Häyry and Takala, 2005; Kopelman, 2005; Selgelid, 2005; Williams, 2005; Wolinsky, 2006; Eriksson et al, 2008, among others).

One of the critical points is whether human rights discourse is suitable for providing a universal platform for global norms. For example, Asian bioethicists have pointed out that cultural differences should be included in formulating research norms. Several bioethicists have pointed out that for cultures that have a societal and family orientation, where the family may indeed be the highest value and the core of all social, cultural and economic activity, an individualistic approach with the focus on autonomy may be unacceptable. In several other cultures, values such as beneficence or solidarity, and emotional security may be ranked more highly.

Ren-Zong Qiu, the editor of the volume *Bioethics: Asian Perspectives – A Quest for Moral Diversity* argues in the preface that:

> Confucian or Buddhist cultures may share with Judaeo-Christian cultures such values or rules as 'Do not kill the innocent' … However, for Confucians, the rule 'Do not kill the innocent' does not include a prohibition of killing fetuses, whereas for Buddhists the rule extends to not killing all forms of animals. (Qiu, 2004: 1)

His conclusion is that, 'Although such shared values are useful in practice, their content is too poor to constitute an overarching universal ethics or global bioethics' (Qiu, 2004: 1). Thus, the search for universal principles has come under continuous attack from different points of view. First, it is seen by dominant Western ethical frameworks as an attempt at moral colonisation.

Second, it has been argued by several authors that there is only a seeming consensus on principles, hiding the fact that moral and political dilemmas continue to exist. There may be agreement about which principles should be respected but since there is no agreement on what these principles mean, they cannot be applied as general guidelines (Häyry and Takala, 2005; Plomer, 2005). Third, there are also problems with the balancing of the principles, should they come into conflict (Holm, 1995).

Pluralism in research integrity

Ethical codes and guidelines for researchers encompass very different values. None of these values is absolute, given that values can come into conflict; researchers must decide which of the conflicting values should rule in the given situation. Conflicts may ensue between freedom and responsibility, privacy and safety, openness and loyalty.

There is also no shared understanding of which virtues scientists should develop. Edmund Pellegrino (1992: 3) has emphasised that the virtues important for a researcher are objectivity, critical thinking, honesty with respect to data, freedom from prejudice and sharing knowledge with the scientific community.

Approaches based on virtue ethics stress the importance of the scientist as a human being. Bruce Macfarlane indicates that the distinction between being a scientist and playing other roles is artificial (Macfarlane, 2009: 45). Ron Iphofen (2011: 4) emphasises that it is difficult to believe that someone could be a good person and not equally 'good' as a scientist, because there is no way to distinguish between how someone behaves personally in their off-hours, for example at home, and the standards to which they should be expected to adhere as a professional.

Macfarlane has skilfully articulated what it means to be an ethical person:

> To be ethical means to create a deep and personal understanding of one's values. … The best solutions to ethical problems can only come through practice, experience, learning from the good and bad examples of others. Being an ethical researcher presumes an authentic bond with one's own convictions and the basic values of one's area of specialisation. In this sense ethics resembles jazz, which is something more than following notes on a page. It calls for the skill of improvisation and independent thinking. (Macfarlane, 2010: A30)

It has been argued that there is:

> irreducible *value pluralism* in research integrity: there is a variety of different kinds of values underlying codes of conduct that cannot be

> reduced to each other and that give rise to a plurality of virtues and norms, some of which are either incommensurable or in conflict with each other. (Peels et al 2019: 1; original emphasis)

We agree with Peels et al (2019) that there is a plurality of values and that consequently should lead to a plurality of norms.

In our PRO-RES project we analysed existing codes and guidelines, identifying 24 values and 27 issues. One of the main points that emerged from the analysis of the documents was that there is 'conceptual ambiguity and terminological inconsistency' (Parder and Juurik, 2019: 3), meaning that different concepts have been used to express similar things. For example, to address external influences and possible biases in decisions, 'impartiality', 'independence', 'objectivity' and 'fairness' are used; to stress treating others fairly, notions such as 'justice', 'equality' and 'equity' are used. Another finding from the codes and guidelines analysed was that no clear distinction is made between values as virtues or as principles for good conduct. For example honesty, accuracy, rigour can be interpreted in two ways: the presumed character traits of the researcher and the standards the researcher needs to follow.

A plurality of understandings can even be found at the level of a single notion. For example, we can extract autonomy from the aforementioned PRO-RES analysis on codes and guidelines. Parder and Juurik (2019) show that although autonomy is most often related to the perspective of research subjects, it has also been ascribed to individual researchers, research organisations and research in general. Most often, free, informed consent has been seen as the means to respect autonomy. However, this analysis shows that respect for autonomy may also mean: 'advance consideration of factors and constraints that may diminish personal autonomy'; 'minimizing or avoiding possible risks and harm to personal autonomy'; 'use of additional measures to protect the interests of individuals with diminished autonomy', 'respecting the autonomy of groups and communities' and 'minimising the use of proxies' (Parder and Juurik, 2019: 13–14). At the level of researchers and research in general the focus is on the following: 'responsibility as a part of professional autonomy'; 'a tension between autonomy and public accountability'; 'a possible tension between *autonomy* and financing sources' (Parder and Juurik, 2019: 14).

The findings of the PRO-RES analysis are aligned with what Peels et al (2019) have designated *value pluralism* in research integrity. They have pointed out that codes of conduct for research contain both metaphysical and axiological pluralism. Metaphysical pluralism means that codes include values, norms and virtues. Axiological pluralism emerges when we notice that there are different categories of values, norms and virtues: epistemic, moral, professional, social and legal. Opponents argue that norms from various categories can come into conflict or there can be conflicts within the same category of norms.

Horbach and Halffman have marked it as a problem that the approaches of scientists and policymakers to research integrity go in different directions: ‘Scientists tend to present integrity as a virtue that must be kindled, while policy documents and newspapers stress norm enforcement’ (Horbach and Halffman, 2017: 1461). The problem with having too many and too diverse views on research integrity is that this makes it difficult to agree on what is meant by ‘good science’ and ethical research and what the implications may be for relevant stakeholders.

How can the plurality of values be maintained at the level of norms?

We have seen that there is a considerable plurality of values and value systems in both research ethics and research integrity. Since norms are translations of values into action guides, moral norms for research should also be regarded as plural. Currently policymakers talk about the need to bring governance more into line with the values of citizens. This is considered important in order to increase public trust and acceptance of policy decisions. However, value-based governance implies that one takes into consideration the fact that there is a plurality of values, values may conflict, and that various ethical frameworks provide different solutions.

Unfortunately, this pluralism is often ignored when developing global regulations or policy recommendations. For example, the recent EU *Ethics Guidelines on Trustworthy AI* (European Commission, 2019) employ an individual rights-based approach, ignoring the fact that since the turn of the millennium, after a long period of dominance, liberal individualism, conceptually based in autonomy, dignity and privacy have come under attack from communitarian ideologies, which promote a more salient role for concepts of solidarity, community and public interest. The *Guidelines* list four principles: respect for human autonomy, prevention of harm, fairness, and explicability (European Commission, 2019: 12), but they do not show where these principles come from. There is also some inconsistency: on the one hand the *Guidelines* stress that AI systems should respect the plurality of values and choices of individuals (European Commission, 2019: 11); on the other, they claim that certain fundamental rights and principles, such as human dignity, are absolute and cannot be subject to a balancing exercise (European Commission, 2019: 13). Also, the *Guidelines* are not helpful in advising what should be done if the principles conflict (Sutrop, 2019: 51).

It has been pointed out that it is also a problem that most of the ethics guidelines for AI have been developed by countries with liberal democracies. Again, this raises the question of whether the plurality of values is taken into account.

In research ethics the initial policy was to make all scientists follow principles and norms worked out in one field of research and automatically carried over to others. For example, the requirement of informed consent, initially drawn up in the World Medical Association Helsinki Declaration as a statement of ethical principles for medical research involving human subjects, has been made a central element of ethics review in all fields of research. Two difficulties have emerged: first, a principle adopted in one area of enquiry cannot be extended in a literal sense to other areas of science; second, one should also take contextual factors into account. Furthermore, in some countries where community and family play a significant role, there have been protests that the principle of informed consent undermines the importance of the community. How community consent is related to individual consent remains an open issue:

> There has been a growing appreciation of the importance of community leaders and families in the context of decision-making. While the process of going through such community gatekeepers does not take away from the importance of the individual's understanding of and willingness to participate in the research, it adds an element of security in traditional societies where communal consciousness and living is the norm. (Bhutta, 2004: 774)

Most critics believe that community consent should not substitute individual consent but rather complement it.

In some other cases, for example when doing research on discriminated groups or in totalitarian regimes, collecting and archiving research subjects' written informed consent may put them at risk. Also, the possibility of obtaining informed consent may depend on the research methodology. For example, in psychological experiments one cannot inform research participants about the set-up of the experiment, because then they would change their behaviour and the experiment would fail.

Even in biomedical research, where the principle of informed consent was first adopted, it has recently become clear that new areas of scholarly enquiry and changing research contexts call for the modification or reformulation of the concept of informed consent. New forms of consent (open, broad, dynamic or meta-consent) have been proposed so as to better meet the needs of research and protect important values – non-maleficence, liberty, personal autonomy, human dignity and trust (Sutrop and Lõuk, 2020). So, the best way to avoid one-size-fits-all solutions is to begin with reflecting on what values need protection in specific research context. Instead of importing requirements from one field of research to another, one should first carry out a proper ethical analysis with the aim of identifying risks and carefully considering what kind of ethical issues are raised by the proposed methodology and what might be the potential social and ethical implications of research.

In the field of research integrity there has been a discussion about whether it would be better to have one global code of conduct for all scientists instead of having numerous codes. There is a boom of creating new codes of conduct that can be categorised either according to scope or discipline. Although narrow discipline-based approaches are not very common, there are still some examples (for example the American Psychological Association's code of conduct: APA, 2017).

There is greater variety of codes based on scope. First, distinctions can be made as to whom the codes are addressed – are they intended for scientists, research institutions (such as the Bonn PRINTEGER Statement (Forsberg et al, 2018); the Code of Practice for Research at the Open University (2017)) or various stakeholders (for example 'Cooperation Between Research Institutions and Journals on Research Integrity Cases: Guidance from the Committee on Publication Ethics' [Wager and Kleinert, 2012]). Recently, there has been a tendency to develop codes which combine the levels of a scientist and a research institution (examples include the Danish Code of Conduct for Research Integrity [Danish Ministry of Higher Education and Science, 2014], the Estonian Code of Conduct for Research Integrity [Centre for Ethics University of Tartu, and Estonian Research Council, 2017]). Finally, the criteria can be established at the level where they should work, either at the level of the professional society or association (for example the *Code of Conduct* of the Market Research Society, 2019, or *Code of Ethics and Conduct* by the British Psychological Society, 2018), at a national level (for example the Netherlands Code of Conduct for Academic Practice [VSNU, 2014], Australian Code for the Responsible Conduct of Research), at a regional level (European Code of Conduct for Research Integrity [ALLEA, 2017]) or a global level (such as initiatives from the World Conference on Research Integrity [WCRI, 2010, 2013, 2019]).

The first WCRI in 2005 has been described as 'an initial attempt at providing World awareness of research integrity and discussing strategies for harmonising policies and fostering responsible conduct in research' (Mayer and Steneck 2012: v). The WCRI has by now grown into a global effort. Over the last 14 years, six international conferences have been held, where global guidance has been offered (these include the Singapore Statement in 2010 [WCRI, 2010]; the Montreal Statement in 2013 [WCRI, 2013]; and Hong Kong principles in 2019 [WCRI, 2019]).

A global code is an attractive option, since it seemingly establishes universal standards for ethical research but it must remain very general and therefore cannot cover all topics in a detailed way. Thus, it leaves plenty of room for discipline-specific or institutional codes. A universal code of conduct for all scientists cannot take into account all interdisciplinary and intercultural differences. Also, a unity of values cannot be achieved by adopting one general code of conduct. However, any attempt to formulate universal values and principles that should be followed by all scientists may promote discussion which may in turn help to get closer to agreement.

Although there exists a descriptive multiplicity of values, this does not yet mean that all values that people hold are good or just. We can acknowledge that there has also been a development in understanding of what is considered ethically permissible in Western countries. Ruth Macklin (1999) has claimed that moral progress is evidenced in the adoption of two principles: the principles of humanity and humaneness. Of course, self-critical reflection is also necessary so that we do not find ourselves holding values and principles only because they are our own, but we should rather attempt to evaluate them from the standpoint of an unbiased observer. By means of reflection and education we may attempt to convince people of the importance of universal values.

To sum up, on the one hand, growing interest in WCRIs and the adoption of global statements at these conferences has shown that scientific misconduct is a shared concern and that there is a willingness to develop global norms and standards for good research. On the other hand, it is not clear whether the initiative of WRCI organisers to formulate global codes of conduct has been similarly welcomed by individual researchers in different fields of research working in different parts of the world. As already noted, in order to have motivational force, norms and regulations should correspond to scientists' values. Lack of such correspondence may indicate either lack of research integrity or even open the door to misconduct: a scientist may show active resistance against regulations and tend to violate norms that differ from their personal ones. Since the codes adopted at the national level show a great variety of values, it can be assumed that not all scientists share the values expressed in these global statements. As long as there is a prevailing plurality of values and value systems, and people rank values differently, codes of conduct for scientists must also reflect such differences.

Conclusion

The task of research ethics and research integrity is to make researchers think about the character traits (virtues) required of a scientist and the values and principles that should be observed so that science can work. The tick-the-right-box approach diverts attention from the reasons why such requirements are necessary and why certain significant values are to be protected.

It is important that research ethics and research integrity frameworks be built upon trust, one of the key themes of this volume. Scientists should not be treated only as possible violators, but as professionals cognisant of ethical principles who deem it important to adhere to professional ethics. Without trust and cooperation, doing science is impossible.

At the same time one should not underestimate risks nor trust scientists blindly, since numerous violations indicate that there are those who either do not know about ethical requirements or who do not wish to fulfil them. Policymakers must take this into account as they determine which research they rely on. The question is in the proper placement of trust, an appropriate assessment of the trustworthiness of scientists and the risks of violation. Control and supervision of ethical behaviour should not be excessive, but it should rather stress the sense of responsibility of scientists and research institutions.

It is increasingly being understood that it is insufficient for scientists to have the right values, know the ethical requirements and wish to follow them: all of these should find support in the organisational culture of the institution for which they work. Leaders of institutions are responsible for shaping integrity culture, and they must guarantee training in how to do science ethically. Likewise, it is important that there be incentives to behave ethically (that is, that the quality rather than the quantity of publications be assessed and that positive and negative sanctions support ethical behaviour). Policymakers can contribute by ensuring that policies promote the highest ethical standards in research. Here we can also note the responsibility of those who shape research politics in terms of how institutions are assessed (evaluation, rankings) and the responsibility of those who finance research.

Acknowledgements

This paper grew out of research undertaken for the European Commission Horizon 2020 project 'PRO-RES – PROmoting integrity in the use of RESearch results in evidence-based policy: a focus on non-medical research'. We also acknowledge the support provided by the Centre for Excellence in Estonian Studies (European Union, European Regional Development Fund). Due to the DAAD Scholarship Margit Sutrop was able to work on this paper during her stay at the International Centre for Ethics in Sciences at the University of Tübingen. We would also like to thank Prof. Thomas Potthast (University of Tübingen) for his guidance on the conceptual analysis of values, principles, virtues and norms and our colleagues Dr Mari-Liisa Parder, Heidy Meriste and Marten Juurik at the Centre for Ethics at the University of Tartu for helpful discussions and feedback. We are very grateful for the editor Dr Ron Iphofen for his support and guidance. We would also like to thank Prof. Tiina Kirss for her careful editing and help with English expression.

References

ALLEA (All European Academies) (2017) *The European Code of Conduct for Research Integrity*, available from: https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf [accessed 12 August 2021].

Anderson, M., Shaw, M.A., Steneck, N.H., Konkle, E. and Kamata, T. (2013) 'Research integrity and misconduct in academic profession', in M. Paulsen (ed) *Higher Education: Handbook of Theory and Research 28*, Dordrecht: Springer, pp 217–61.

Andorno, R. (2007) 'Global bioethics at UNESCO: in defence of the Universal Declaration on Bioethics and Human Rights', *Journal of Medical Ethics*, 33(3): 150–4.

APA (American Psychological Association) (2017) *Ethical Principles of Psychologists and Code of Conduct*, available from: https://www.apa.org/ethics/code [accessed 12 August 2021].

Baker, R. (2005) 'International bioethics and human rights: reflections on a proposed Universal Declaration on Bioethics and Human Rights', *Politics and Ethics Review*, 1(2): 188–96.

Banks, S. (2015) 'From research integrity to researcher integrity: issues of conduct, competence and commitment', Virtue Ethics in the Practice and Review of Social Science Research, 1 May 2015, Academy of Social Sciences, available from: https://web.archive.org/web/20201108160147/https://www.acss.org.uk/wp-content/uploads/2015/03/Banks-From-research-integrity-to-researcher-integrity-AcSS-BSA-Virtue-ethics-1st-May-2015.pdf [accessed 12 August 2021].

Beauchamp, T.L. (2010a) 'Principles and other emerging paradigms in bioethics', in T.L. Beauchamp (ed) *Standing on Principles: Collected Essays*, Oxford: Oxford University Press, pp 153–74.

Beauchamp, T.L. (2010b) 'The origins and evolution of the *Belmont Report*', in T.L. Beauchamp (ed) *Standing on Principles: Collected Essays*, Oxford: Oxford University Press, pp 3–17.

Beauchamp, T.L. and Childress, J.F. (1979) *Principles of Biomedical Ethics*, New York: Oxford University Press.

Beauchamp, T.L. and Childress, J.F. (2009) *Principles of Biomedical Ethics* (6th edn), New York: Oxford University Press.

Bhutta, Z.A. (2004) 'Beyond informed consent', *Bulletin of the World Health Organization*, 82(10): 771–7.

Boulter, L. (2020) 'What research institutions can do to foster research integrity', *Science and Engineering Ethics*, 26(4): 2363–9.

Braun, R., Ravn, T. and Frankus, E. (2020) 'What constitutes expertise in research ethics and integrity?', *Research Ethics*, 16(1/2): 1–16.

Briggle, A. and Mitcham, C. (2012) *Ethics and Science: An Introduction*, Cambridge: Cambridge University Press.

British Psychological Society (2018) *Code of Ethics and Conduct*, available from: https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct [accessed 12 August 2021].

Brunning, L. (2019) 'The avoidance approach to plural value', *Theoria*, 66(160): 53–70.

Buljan, I., Pina, D.G. and Marušić, A. (2021) 'Ethics issues identified by applicants and ethics experts in Horizon 2020 grant proposals', *F1000Research*, 10: art 471, available from: https://f1000research.com/articles/10-471 [accessed 12 August 2021].

Centre for Ethics University of Tartu, and Estonian Research Council (2017) *Estonian Code of Conduct for Research Integrity*, available from: https://www.eetika.ee/en/ethics-estonia/estonian-code-conduct-research-integrity [accessed 12 August 2021].

Clouser, K.D and Gert, B. (1990) 'A critique of principlism', *Journal of Medicine and Philosophy*, 15(2): 219–36.

Code of Practice for Research at the Open University (2017) Available from: http://www.open.ac.uk/research/sites/www.open.ac.uk.research/files/files/ecms/research-pr/web-content/Code-of-Practice-for-Research-at-The-Open-University-FINAL-for-the-external-research-website-July-2017.pdf [accessed 2 March 2022].

Council of Europe (1997) *Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine*, available from: https://rm.coe.int/168007cf98 [accessed 12 August 2021].

Dahal, B. (2020) 'Research ethics: a perspective of South Asian context', *Edukacja*, 1(152): 9–20.

Danish Ministry of Higher Education and Science (2014) Danish Code of Conduct for Research Integrity, available from: https://ufm.dk/en/publications/2014/the-danish-code-of-conduct-for-research-integrity [accessed 12 August 2021].

Davis, S.R. (2019) 'An ethics of the system: talking to scientists about research integrity', *Science and Engineering Ethics*, 25(4): 1235–53.

De Albuquerque Rocha, K. and Vasconcelos, S.M.R. (2019) 'Compliance with National Ethics Requirements for human-subject research in non-biomedical sciences in Brazil: a changing culture?', *Science and Engineering Ethics*, 25(3): 693–705.

Emanuel, E., Crouch, R.A., Arras, J.D., Moreno, J.D. and Grady, C. (eds) (2011) *Ethical and Regulatory Aspects of Clinical Research: Readings and Commentary*, Baltimore, MD: Johns Hopkins University Press.

Eriksson, S., Höglund, A.T. and Helgesson, G. (2008) 'Do ethical guidelines give guidance? A critical examination of eight ethics regulation', *Cambridge Quarterly of Healthcare Ethics*, 17(1): 15–29.

European Commission (2019) *Ethics Guidelines for Trustworthy AI*, available from: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 [accessed 12 August 2021].

Forsberg, E.M., Anthun, F.O. et al (2018) 'Working with research integrity – guidance for research performing organisations: the Bonn PRINTEGER Statement', *Science and Engineering Ethics*, 24(4): 1023–34.

Grant, C. and Gazdula, J. (2020) 'Research ethics in sub-Saharan Africa: graduate business students' attitudes and beliefs', *Journal of Ethnographic & Qualitative Research*, 14(1): 260–73.

Haggerty, K.D. (2004) 'Ethics creep: governing social science research in the name of ethics', *Qualitative Sociology*, 27(4): 391–414.

Häyry, M. (2003) 'European values in bioethics: why, what and how to be used?', *Theoretical Medicine*, 24(3): 199–214.

Häyry, M. (2004) 'Another look at dignity', *Cambridge Quarterly of Healthcare Ethics*, 13(1): 7-14.

Häyry, M. and Takala, T. (2005) 'Human dignity, bioethics, and human rights', *Developing World Bioethics*, 5(3): 225–33.

Holm, S. (1995) 'Not just autonomy: the principle of American biomedical ethics', *Journal of Medical Ethics*, 21(6): 332–8.

Horbach, S.P.J.M. and Halffman, W. (2017) 'Promoting virtue or punishing fraud: mapping contrasts in the language of "scientific integrity"', *Science and Engineering Ethics*, 23(6): 1461–85.

Institute of Medicine National Research Council of the National Academies (2002) *Integrity in Scientific Research: Creating an Environment that Promotes Responsible Conduct*, Washington, DC: National Academy of Sciences.

Iphofen, R. (2011) *Ethical Decision-Making in Social Research: A Practical Guide*, London: Palgrave Macmillan.

Iphofen, R. (2020) 'Regulating research: governance and ethics across the spectrum', in R. Iphofen (ed) *Handbook of Research Ethics and Integrity*, pp 17–32.

Israel, M. (2015) *Research Ethics and Integrity for Social Scientists* (2nd edn), London: SAGE.

Joas, H. (2000) *The Genesis of Values*, Chicago: University of Chicago Press.

Jordan, S.R. and Gray, P.W. (2013) 'Research integrity in greater China: surveying regulations, perceptions and knowledge of research integrity from a Hong Kong perspective', *Developing World Bioethics*, 13(3): 125–37.

Kemp, P., Rendtorff, J. and Johansen, M.N. (eds) (2000) *Bioethics and Biolaw: Four Ethical Principles* (vol. 2), Copenhagen: Rhodos International Science and Art Publishers & Centre for Ethics and Law.

Kopelman, L.M. (2005) 'The incompatibility of the United Nations' goals and conventionalist ethical relativism', *Developing World Bioethics*, 5(3): 234–43.

Macfarlane, B. (2009) *Researching with Integrity: The Ethics of Academic Enquiry*, New York: Routledge.

Macfarlane, B. (2010) 'Researching with integrity', *Chronicle of Higher Education*, 56(3): A30.

Macklin, R. (1999) *Against Relativism: Cultural Diversity and the Search for Ethical Universals in Medicine*, New York: Oxford University Press.

Macklin, R. (2005) 'Yet another guideline? The UNESCO draft declaration', *Developing World Bioethics*, 5(3): 244–50.

Market Research Society (2019) *Code of Conduct*, October, available from: https://www.mrs.org.uk/pdf/MRS-Code-of-Conduct-2019.pdf [accessed 12 August 2021].

Mayer, T. and Steneck, N. (2012) 'Preface', in T. Mayer and N. Steneck (eds) *Promoting Research Integrity in a Global Environment*, Hackensack, NJ: World Scientific, pp v–vii.

Meisch, S. and Potthast, T. (2010) Values in Society EC Project: The Landscape and Isobars of European Values in Relation to Science and Technology, available from: https://www.researchgate.net/publication/324331369_Values_in_Society [accessed 12 August 2021].

Meisch, S., Beck, R. and Potthast, T. (2011) *Towards a Pragmatically Justified Theory of Values for Governance*, available from: https://www.researchgate.net/publication/324331427_Towards_a_pragmatically_justified_theory_of_values_for_governance [accessed 12 August 2021].

Meriste, H., Parder, M.L., Lõuk, K., Simm, K., Lilles-Heinsar, L., Veski, L., Soone, M., Juurik, M. and Sutrop, M. (2016) *Normative Analysis of* R*esearch* I*ntegrity and* M*isconduct,* PRINTEGER Deliverable D.2.3, available from: https://printeger.eu/wp-content/uploads/2016/10/D2.3.pdf [accessed 12 August 2021].

National Academy of Sciences (2002) *Integrity in Scientific Research: Creating an Environment That Promotes Respinsible Conduct*, Washington, DC: National Academy of Sciences.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1976) *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, available from: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html [accessed 12 August 2021].

Orsi, F. (2015) *Value Theory*, London: Bloomsbury Academic.

Parder, M.-L. and Juurik, M. (2019) *Reporting on Existing Codes and Guidelines*, Zenodo, https://doi.org/10.5281/zenodo.3560777 [accessed 12 August 2021].

Peels, R. de Ridder, J., Haven, T. and Bouter, L. (2019) 'Value pluralism in research integrity', *Research Integrity and Peer Review*, 4: art 18, available from: https://doi.org/10.1186/s41073-019-0076-4 [accessed 12 August 2021].

Pellegrino, E.D. (1992) 'Character and the ethical conduct of research', *Accountability in Research*, 2(1): 1–11.

Plomer, A. (2005) *The Law and Ethics of Medical Research: International Bioethics and Human Rights*, London: Cavendish.

Qiu, R.Z. (ed) (2004) *Bioethics: Asian Perspectives: A Quest for Moral Diversity*, Dordrecht: Springer.

Rehbock, T. (2011) 'Limits of autonomy in biomedical ethics? Conceptual clarifications', *Cambridge Quarterly of Healthcare Ethics*, 20(4): 524–32.

Rendtorff, J.D. (2002) 'Basic ethical principles in European bioethics and biolaw: autonomy, dignity, integrity and vulnerability – towards a foundation of bioethics and biolaw', *Medicine, Health Care and Philosophy*, 5(3): 235–44.

Rendtorff, J. D. and Kemp, P. (2000) *Basic Ethical Principles in European Bioethics and Biolaw, Vol I-II*. Barcelona and Copenhagen: Institut Borja di bioètica and Centre for Ethics and Law.

Resnik, D.B. (1998) *The Ethics of Science: An Introduction*, London: Routledge.

Resnik, D.B. (2007) *The Price of Truth: How Money Affects the Norms of Science*, Oxford: Oxford University Press.

Resnik, D.B. (2009) 'International standards for research integrity: an idea whose time has come?', *Accountability in Research*, 16(4): 218–28.

Resnik, D.B. (2012) 'Ethical virtues in scientific research', *Accountability in Research*, 19(6): 329–43.

Schrag, Z.M. (2010) *Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965–2009*, Baltimore, MD: Johns Hopkins University Press.

Schrag, Z.M. (2011) 'The case against ethics review in the social sciences', *Research Ethics*, 7(4): 120–31.

Selgelid, M.J. (2005) 'Universal norms and conflicting values', *Developing World Bioethics*, 5(3): 267–73.

Shamoo, A.E. and Resnik, D.B. (2015) *Responsible Conduct of Research* (3rd edn), Oxford: Oxford University Press.

Shaw, D. (2019) 'The quest for clarity in research integrity: a conceptual schema', *Science and Engineering Ethics*, 25(4): 1085–93.

Sutrop, M. (2011a) 'Changing ethical frameworks: from individual rights to the common good?', *Cambridge Quarterly of Healthcare Ethics*, 20(4): 533–45.

Sutrop, M. (2011b) 'Viewpoint: how to avoid a dichotomy between autonomy and beneficence – from liberalism to communitarianism and beyond', *Journal of Internal Medicine*, 269(4): 375–9.

Sutrop, M. (2019) 'Should we trust artificial intelligence?', *TRAMES*, 23(4): 499–522.

Sutrop, M. and Lõuk, K. (2020) 'Informed consent and ethical research', in R. Iphofen (ed), *Handbook of Research Ethics and Scientific Integrity*, Cham: Springer, pp 213–32.

Ten Have, H. (2005) 'The Universal Declaration on Bioethics and Human Rights', *International Social Science Journal*, 57(186): 745–53.

Ten Have, H. (2006) 'The activities of UNESCO in the area of ethics', *Kennedy Institute of Ethics Journal*, 16(4): 333–51.

UNESCO (2005) *Universal Declaration on Bioethics and Human Rights*, available from: http://portal.unesco.org/en/ev.php-URL_ID=31058&URL_DO=DO_TOPIC&URL_SECTION=201.html [accessed 12 August 2021].

Viganò, E. and Lombardi Vallauri, E. (2020) 'Moral and vocational dilemmas meet the common currency hypothesis: a contribution to value commensurability', *Review of Philosophy and Psychology*, 11(1): 83–102.

VSNU (Association of Universities in the Netherlands) (2014) *The Netherlands Code of Conduct for Academic Practice*, available from: https://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code%20of_Conduct_for_Academic_Practice_2004_ (version2014).pdf [accessed 12 August 2021].

Wager, E. and Kleinert, S. on behalf of COPE Council (2012) Cooperation Between Research Institutions and Journals on Research Integrity Cases: Guidance from the Committee on Publication Ethics (COPE), available from https://publicationethics.org/files/Research_institutions_guidelines_final_0_0.pdf [accessed 12 August 2021].

WCRI (World Conferences on Research Integrity) (2010) *Singapore Statement*, available from: https://wcrif.org/guidance/singapore-statement [accessed 12 August 2021].

WCRI (World Conferences on Research Integrity) (2013) *Montreal Statement*, available from: https://wcrif.org/documents/354-montreal-statement-english/file [accessed 12 August 2021].

WCRI (World Conferences on Research Integrity) (2019) *Hong Kong Principles*, available from: https://www.wcrif.org/guidance/hong-kong-principles [accessed 12 August 2021].

Werhane, P. and Doering, J. (1997) 'Conflicts of interest and conflicts of commitment', in D. Elliot and J.D. Stern (eds) *Research Ethics: A Reader*, London: University Press of New England, pp 165–89.

Williams, J.R. (2005) 'UNESCO's proposed declaration on bioethics and human rights: a bland compromise', *Developing World Bioethics*, 5(3): 210–15.

WMA (World Medical Association) (1964/2013) *Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects*, available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 12 August 2021].

Wolinsky, H. (2006) 'Bioethics for the world: UNESCO's Universal Declaration on Bioethics and Human Rights has far-reaching goals, and has met with widespread opposition', *EMBO Reports*, 7(4): 354–8.

Zhang, J.J. (2017) 'Research ethics and ethical research: some observations from the Global South', *Journal of Geography in Higher Education*, 41(1): 147–54.

2

One size fits all? The problems of offering ethical guidance to everyone

Paul Spicker

Introduction

The process of ethical review has become a routine aspect of academic studies, but there are serious concerns about how it is being done. A clutch of articles has complained that the standards being applied are inappropriate, restrictive and in some cases downright silly (see Center for Advanced Study, 2005; Tysome, 2006; Murphy and Dingwall, 2007; Hammersley, 2009; Boden et al, 2009; and Schrag, 2011). It is always difficult to generalise across a series of disparate arguments, but looking at these critiques as a whole, what they seem to have in common is a sense of controlled fury. Schrag puts it in these terms:

> The first thing to understand about the critique of ethics committees is that it is grounded in bitter, bitter experience. People who devote their lives to the study of others are often quite concerned with ethics, and when they learn that their universities maintain ethics committees, their first reaction is often eager cooperation. But that goodwill can evaporate quickly when a researcher loses an afternoon to online training that is obviously irrelevant to the ethical challenges she faces, or when a committee imposes reporting requirements or restrictions that make the work difficult or impossible. (Schrag, 2011: 122)

Part of the problem is that ethical review tends to be supplemented by other concerns, which have more to do with institutional governance than with ethics. But the core of the problem is that the ethical positions that are being adopted are often misjudged or misconceived. Hammersley argues that research ethics committees (also known as institutional review boards or IRBs) are applying ethics codes much too literally, in a way that verges on the unethical – unable to address principles in context, and failing to identify ethical conflicts (Hammersley, 2006). It is difficult to say how general this sort of thing is, because researchers unfortunately are likely to opt out of the process; there is no point working on a proposal that is sure to be rejected.

> The anecdotal litany of IRB zeal includes members of preliterate tribes being asked to sign consent forms; faculty members 'investigated' for writing about their classroom experiences years earlier without advance IRB approval; projects so delayed that students were unable to complete their degrees … Mission creep damages the entire compliance system, because researchers find IRB requirements to be overwhelming and sometimes illogical. One example of this is consent forms running 20 pages or more. They are so long and detailed that most subjects, it has been observed, sign without reading them. (Centre for Advanced Study, 2005: 5, 12)

That last is a shining example of what Corrigan calls 'empty ethics' (Corrigan, 2003) – the triumph of style over substance. Consent is supposed to protect the interests of respondents, but it cannot cover all the circumstances in which the information may be used. Researchers have a duty to protect participants regardless of any formal procedure, and written consent is no guarantee that the rights of a participant will be respected (ASA, 2001: 2).

Conventional ethical guidance, and where it has gone wrong

Conventional representations of research ethics have tended to follow the lead given by medical sciences. Much of the received wisdom in the discussion of research ethics begins with the Nuremberg Code (1947). The Nuremberg trials established ten principles, covering:

- *the importance of the scientific basis of experiments:*
    - 'the experiment should be such as to yield fruitful results for the good of society'
    - 'the experiment should be so designed and based on the results of animal experimentation and a knowledge of the natural history of the disease or other problem under study'
    - 'the degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem to be solved by the experiment'
    - 'the experiment should be conducted only by scientifically qualified persons …'
- *the duty of the scientist to protect subjects:*
    - 'the experiment should be so conducted as to avoid all unnecessary physical and mental suffering and injury'

  - 'no experiment should be conducted where there is an a priori reason to believe that death or disabling injury will occur'
  - 'proper preparations should be made and adequate facilities provided to protect the experimental subject'
  - 'during the course of the experiment the scientist in charge must be prepared to terminate the experiment'
- *the rights of human subjects to have a say:*
  - 'the voluntary consent of the human subject is absolutely essential'
  - 'during the course of the experiment the human subject should be at liberty to bring the experiment to an end'

The Nuremberg rules were a response to a particular set of abuses. They emphasise the process of research and the relationship of the researcher to people who participated in the research. They refer primarily to research 'experiments' – the word is found in nine of the ten principles. In an experiment, the researcher introduces some factor into a situation to see what effect it will have. That has an immediate ethical implication. If the researcher is setting out to do things to people, there is an obvious duty not to do bad things: *primum non nocere*. However, even within medical science, experimentation with human participants is only a small part of the range of research activity. Researchers can affect the position of people who are being studied even if that is not what they intend to do: that is the lesson of the 'Hawthorne effect' (Olson et al, 2004). But changing other people is not the primary purpose of most research, and even when it is an objective, the people most affected are not necessarily participants in the research process.

Most ethical codes in the social sciences make recommendations about:

- the impact of research, including:
  - the potential implications of research for participants
  - the potential implications of research for non-participants, and
  - the uses to which research can be put
- the way that participants are treated:
  - informed consent
  - confidentiality and anonymity, and
  - special consideration of vulnerable respondents
- disciplinary considerations, asking researchers to:
  - maintain research of high quality
  - display competence
  - act responsibly towards others in their field, and
  - advance their discipline
- the research relationship, including:
  - the responsibilities of the researcher to the body commissioning the research

- responsibilities to the host institution
- commitments to fellow researchers
- conflicts of interest, and
- integrity in dealing with participants and stakeholders.

The emphasis on the treatment of participants reflects the particular influence of the Nuremberg rules, and arguably the broader influence of medical sciences on other fields of activity. These rules relating to participants impinge in particular on the research process, and because of that they have been taken to apply in almost all research with human subjects. However, they have serious shortcomings. On one hand, they cannot engage with all the ethical issues that arise in the course of research; but on the other, they have been taken to impose limitations on the process of research of all kinds. Among the most prominent restrictions, from a large helping of inappropriate criteria, are rules concerning competence, voluntary participation and informed consent.

Competence

The Socio-Legal Studies Association suggests that it is unethical to do work which one is not competent to tackle. Members should not undertake work of a kind that they are not competent to carry out and should not ask sociolegal researchers under their supervision or guidance to carry out work which those researchers are not competent to carry out, or they themselves are not competent to supervise (SLSA, 2009: 2.2.1). If that was taken literally, learning how to research through practice would be illegitimate, and the demands of a PhD would probably be unachievable: no one can be sure that original insight is within their competence until they produce it. I would be far more worried about a student who was convinced that nothing was beyond their competence than one who had reasonable doubts.

Much of the point of doing research is to find out about things we don't know, and that always raises the possibility that things may be uncovered that we don't know how to interpret or respond to. The British Sociological Association treats competence as a matter of professional integrity: 'While recognising that training and skills are necessary for sociological practice – and, particularly for conduct of social research – sociologists should themselves recognise the boundaries of their professional competence. They should not accept work of a kind that they are not qualified to carry out' (The British Sociologial Association, 2017: 4).

This is dangerous ground. Incompetence, Kruger and Dunning argue, is commonly marked by a failure to understand what competence requires. People who are incompetent often don't realise it; the same skills and knowledge which would lead them to be self-critical are precisely the skills and knowledge they don't have. 'We propose that those with limited knowledge in a domain suffer a dual burden: not only do they reach mistaken conclusions and make regrettable errors, but their incompetence robs them of the ability to realize it' (Kruger and Dunning, 1999: 1121).

It seems, then, that we are demanding ethical compliance in situations when almost by definition people lack the insight or information to know that their actions might be unethical. That is a recipe for conflict and confusion – not all of it on the part of the researchers. There may be researchers who lack the insight to understand why other people are saying that their work is inadequate; but equally, there are ethical review committees where possibly no one has the specialist knowledge or disciplinary background to understand what the researcher is doing, and they put obstacles in the way of the work (Israel, 2004). Likewise, there are disciplinary researchers who do not understand the scope of alternative methodologies. For example, there are those who are ready to demand that other researchers should confine their research to an approved style of research: 'Graduate students in psychology are routinely taught the importance of delineating one's hypotheses in advance (i.e., prior to collecting data). Established researchers continue to regard it as questionable and possibly unethical to theorize after one's empirical results are known' (Baumeister and Leary, 1997: 313). It seems that by their lights the only competent way to conduct research is deductively. By that test, qualitative, inductive, intensive, abductive, empowering and practice-based research are all 'questionable and possibly unethical'. That says more about the limitations of those imposing these standards than it does about ethical methodology.

Voluntary participation

The UK Research Councils have told universities that 'Research participants must take part voluntarily, free from any coercion,' and that compliance with their recommendations is mandatory. So, the British Society of Criminology states that the people who are being researched must be able to 'Take part in research voluntarily, free from any concern and be able to give freely informed consent in all but exceptional circumstances (exceptional in this context relates to exceptional importance of the topic rather than difficulty of gaining access)' (BSC, 2015). That seems to say that a criminologist cannot report on a trial unless everyone has agreed to be part of both the trial process and the research.

There are different kinds of involuntary participation in research, and it may be helpful to distinguish disengaged and engaged forms. People can be part of research simply because they are being observed, or their data are being used, or they have a part in a situation (for example, someone involved in the operation of a law court or a person in a crowd). This is sometimes, not always helpfully, referred to as 'covert' research, which seems to imply that the researcher is doing something surreptitious; more precisely, it is research where there is 'limited disclosure' (NHMRC, 2018: 2.3), and the subject may not even know that research is going on. There tends to be a presumption that this is illegitimate, but that is far from clear; there are many circumstances where a researcher may not think it appropriate to engage with the involuntary participant, either because it will change the behaviour that the researcher legitimately wishes to examine, or because there is simply no need to do so. This is typically the case for work done in the public sphere – a broad term which stretches to behaviour in public, socially defined public acts (such as motoring), public affairs and published material.

More problematically, there are also circumstances where there may legitimately be involuntary engagement with the research process. In organisational research, individual respondents are constrained by their organisational roles, their contractual relationships with the organisation, and potentially by legal restrictions. This means that the participation and consent that is required is offered by the organisation, not from the individuals engaged with it – and comments and findings will routinely be reported to the organisation. In other cases, there is a strong case for organisations to be publicly accountable: the Canadian Tri-Boards say explicitly that in critical inquiry 'that is, the analysis of social structures or activities, public policies or other social phenomena', 'permission is not required from an institution, organization or other group in order to conduct research on them' (Government of Canada, 2018, article 3.6). The work done by government is, in general terms, fair game: many researchers would consider that as citizens they have a duty to scrutinise, publicise and hold the government to account. Some information relates to circumstances that the researcher (or any member of the public) could legitimately seek. Freedom of Information requests, made to government, are the right of every citizen. (One of my very first research projects was based on material that local authorities had a statutory obligation to publish and make available; and so my first step, in asking for material, was to find out whether or not the local authorities were complying with the law.) Government activity takes place in the public domain, and consequently it is open to public scrutiny. The idea that no one can find things out about people without asking them personally first is wildly overgeneralised, obstructive and, at times, improper. The examination, scrutiny and criticism of public information are fundamental to a free society.

Informed consent

Voluntary participation depends on consent, but beyond that there is a demand for consent to be 'informed', and that calls for more than acquiescence. Informed consent means that the participants should be told what the research is for, why it is being done and how the information will be used, and that calls for a process that is distinct from, and prior to, empirical research. A prominent example is the policy of the Massachusetts Institute of Technology's (MIT's) Committee on the Use of Human Subjects, which requires that nearly all studies involving human subjects must obtain informed consent, usually in writing (COUHES, 2019). The American Anthropological Association puts it in these terms:

> Minimally, informed consent includes sharing with potential participants the research goals, methods, funding sources or sponsors, expected outcomes, anticipated impacts of the research, and the rights and responsibilities of research participants. … Researchers must present to research participants the possible impacts of participation, and make clear that despite their best efforts, confidentiality may be compromised or outcomes may differ from those anticipated. These expectations apply to all field data, regardless of medium. (AAA, 2012)

Grugulis tells us what this means in practice:

> I spent last year conducting an ethnography of a computer games company, watching the way people learned skills and the way they were managed. No under-18s, no members of vulnerable groups, no illegal activities. Everyone was told who I was in advance by the company, both company and individuals would be anonymised in any publications and before observing anyone I would ask their permission. … Enter the ethics committee. They insisted on full written consent from every worker in the offices (about 250), every delivery person and – on the occasions I went off for a chat with informants – every barrista who served us coffee and waitress who brought us pizzas (no, seriously). (Grugulis, 2011)

This is yet another example of a set of rules that do not sit well with the process of research in practice. It is not just that the rule is impractical; what it asks people to do may be impossible. There are many circumstances in which an academic researcher cannot actually tell the participant how the information is going to be used. That happens in part because the information from any individual respondent or participant can usually be understood only in the context of information retrieved from other participants – that is how data analysis is done. 'Social science data', Boden and her colleagues write,

> is not simply 'out there' waiting to be harvested by researchers. Rather, it is constructed in the course of its collection and subsequent analysis. … Social research data is thus necessarily continuously constructed, defined

> and redefined throughout the entire research process and is framed by and produced through social relationships and ongoing processes of analysis and writing, into and through the literature. The requirement for a priori definition of 'data' therefore runs the risk that the processes of ethical bureaucracy will prohibit or inhibit the collection/construction of data as currently understood. (Boden et al, 2009: 738)

It may well be that the researcher does not know what is going to be used in research, and what is not. Whyte's *Street Corner Society*, a sociological classic based on participant observation, explains one of the main justifications of the approach: researchers learn answers to questions that it would not have occurred to them to ask (Whyte, 1955). Researchers may not even know that what they are doing is research. Any researcher's knowledge is based on a range of experiences. Like many people in my field, I was a practitioner before I was an academic. I have written about issues that were based on my own observation or practice – for example, pieces I have written about compulsory treatment for mental illness or the distress of people trying to get access to social housing. The people I have written about did not know I was going to write about them; nor, at the time, did I. I have anonymised the details, and altered some salient facts, so that the people I am writing about are not in any way identifiable; but there was no consent.

Exemptions and exceptions

There is a further set of rules that demands consideration. Faced with the construction of a clumsy and overgeneralised set of rules governing research activity, it is not uncommon for exceptions to be made. Committees that supervise research into healthcare know very well that the same kinds of consideration do not apply to protocols for pharmaceutical testing or invasive surgery as they do to population studies or evaluation of local services. I have had research 'nodded through' ethical review in the past – questionably, because it does not follow that if people are not being cut open or drugged, there are no more ethical issues to consider (see, for example, Spicker and Gordon, 1998: Ch. 11). The MIT procedures allow for requirements for informed consent to be waived if there is no more than a minimal risk of harm to subjects (COUHES, 2019). The US Department of Health and Human Services (DHHS) already exempts 'benign' behavioural research, whatever that may be, and has consulted about introducing an exemption for minimal risk into the 'common rule' applying to research in government activity (DHHS, 2018). It gives a blanket exemption to work relating to public officials, candidates for public office and public service programmes (DHHS, 1991). Beyond that, going further than most European researchers would advocate, it treats some ethically sensitive work as exempt from ethical examination.

> The following activities are deemed not to be research:
>
> 1. Scholarly and journalistic activities (such as oral history, journalism, biography, literary criticism, legal research and historical scholarship) …;
> 2. Public health surveillance activities …;
> 3. Collection and analysis of information for a criminal justice agency for activities authorized by law …
> 4. Authorized operational activities … in support of intelligence, homeland security, defense or other national security missions. (DHHS, 2018)

Having sat for some years on a university's ethical review board, this all rings alarm bells. The greatest risks of harm seem to me to stem from projects where the researcher is dismissive, or more usually not aware, of potential ethical problems. Exemptions allow this kind of work to pass under the bar. There is no legitimate ethical position that holds that ethical arguments should be deemed never to apply.

The points I have been discussing – competence, voluntary participation and informed consent – can all be read back to the Nuremberg Code. There are obvious reasons why people should not normally be subjected to invasive surgery if they have not agreed to it – that is not true in all circumstances, but it can be accepted as the default position for elective surgery at least. Medical interventions need to be done by competent professionals, subjects must be free to withdraw, everyone must consent to the procedure.

Having said that, the problems identified here are symptomatic of a deeper problem than the interpretation of the Nuremberg Code itself. In the process of constructing generalised rules, the way those rules have been expressed often seems to be at odds with the demands of research in practice. The problem is not that these principles are plainly and flatly wrong – they are right in some contexts and wrong in others – but they are wrong often enough to raise questions about the validity of codes that rely on them. The social scientists who have enshrined these principles in their guidelines might perhaps have found some cases where the rule seemed appropriate, but if so they have ignored the rest. I think it more likely that they have framed their guidance in these terms because expressing ethical principles is difficult, they found similar rules in other codes, and they thought these principles were broadly accepted. However, neither of those explanations would justify the codes in their present form.

Ethical research

Ethical conduct is typically judged in three ways. The first approach is consequentialist: actions are considered in the light of their likely or intended benefits, or potential harms. The desire to bring about good consequences is referred to as 'beneficence': the question of who benefits, and who is harmed, by the research. The RESPECT code, for example, suggests:

> It should be an overriding aim of socio-economic research that the results should benefit society, either directly or by generally improving human knowledge and understanding. It follows from this aim that in the conduct of the research, researchers should aim to avoid or minimise social harm to groups and individuals. With this in mind, socio-economic researchers and their funders should reflect on the consequences of participation in the research for all research subjects and stakeholders. (RESPECT Project, 2004)

The text of this guidance was developed through a process of negotiation by a group of people who came to it with different issues and agendas, and this brief passage segues across three different interpretations of beneficence. The first sentence is about beneficence in a general sense – the substantive benefits of research. The second sentence reduces this to a much less demanding criterion: 'non-maleficence', or doing no harm. (Even that may not be possible. Some research may be intended specifically to make its subjects worse off – for example, research on the effectiveness of punishment or economic disincentives.) The third sentence is about something different again: the position of people who participate in the research. It often happens that that the people participating in the research are not the subjects of it – a distinction that tends to be lost when well-meaning researchers complain about the use of the term 'subjects' – and, because scientific knowledge leans toward generalisation, it is only to be expected that the people who are most affected by research are not the participants.

The second approach is 'deontological' or principled: there are moral norms, codes and rules to follow, and conduct is ethical when it is consistent with those norms, and unethical when it is not. This has long been the primary paradigm in the governance of research ethics. But the staggering generality of the task should give us pause. Are there common rules, principles or values that can be applied across the range of all kinds of research? Some things are bound to be missed. No matter how good and well-constructed the material, a code of this sort cannot ever deal with the full range of ethical issues and problems that a researcher might encounter. The Social Research Association puts the point directly:

> no declaration could successfully impose a rigid set of rules to which social researchers everywhere should be expected to adhere, and this document does not attempt to do so. The aim of these guidelines is to enable the social researcher's individual ethical judgments and decisions

> to be informed by shared values and experience, rather than to be imposed by the profession. (SRA, 2003: 10)

Taking ethics seriously calls for a different approach to ethical rules. The British Psychological Society's (BPS's) professional guidance used to be highly prescriptive: for example, its 2009 code specifies on informed consent that psychologists should ensure that clients, particularly children and vulnerable adults, are given ample opportunity to understand the nature, purpose and anticipated consequences of any professional services or research participation, so that they may give informed consent to the extent that their capabilities allow. There were two pages on that topic alone. The BPS has radically reformed its approach in recent years – stripping down to core principles and indicating outlines rather than detailed rules (BPS, 2018). It proposes four central principles: respect, competence, responsibility and integrity. The principle of respect is explained in these terms:

> Respect for dignity recognises the inherent worth of all human beings, regardless of perceived or real differences in social status, ethnic origin, gender, capacities, or any other such group-based characteristics. This inherent worth means that all human beings are worthy of equal moral consideration.

Under this general heading, the BPS asks psychologists only to consider:

- Privacy and confidentiality;
- Respect;
- Communities and shared values within them;
- Impacts on the broader environment – living or otherwise;
- Issues of power;
- Consent;
- Self-determination;
- The importance of compassionate care, including empathy, sympathy, generosity, openness, distress tolerance, commitment and courage. (BPS, 2018: 3.1)

This does not try to tell people that there are rules and procedures that they must follow; rather, there are general principles that need to be thought about and taken into account. An ethical psychologist is asked to engage with ethical issues, to treat ethical consideration as a process, and to take a number of issues into consideration. The whole approach tends in the direction of the third main approach to ethical conduct: virtue ethics. It may not be possible to establish every principle, or to be aware of every possible consequence. What matters is that people try to do what is right, even if they get it wrong, and that they are alert to the issues. This is how we teach people to be moral; not by rules, not by weighing costs and benefits, but by trying to behave well.

Virtue ethics begins, not from general principles, but from the person and the context where ethical principles might apply. Lawton and his colleagues see it as a difference between compliance-based and integrity-based ethical systems. Integrity-based systems emphasise moral sensitivity, the exercise of moral reasoning or judgement, moral motivation (that is, the place accorded to moral values) and moral character (Lawton et al, 2012). People who want to behave ethically need to engage with ethics and values, to examine the dimensions of ethical problems and anticipate their implications. The approach is based on a recognition of the conditional character of ethical behaviour, asking people who want to behave ethically to think about their behaviour in context. Virtue is not a guarantee against unethical behaviour, but it means at least that researchers will try to do what is right. There may be different ways of promoting ethical research in practice, but the common factor in all the stages will not be the programme of research, which is liable to change and morph as new information emerges; it is the presence of the researcher, and it is on the conduct and approach of the researcher that responsiveness to ethical issues depends. It can be fostered by education and training, but responsiveness depends on continuing engagement in ethical reflection.

Rethinking ethical research

Ethical review processes tend to be based on the assumption that the research will be designed by a principal researcher, that review can take place prior to engagement, and that the researcher can then manage the ethical implications in accordance with the guidance offered by ethical codes and the judgement of the review panel. That only works for a limited class of research activity. We need to think about research ethics in context, and to do that we need to think differently about research.

The place to begin is with a simple question: what is research? We need to understand that research happens in all kinds of settings and circumstances, and it takes many different forms. The Common Rule applied by the US Federal Government defines research in these terms: 'systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge' (DHHS, 2018). That is wider than the conventional focus on research with human subjects, but it is still too narrow. Research is about more than the production of generalisable knowledge: investigations and evaluations are more usually particular rather than general. In simple terms, research is the task and process of finding things out. That might take in (among other things) scholarship, exploration, discovery, experimentation, analysis, disciplinary development, and practice. Sometimes researchers know what they want to investigate or examine;

sometimes they do not. Some research is concerned about relationships and generative mechanisms; some is about description and classification. Some research is carefully planned; some is serendipitous. Some research answers questions; some discovers the questions that might be asked later. Policymakers seeking 'ethical evidence' will often encounter many of these forms of research.

The second basic question that we need to consider: how do we know when research is happening? That question is not always obviously or straightforwardly answered. Long before anyone can come up with a 'research proposal', they need to find things out about the subject they are examining. My own work on public services has been built on a process of networking, discussion, engagement, preparing students for practice, visiting offices, asking questions and listening. In a nutshell, the process of preparing research itself begins with research. This matters, not just because it means that a process that focuses on formal research proposals cannot possibly cover the field, but because things can go wrong at any or every stage – the conception of the research, the development of aims, the construction of a plan of work, engagement with the subject matter, reporting and dissemination, and application. It follows that engagement with ethics needs to be continuous, and not confined to any single-stage process.

Third: who is a researcher? Many disciplinary codes, and processes such as ethical review committees, are intended to govern the behaviour of academic researchers. Research is done by many more people than that: journalists, investigators, lawyers, regulators, independent writers, practitioners, students and schoolchildren. If ethical conduct in research is to be identified by the kind of work being done, rather than by the person who is doing it, ethical principles have to apply to every possible group. This is reflected in how this volume addresses a broad range of practitioners, not just academic researchers. It makes no sense to say, for example, that a journalist or writer can legitimately produce a critical biography of an author, but an academic researcher cannot; or even that a child in primary school can be asked by their teacher to interview their grandparents, but that an academic who puts the same questions to the same people has to go through an ethical review process. If it's unethical for some, then, by whatever criteria we may apply, it's presumably unethical for others, too. Policymakers should be concerned about the ethics of all of these activities, not just academic research.

A fourth question: who is responsible for ethical conduct? The researcher is rarely the only person involved in managing the data and the presentation, and cannot presume to control the process: commissioned research has to be reported as it progresses, while in research units the information gets carried back to a research team. Research findings have to be processed, submitted to funders or editors, and may well be subject to peer review. Everyone involved in the process has to take some ethical responsibility. Researchers, stakeholders, peers and the users of research, including think tanks and policymakers, all have things to say about the process of research; and where there are research participants, there is a strong case for bringing them into the discussion. The key is discursive engagement – encouraging ethical reflection by everyone involved.

Ethical consideration has to be seen as an integral part of research and practice. Because ethical rules cannot be treated as fixed and predictable, there needs to be an ethical discourse to be sure that researchers are aware of, and sensitive to, the ethical dimensions of their work. Ethical issues can arise at every stage of research: education, conception, development, proposal, process, conclusion and dissemination. The key to moral conduct rests, then, not in a single-stage process of review, but in continuous discursive engagement.

References

AAA (American Anthropological Association) (2012) 'Principles of professional responsibility', available from: http://ethics.americananthro.org/category/statement/ [accessed 25 January 2022].

ASA (Association of Social Anthropologists of the UK and the Commonwealth) (2001) 'Ethical guidelines for good research practice', available from: https://www.theasa.org/ethics/guidelines.shtml [accessed 25 January 2022].

Baumeister, R.F. and Leary, M.R. (1997) 'Writing narrative literature reviews', *Review of General Psychology*, 1(3): 311–20.

Boden, R., Epstein, E. and Latimer, J. (2009) 'Accounting for ethos or programmes for conduct? The brave new world of research ethics committees', *Sociological Review*, 57(4): 727–49.

BPS (British Psychological Society) (2018) *Code of Ethics and Conduct*, available from: https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct [accessed 25 January 2022].

BSC (British Society of Criminology) (2015) 'Statement of ethics for researchers', available from: https://www.britsoccrim.org/ethics/ [accessed 25 January 2022].

British Sociological Association (2017) 'Statement of ethical practice', available from: www.britsoc.co.uk [accessed 25 January 2022].

Center for Advanced Study (2005) *The Illinois White Paper: Improving the System for Protecting Human Subjects – Counteracting IRB 'Mission Creep'*, available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=902995# [accessed 25 January 2022].

COUHES (Committee on the Use of Humans as Experimental Subjects) (2019) 'Overview', available from: http://web.mit.edu/committees/couhes/index.shtml [accessed 25 January 2022].

Corrigan, O. (2003) 'Empty ethics: the problem with informed consent', *Sociology of Health and Illness*, 25(3): 769–92.

DHHS (US Department of Health and Human Services) (1991) Code of Federal Regulations: Title 45: Public Welfare: Part 46: Protection of Human Subjects, Washington, DC: Department of Health and Human Services, available from: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html [accessed 25 January 2022].

DHHS (US Department of Health and Human Services) (2018) Basic HHS Policy for Protection of Human Subjects, Washington, DC: HHS Office for Human Research Protections, available from: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/revised-common-rule-regulatory-text/index.html [accessed 25 January 2022].

Government of Canada (2018) Tri-council Policy Statement: Ethical Conduct for Research Involving Humans, Ottawa: Government of Canada, available from: www.pre.ethics.gc.ca/english/policystatement/policystatement.cfm [accessed 25 January 2022].

Grugulis, I. (2011) 'Research ethics and James Bond', Social Science Space [blog], 6 January, available from: http://www.socialsciencespace.com/2011/01/research-ethics-and-james-bond/ [accessed 25 January 2022].

Hammersley, M. (2006) 'Are ethical committees ethical?', *Qualitative Researcher*, 2: 3–8.

Hammersley, M. (2009) 'Against the ethicists: on the evils of ethical regulation', *International Journal of Social Research Methodology*, 12(3): 211–25.

Israel, M. (2004) 'Ethics and the governance of criminological research in Australia', Sydney: NSW Bureau of Crime Statistics Research, available from: https://www.bocsar.nsw.gov.au/Pages/bocsar_publication/Pub_Summary/General-Series/R55-Ethics-and-the-governance-of-criminological-research-in-Australia.aspx [accessed 25 January 2022].

Kruger, J. and Dunning, D. (1999) 'Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments', *Journal of Personality and Social Psychology*, 77(6): 1121–34

Lawton, A., Rayner, J. and Lasthuizen, K. (2012) *Ethics and Management in the Public Sector*, Abingdon: Routledge.

Murphy, E. and Dingwall, R. (2007) 'Informed consent, anticipatory regulation and ethnographic practice', *Social Science & Medicine*, 65(11): 2223–34.

NHMRC (National Health and Medical Research Council [Australia]) (2018) 'National statement of ethical conduct in human research', available from: [https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018# [accessed 25 January 2022].

Nuremberg Code (1947) available from: https://www.ushmm.org/information/exhibitions/online-exhibitions/special-focus/doctors-trial/nuremberg-code [accessed 2 April 2022].

Olson, R., Verley, J., Santos, L. and Salas, C. (2004) 'What we teach students about the Hawthorne studies: a review of content within a sample of introductory I-O and OB textbooks', *The Industrial-Organizational Psychologist*, 41(3): 23–39.

RESPECT Project (2004) RESPECT Code of Practice for socio-economic research, available from: https://www.respectproject.org/code/respect_code.pdf [accessed 25 January 2022].

Schrag, Z.M. (2011) 'The case against ethics review in the social sciences', *Research Ethics*, 7(4): 120–31.

SLSA (Socio-Legal Studies Association) (2009) Statement of Principles of Ethical Research Practice, available from: https://www.slsa.ac.uk/index.php/ethics-statement [accessed 25 January 2022].

Spicker, P. and Gordon, D.S. (1998) *Planning for the Needs of People with Dementia*, Aldershot: Avebury.

SRA (Social Research Association) (2003) *Ethical Guidelines*, available from: https://the-sra.org.uk/common/Uploaded%20files/ethical%20guidelines%202003.pdf [accessed 25 January 2022].

Tysome, T. (2006) 'Ethics guards are "stifling" creativity', *Times Higher Education Supplement*, 25 August, available from: https://www.timeshighereducation.com/news/ethics-guards-are-stifling-creativity/204946.article [accessed 25 January 2022].

Whyte, W.F. (1955) *Street Corner Society*, Chicago: University of Chicago Press.

3

Trust in institutions or the scientist? The drivers and mechanisms of trust in research and innovation as a means to improve related political outcomes

Caroline Gans-Combe

Introduction

This chapter aims to give an account of the problems relating to the trust placed in science and scientists in a context where there is a strong demand for results (as in a pandemic) and to propose ways of improving this trust, which is undermined, on the one hand, by the scientific practices themselves (such as the expression of contradictory views, publicised controversies and so on) and, on the other hand, by the use that the general public can make of these practices.

The chapter is built on a research approach constructed on a two-phase methodology:

- an extensive literature review on existing views pertaining to trust in science and a case study aimed at making visible the components of mistrust in science and/or scientists;
- built on the results of phase 1, the second phase looks at the mechanism of trust within the knowledge ecosystem – understood as a complex network or interconnected systems – and considers how to rebuild the link between science, the scientist and the civil society. This second phase used a specifically designed survey, testing and validating the latter's demands as well as data analytics processes.

The objective is to clearly identify the drivers and mechanisms of trust in research and innovation as a means to enhance both efficiency and related political outcomes.

Phase 1: The components of mistrust

Data collection and the background literature

Two methods were applied to identify the paradigms on which trust in research is built and the existing mechanisms to avoid unethical use of research results.

For the first approach a Boolean literature search identified data in four different categories: evidence of unethical use of research results in-house and outside, evidence of lack of trust/mistrust in relation to the unethical use of results, existing and non-existing mechanisms to identify risks of unethical use within and outside research organisations, and mechanisms to avoid/manage cases of misuse. The relevant keywords were used to search on Qwant and Google Scholar.

The second method involved a case study on French education. Data sets provided by an open data website and public walls were accessed. Data were collected through a scraping phase using Beautiful Soup and then saved to a .csv file awaiting analysis. The homogeneity was then observed, and dissonance of these likes with the educational choices of the related individuals predicted potential dropout risk. Since data could be collected from anonymous public discussion walls, this risks possible social engineering. One outcome of these restrictions is that they introduced a likelihood of bias in our data resulting from a lack of objectivity on the part of the chosen individuals from whom the data were collected.

The literature analysis showed an overall absence of institutional processes preventing misuse or biased use of research findings. This raised the question of how a dynamic trust could be built and applied to the use of research outcomes. The response was to look at the distribution and causes of mistrust in various contexts. Concerning vaccines, for example, issues regarding security of adjuvants generally arise, and it appears that researchers typically find it hard to respond, tending to propose answers not fully addressing the question: a known driver of frustration (Bateson, 1941). Besides justifying their findings to the public, another strategy scientists build upon to improve the acceptance of their results is peer review, where other qualified people constructively criticise a fellow researcher's work. This approach, while generally targeted towards ensuring a high level of trust for the results, in some cases, fuels mistrust, especially when individuals reviewing the research disagree on certain essentials (Ziegler, 2017). Finally, the frequent use of 'science' as justification and support for political action, tends to strain the scientific voice (Langer et al, 2016).

The literature shows that despite numerous efforts at enhancing trust for research results, current practices appear rather counterproductive. By institutionalising content validation processes (through cherry-picked publication, peer review and so on), by not responding directly and explicitly to questions from non-insiders, and by referring too often to science as the pillar of their action, the actors seeking to build trust in science and its achievements generate the opposite effect: suspicion and mistrust are significantly high, leading to low institutional confidence (Lewicki et al, 1998).

Trust and distrust in different contexts

To understand the different elements of trust, coming up with a clear concept definition is essential. Generally, trust is a belief in the reliability of an individual or entity. In the research context, therefore, trust relates to people's belief in scientists, their findings and their use. Trust could be treated as an action and a belief depending on the specific language used. In the French context, there is no distinction between the act and the feeling of trust. German and English users, however, distinguish between trust as action and as a level of belief (Lewicki et al, 1998).

Numerous researchers have linked trust with economic development. This argument is plausible as people with mutual trust are likely to cooperate in different forms of economic activities, leading to the growth of profitable companies. In addition, high levels of trust reduce monitoring, controlling and litigating costs in economic relations. Researchers investigating the connection between trust and economic outcomes do disagree as to the actual nature of the concept since each group applies the term uniquely (Welsh et al, 2008). For example, while psychologists consider trust a personal trait, sociologists treat it as part of social structure, while economists view it as a mechanism of economic choice.

To reconcile the social and economic arguments for trust, the *conventional* theory was formulated in France in the 1980s. In it, the link between social and economic aspects of trust is established on the basis that economic actors within the institutional framework act as social actors by agreeing on the conventions supposed to govern them (Simon, 1986). Therefore, these social contracts, insisting on the common good over individual interest, form the basis of trust in economic interactions among different parties. In that regard, conventions make the behaviour of people much more predictable, supporting an environment of trust in economic activities.

Trust and its components as described in literature.

The theoretical literature categorises trust as characteristic-based, process-based and institution-based. Characteristic-based trust is trait-dependent, while process-based trust relates to the confidence one party has in the others' actions. Institution-based trust depends on the level of trust an organisation elicits from people. Institutional trust is based on the trustworthiness of the different people working in particular organisations, avoiding mistrust by having highly reputable people in organisations and setting standards for a high level of organisational integrity (Hardy et al, 1998).

Identifying their modes of production is necessary to understand trust and mistrust. The trust or mistrust manifested by an individual depends on the intensity of social ties between individuals (Carucci, 2016). Personal proximity, learning and experience significantly predict the trust or mistrust between individuals. Confidence between individuals is built on their ability to keep promises. Generalised trust, however, results from the social structure, which requires the presence of contracts between parties as a means of entering agreements. In the generalised trust context, therefore, contracts, legal systems and institutional policies are required to build trust between different parties in the broader socio-economic background. In the contractual trust context, the likelihood of unethical behaviour generating mistrust is mediated by a third party, which is, in most cases, the judiciary (Brinkmann and Lentz, 2006).

The judiciary acts as a mechanism of avoiding mistrust by enforcing predetermined legal provisions if contracts are breached. Unfortunately, in the present research context, there are few readable sanctions in the event of a breach of the social contract, explaining the lack of trust in research institutions as economic actors rather than in the conceptual figure that is science (according to Ben-Yehuda and Oliver-Lumerman [2017] less than 7.5 per cent of identified cases reach court).

In contrast to institutional trust, the literature shows that trust between individuals is exclusively based on the personal traits of the persons involved. This depends on a person's belief in another's loyalty rather than rational risk calculation. Lack of trust in such interaction usually results from either party not honouring their obligations. Since there are no contracts, in this case, breach of trust in a personal context is much more complicated than in the institutional context as there are no predetermined guidelines on what should be done in any eventuality. The avoidance of breach of trust in the personal context thus depends on the level of social interaction between the individuals involved in personal agreements (Glover, 2018). As for trust in research findings, legal contracts between researchers and the public might not exist. It is, therefore, the role of the scientific community, the media and politicians to ensure that necessary measures are enacted to establish an atmosphere of trust between researchers and the public.

Findings

There has been a significant increase in the number of research results–based fraud identified over the last few years, from both a research design and a research use perspective. The two most plausible causes are either an actual increase in the number of unethical practices (including the so-called post-truth environment) or an improvement of the vigilance in identifying misconduct towards research at all levels (Fensham, 2014), including fact checkers (d'Ancona, 2017). Vigilance has certainly increased the detection of many cases, wrongful behaviour being the main reason for the increase. The scientific community concurs that despite the application of advanced detection measures, misconduct cases are still growing.

Lack of trust in scientific results: a question of neglect?

One of the leading causes of a lack of trust in scientific results is carelessness, specifically as far as data (Veldkamp et al, 2017) is concerned. Negligence in the way data is collected is structured. However, only a few forms of neglect are commonly highlighted in research. The lack of critical thinking about results and insufficient cross-checking leads to erroneous conclusions and recommendations. There have been cases of lax data management at various stages of the data-handling process, leading to a loss in the integrity of the data, meaning that any analyses will have inherent inaccuracies. Some researchers choose inappropriate methods and analytical tools; thus, the collected data might be inaccurate or prone to being wrongly interpreted. Likewise, cases of non-disclosure of information are important in the overall success of the research. All these factors question the validity of any research findings presented and hence the trustworthiness of scientific research. Finally, it is also important not to neglect the truncated inclusion of scientific results (selection bias) to claim scientific support for a given action (Geddes, 1990).

Lack of trust in scientific results: a question of conflict of interest?

In addition to the causes of mistrust in research outputs already discussed, conflict of interest is another significant factor as far as trustworthiness is concerned (Bekelman et al, 2003). In many cases, the researcher, or the organisation funding the research, might have a vested interest in a particular outcome. Such a situation skews the analysis of results towards the preferred outcomes. It is, however, noted that conflict of interest is not necessarily a misconduct per se unless their interests explicitly bias the researcher at any stage of the research (McTighe, 2019). One avoidance mechanism, in that regard, is being vigilant of one's interests at all stages of the research process and ensuring that these interests do not interfere with the data collection and analysis.

Lack of trust in scientific results: a question of societal demands?

One of the most important findings is the fact that countries compete in terms of the number of yearly publications. In the current age, professional activity is usually subjected to quantitative evaluation, which has led to 'Bibliometrics' as one of the most critical measures of the success of the researchers of a particular country. In this case, 'Bibliometrics' refers to the number of times a country's scientific publications are referenced in new publications. While this does not directly imply untrustworthiness, the consideration of the mere numbers of published works increases the likelihood of research that does not meet international standards being published. Thus most of the research studies in popular journals might not be appropriately validated raising questions about their trustworthiness. An example of the shortcomings of bibliometric indicators was seen in 2014 when *Nature* revealed that some leading publishers were removing over 100 fake articles from their databases (Van Noorden, 2014). This, therefore, points to the possibility of numerous fake articles in less prestigious journals never being identified because nobody will take time to scrutinise them. It is important not to neglect the appetite for support by third parties. The less trustworthy the stakeholders are considered (as shown by the Edelman (2019) Trust Barometer, described in the section 'Trust indicators'), the more likely they are to seek third-party validation of their programme (Bradley, 2011), including in an adversarial way.

Limited and unreliable sanctions against unethical research or unethical use of research

While research misconduct might be perpetrated by individual researchers, the research ecosystem can foster a culture of untrustworthiness (Brown, 2013) since inappropriate use of research outcomes by a third party is seldom considered (Frericks and Höppner, 2019). One of the measures available to combat unethical research behaviour is the law. The legal context could allow mechanisms such as whistleblowing to pave the way for investigations if ethical failures occur. The problem is that very few instances of research-related misbehaviours have been reported. Besides, out of all the cases reported, only a few end up in court, and most go unpunished. Additionally, many institutions have internal ethical guidelines that give direction in case of research misconduct. However, implementation is weak since the persons conducting research are usually involved in their deployment, and these tools – in an astonishing realisation of powerlessness – pay relatively little attention to the misuse of research results (Bird, 2014). Moreover, there are hardly any situations where cases of research misconduct or research results misuse enter the legal system. This means that research-related wrongful acts are not taken as seriously as other kinds of fraud, increasing the chances of the different actors in the research ecosystem participating in unethical behaviour. Whistleblowing is further jeopardised by the fact that denouncing the practices of peers, specifically when they are seniors, usually has severe implications for careers (see Mueller, 2019).

Trust indicators

Several tools have been developed over the years to measure people's trust in scientific research. The first is surveying in a binary (yes/no) or trinary (yes/no/no opinion) format concerning people's perceptions of scientific research. One such tool, the Edelman (2019) Trust Barometer, makes it possible to measure a lack of confidence with research in institutions such as governments, companies, non-governmental organisations (NGOs) and the media. A 2003 survey found that confidence in authorities was diminishing in favour of confidence in peers. In recent years, this trend has strengthened as a result of the advent of social networking sites, leading to more connections between people of similar ideologies (Sturgis and Smith, 2010).

Communication seems to be essential to correcting a lack of trust. In a research study seeking ways to build trust in the organisational context, findings indicate that 82 per cent of the respondents believed that fostering a culture of clear and transparent communication was essential. However, 81 per cent of the respondents also hinted that for trust, communication must be accurate regardless of acceptability (Brion et al, 2019). Regular engagement between employees and managers and the presence of managers during serious problems were identified as other factors for fostering trust.

Responses to the confidence crisis

As far as the legal approach is concerned, countries deal with trust deficiency in science in three different ways: explicitly stated statutory provisions defining the procedures for addressing issues relating to research misuse; no explicit legal national frameworks but specific institutional ethical codes that address issues relating to scientific integrity (in which we include impacts and adequate use of results); and no codified laws or rules relating to these issues, with research ethics issues dealt with on a case-by-case basis.

At the institutional level, governance is found to play a profound role in the overall trust relationship during innovations.

In any research project at the organisational level, there are numerous stakeholders, including public, private, governmental, national and international. The findings indicate that in research, each of these groups protect their respective institutions, hence maintaining a biased view, which makes objective research findings impossible. The EU-funded DEFORM project (Gans-Combe et al, 2019) found that the different stakeholders usually set divergent objectives depending on the specific bias of each group. Focusing on outcomes favourable to each respective group leads to potentially untrustworthy outputs.

Limitations of current governance structures

The governance structures employed in the management of trust in scientific research have several profound failures. The first is that while there are clearly defined policies governing research aspects such as ethics and accountability, there is poor implementation in relation to research results' (mis)use (Fox and Stoett, 2016). The second is that the governance bodies in institutions such as universities place too much emphasis on visible outcomes at the expense of policies meant to ensure the trustworthiness of research in scientific, political and economic contexts. Besides, these governance bodies have not established spaces where controversies involving scientific research can be rigorously addressed. There usually is confusion whenever cases of misconduct arise since the frameworks necessary for implementing corrective measures are non-existent. Finally, there are no tools to measure the bias in the interpretation of results in their use outside the scientific sphere. Thus, political insights and the risk inherent in a discovery (impact studies) are only rarely analysed *ex ante*, which means that they become significant only when they arrive in the public sphere (Nelkin, 1994), as is the case, for example, with CRISPR-Cas9 and biohacking.

Testing the 'misplaced response to the question asked' hypothesis: case study on academic orientation choices and trust impacts

To test our hypotheses, we set out to study an academic situation that today generates one of the most impactful societal problems, the prevalence of NEETs (people not in employment, education or training) and the reasons for their situation, in particular by measuring the correlation between misdirection – failure – and trust.

Academic work involves the dissemination and acquisition of knowledge in both the university and the organisational context. However, most of the research studies focusing on academic work collect knowledge and ignore dissemination. The assertion of lack of trust in the knowledge sector is therefore based on erroneous data since any available information collected does not consider the academic industry in its entirety (Cavestro et al, 2007). The assertion is that the trust deficiency towards the academic world is because education in higher learning institutions does not necessarily give students the knowledge and skills required to meet their objectives and answer questions such as: what path should I take to succeed? Higher education makes a latent promise to provide answers concerning their social positioning (Brown, 2003). This promise is the basis of the contract of trust as defined earlier, between the future students and the educational institution. We assume that as soon as the choice of orientation is made, the ability to succeed or fail in a given curriculum is established, and that the roots of trust or lack of trust in the institution are established.

The data indicate a strong correlation between trust and the perception of students towards the curriculum. From the equation investigating adherence to curriculum, factors such as social ecosystems and learner choices profoundly impact how students are attached to their specific fields of study. More specifically, students are found to be more likely to have positive attitudes towards their courses if their own persona most drives them in selecting them. Higher chances of failure are observed where the learner has to choose specific areas of specialisation due to pressures from their social ecosystem. The intersection between the forced choices and mistrust is argued to emerge from the amplification of negative experiences at any stage of their studies. Negative feelings lead the learner to perceive their academic choices even more negatively, ultimately resulting in frustration and mistrust: 50 per cent of people interviewed in the trust barometer indicate that the 'system' does not work for them, and the percentage of trust in the institution especially has consistently remained at the same level for 20 years; at 65–70 per cent, according to the indicators (Edelman, 2019).

More importantly, we hypothesise that a poor response to the expectations of future learners (whatever their age) misdirects and therefore fails, which would explain why the population of graduates with a university degree or equivalent remains relatively low in Europe (37 per cent: Eurostat, 2020). The learner then enters a vicious circle of amplification of bad experiences (bad orientation→bad social choices→bad careers→non-fulfilment [Cook-Sather, 2002]). We know that accumulation of these negative experiences leads to a lack of trust, completing the chain. If the sense of lack of trust does not seem to be statistically verified when it comes to research activities, perhaps knowledge transfer activities should be addressed; as, indeed, initial analyses seem to verify.

Preliminary conclusion

Overall, the trust towards scientific research has been declining since scientifically inadequate practices or biased interpretation of research results are reported more than examples of prudent approaches, but this assertion is hardly verified statistically. In many cases, unethical practices in the treatment of research findings result from the expectations of various players. When the results contradict the expectations of stakeholders, primarily the funders, then researchers, in many cases, might be tempted to alter the findings and align them with their promoters' interests. Such bias generally leads to inaccurate conclusions and recommendations by researchers according to interested third parties. Public exposure of these cases, such as during the

2007/08 financial crisis (Campbell, 2019), leads to the spread of mistrust towards all scientific researchers regardless of the true situation.

Most of the ethical issues in research emerge at the institutional level. A number of sentiments indicate the lack of clear procedures to fully address ethical issues at the institutional level, that is, by including issues related to the potential impact of research outputs and their use by third parties. Additionally, even when there are clear procedures, the transfer of fraud or misconduct cases and improper research use to the legal system are very few. Convictions for such 'malpractices' are thus virtually non-existent, meaning that researchers and/or research end users generally have no fear even if they present untrue findings, and it is even more the case for the end users of such results who build their decision on them, such as in policymaking: no policymakers advocating fake scientific results have ever been brought to court. This shows weaknesses in the enforcement of research policies both in institutions and in the law.

The issue of trust also arises as a result of societal demands

Countries compete for international rankings as this is paramount to their attractiveness and competitiveness. For this, it is necessary to mass publish. Aiming merely for a large number of publications, researchers engage in numerous studies usually of limited quality. Published in second-rate journals, these are typically inadequately scrutinised to reveal deficiencies. Even a well-respected journal such as *Nature* has withdrawn articles subsequently show to be deficient. Such exposure further increases the level of distrust towards scientific outputs.

Some of the plausible interventions suggested as a response to ethical failures in the treatment of research results include whistleblowing, establishing ethical guidelines and creating robust governance structures.

Another intervention effective in different contexts in the past is peer-reviewing research work. The problem arises, such as on climate change, when the different individuals disagree on the methodologies and the validity of the findings.

For the establishment of ethical guidelines, clearly defined policies are essential to instruct all actors of the research chain in terms of the treatment of findings. However, the principal limit to guidelines is that most of them do not cover the issue of research outside the institution (impact of research use) or even establish sanctions in the event of fraudulent communication. In recent cases of a deliberately misleading interpretation of climate change, NASA did not even demand a right of reply, allowing fanciful interpretations on the subject to still flourish (Berardelli, 2020). Most institutions poorly implement guidelines leading to malpractices, leading to distrust.

Furthermore, governance structures are not properly defined to regulate supranational-level research policies and most importantly research output usages. Different countries apply different approaches to governance in research. Hence, no simple readability of the trust can be placed in a particular research result or use.

Notwithstanding all these situations, which actually harm the research world more than science itself, perhaps the problems of trust in the scientific ecosystem are nested also in the fact that the knowledge industry is not statistically approached in its entirety. Our case study thus shows that the problem of trust also arises regarding transmission of knowledge, constituting another important aspect of academic activity.

Three recommendations could be made towards a solution:

- consider the 'knowledge industry' as a whole when analysing situations and attempting to provide solutions;
- extend the use of research results the operating rules already applied to make such investigative processes transparent – in particular for policymaking purposes – by providing for systematic procedures in the event of misuse of research outputs (in particular for post-truth cases), which would be simpler as the rules already exist (public corrigendum and so on);
- set up a recommendation tool rating the trust, which can be placed in a given research governance and results use process, based on objective criteria, such as the ethical ratings set up by Vigeo Eiris©.

For effective implementation, these would require adoption by the key stakeholders. Their appetite to adopt this type of tool needs to be explored.

Phase 2: The mechanism of trust within the knowledge ecosystem and validation of the demands of the civil society

This phase is an effort to analyse the mechanism of trust within the knowledge ecosystem and validate civil society's demands in terms of tools to restore confidence in the knowledge economy.

Due to a lack of trust in the knowledge and research industry, civil society loses confidence and becomes more distant and unclear in its attitude towards the industry though it generally trusts 'science' but not the actors involved. Our research project aimed to construct tools and solutions to enhance the bonds of trust between the stakeholders from civil society and the knowledge industry.

In the previous phase, analysis of the mechanisms of trust within the knowledge ecosystem has shown several strong expectations of civil society, particularly in terms of transparency.

Today's demands are centred on the possibility of proposing objective indicators and tools to measure the relevance of critical paths serving as an operational foundation for increased trust while focusing on stakeholders' demands. The following section deals with several issues such as the identification of stakeholders' understanding of knowledge economy (1) as a whole, (2) as per its components, (3) the trustworthy players in the ecosystem and (4) the parameters of lack of trust. Furthermore, an effort is made to identify the methodologies and tools for building trust and confidence and enhance adherence towards such tools. All these steps might help provide insights on the remediation process in line with stakeholders' views. Considering all the arguments, our concluding report tried to establish a remediation tool and provide sophisticated solutions to the knowledge economy's lack of trust.

Methodology

To understand who civil society sees as stakeholders in the knowledge economy, the research team deployed the following method. The targeted population included European civil society and those who affect it in any way possible. For the study, state-of-the-art standard Cochran's formula (Cochran and Carroll, 1953) was applied to extrapolate the appropriate sample size required to represent the whole population's view adequately. When calculated, the minimal requested sample was 1,359 individuals. To avoid sample error, 1,688 invitations were issued, to which 1,479 replies were received, giving a response rate of 87.61 per cent. The semi-randomised cohort comprised adults with an annual household income of €20,000 or more. Of the sample, 596 came from France, 550 from Germany and 542 from other EU Member States. This over-representation of the former two countries' has been weighted so that each country's results have the same relative weight as their population. Considering all the errors and responses, a sample of more than 1,359 was obtained, fulfilling the minimum expected sample. The process of questionnaire distribution was significantly difficult, and various methods were applied. Two delivery methods, including usual distribution channels (network and alumni database) and social pooling tools (SurveyCircle and so on) proved inadequate in providing statistically viable returns. Finally, the survey through professional panellists (Zoho Survey and Prolific) provided statistically convincing data. The raw data was in both text and numeric form. For refining data, we transformed the text strings to numerical equivalents by using Python DF and split the lines into single values of analysis using Python standard functions split.

The results were generated in which the first consideration was for understanding of the knowledge economy by stakeholders in civil society. Output tests affirmed the absence of any bias or errors.

Stakeholders see knowledge-based society as primarily the result of innovative approaches, regardless of the field, most promoted by the academic world, the sphere of education and the beneficiaries of this knowledge with a non-preponderant industrial research component and participation. Does this mean that industrial players are not perceived as innovative, or that they have little involvement in the strategic orientations underpinned by the rationale of knowledge creation? This remains to be verified. Influencers are not perceived as part of the innovation and knowledge ecosystem.

With this level of understanding of stakeholders, the team moved to the second step related to understanding the determinants of the lack of trust among civil society and holding those responsible who contribute to it.

The question of trust: the knowledge economy and knowledge ecosystem, a fundamental distinction

Literature abounds regarding civil society's trust in science, or rather scientists (Brewer and Ley, 2013). However, institutions remain under question. The 'scientist' as a person has more or less the same approval rating over the last 30 years: around two thirds of respondents in all surveys, from the old Eurobarometer 55.2 of 2001 to the more recent approaches available in Switzerland (Science Barometer Switzerland, 2019), Germany (Science Barometer – A Representative Survey of German Citizens on Science and Research, 2017), Ireland (Science in Ireland, 2015) and many other European countries. Our analyses show that institutions are not in this configuration.

Any confusion between the individual and their research 'field' is in line with a question raised long ago of the personification of the scientific approach (Campbell, 1975). This has guided the industry's rhetoric for more than half a century, including the management of its crises, particularly in terms of responsible research. As noted in the DEFORM project, the individual alone (and not his or her environment) often suffers from sanctions related to real or supposed breaches, without the intrinsic causes of such breaches being considered. This in no way excuses unwelcome, if not fraudulent, practices but may explain them and thus make prevention possible by amending the approach or establishing processes to anticipate them (Ashmawy, 2018).

We find the deep roots of the dichotomy in this lack of questioning of institutions that rely on individuals to avoid having to evolve (Zemba et al, 2006). This can be seen between the literature that insists on the continuity of trust in the scientist (the person) and his field (science), without addressing how domains and people fit into society, particularly through research organisations and dedicated ecosystems. The breakdown of trust is happening at this level and has been for at least two decades (Millstone and Van Zwanenberg, 2000).

This proxy logic has created paradoxical situations, as witnessed during the COVID-19 crisis where institutions were mistrusted, while fully fledged members of these institutions were selectively listened to with deference. The French case of Professor Raoult is emblematic in this respect. The latter is well-established in the hospital ecosystem (he is a university professor and hospital practitioner – PUPH in French – at La Timone in Marseilles), yet he has not suffered from the population's failure of confidence in the institution of which he is a pillar (Ramli, 2019).

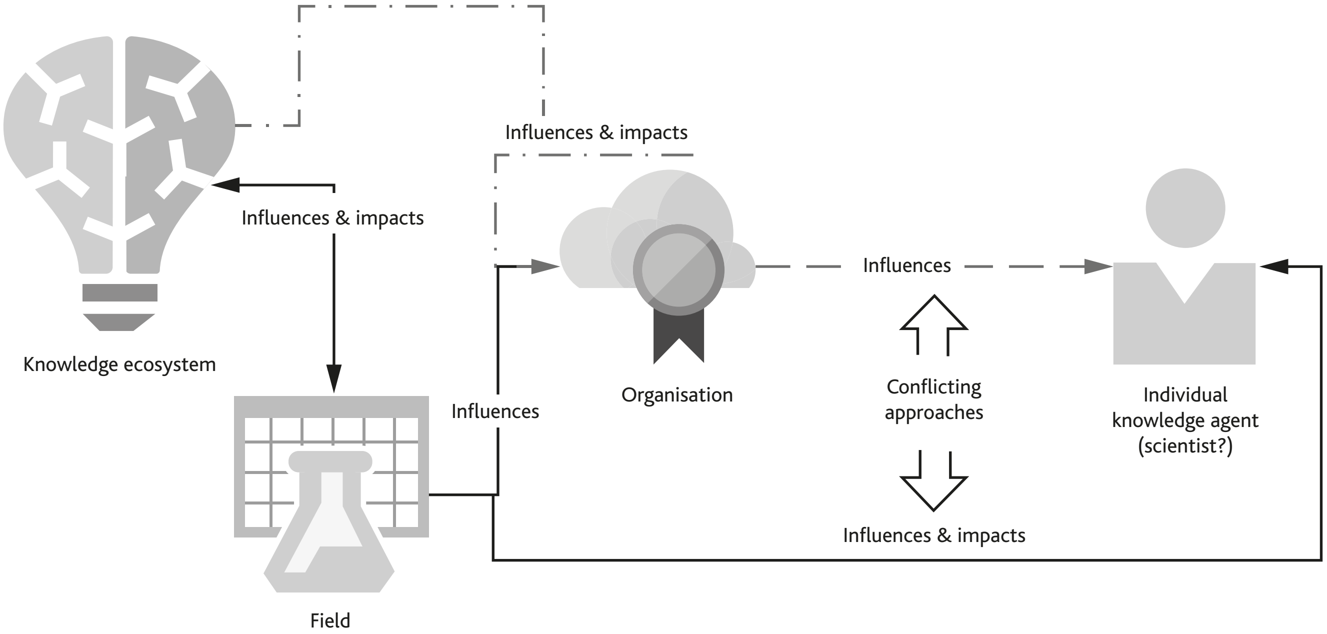
By focusing on the individual rather than the institution, even though failures often stem directly from its governance problems, the latter has sown the seeds of mistrust, particularly because permanence is no longer a concept accepted by stakeholders being called upon every day to comply with 'changes' imposed in ever-increasing numbers (Clarke, 2015). In other words, by claiming to be immutable (or unchallengeable), institutions have sown the seeds of a lack of confidence in them (Araiza, 1997), which allows actors dependent on these organisations to position themselves in strong opposition to them, only serving to increase doubts. Thus, this situation is symptomatic of a growing conflict between individual knowledge actors (scientists) wishing to be fully integrated into the knowledge ecosystem, without the filter of institutions, and the latter, which are determined to preserve their preponderance (Figure 3.1).

Conflicting approaches in the knowledge ecosystem

Increasingly, it is integrating into an ecosystem – more than just belonging to one – that in a way allows an actor, even a scientist, to be impactful (Posner et al, 2016). This institutional practice of individualisation/personalisation of the lack of trust has adversely given primacy to the individual over the ecosystem (if the latter faces the blame alone, why would it not benefit alone from the successes?), making it difficult to identify the components of the lack of trust, and therefore the targets of the actions for restoring this trust (Puusa and Tolvanen, 2006). Even if it already seems interesting for the institutions to think about escaping the personification of trust or lack of trust, it is necessary to identify, in the process of building trust, what the responsibility of the ecosystem and the actors is.

We have therefore formulated three questions about trust in the knowledge economy and its actors in general and by typology. It is no longer science that is being questioned but its different stakeholders. This is a strong societal trend in the age of citizen science that tends to propose involving citizens in the process of knowledge creation to increase confidence in its outputs. Thus:

Figure 3.1: Interrelationship between the individual and their institution in the knowledge ecosystem



1. Does trust in the field of science imply trust in the different players concerned with the subject matter?
2. What component of the ecosystem receives the most positive feedback from citizens? Hence, which level of trust is placed in the various players making up the knowledge economy?
3. Which component of the ecosystem is more responsible for the lack of trust? Thus identifying the boundary to the lack of trust.

Does trust in the field of science imply trust in the different players concerned with the subject matter?

It is not the scientific ecosystem that is currently in question. There is a global adherence, a generalised confidence in 'science' when not represented by an operator. It is moreover probably on the basis of this conceptual trust that citizens can appropriate – whether rightly or wrongly – scientific controversies to feed their own doubts (Nguyen and Catalan, 2020) and express them publicly through social networks (Waszak et al, 2018) by inventing their own 'scientific demonstrations', which may be oriented if not fallacious, but which are always capable of being accepted by the public because of the alleged authority and knowledge of the bearers of this type of message (Scheufele and Krause, 2019): 'doctors' who promote alternative cures for lethal diseases without adequate clinical evidence; child vaccination claimed to cause autism despite mounting evidence to the contrary (Hopf et al, 2019).

But this is not our point. If science as such is not questioned by the citizen (so little questioned that politicians tend to hide behind it when an unexpected crisis occurs [Besley and Velasco, 2020]), one can very legitimately wonder whether it is not the actors of the scientific world who are at the root of the often-violent rejection of scientific outputs or conclusions by some populations (Camargo and Grant, 2015). The debate is not new, but if the literature abounds in theoretical readings and solutions, the lack of an empirical approach to solving these questions is the probable root of the doubt that persists not regarding science but its various operators.

We have therefore questioned the level of confidence of Europeans in the said operators. For this, again with the help of the SciPy packages, we have undertaken a certain number of analyses relating to the answers of our survey participants.

As a reminder, our approach consists of confirming the general confidence granted to the scientific field and its actors without distinction (the ecosystem of knowledge: Q2 and Q3), and, on the other hand, in giving a 'score' to a level of confidence towards the different actors identified in the previous questions (Q4). We considered each return separately and then carried out parametric comparisons using state-of-the-art statistical data processing tools: filtering methodologies (univariate analysis of variables: sorting, distribution, histograms, bi- and multidimensional variate analysis of variables: sorting correlations, correlations, the test of the flat- and cross-sorting, as well as factor analysis).

To facilitate the reading of what will follow, a quick reminder of what can be expected from the latter tests is given next. Flat-sorting is the most basic method, consisting of obtaining a statistical measure question by question (the answer given by country to each question). It provides information on the criteria most important to respondents. Statistics are measured simply by dividing the number of responses per criterion by the total number of responses. Cross-sorting is the integration of an additional variable to the flat-sort to obtain statistics specific to each segment of the sample population. The additional factor we have to consider here is the country of origin of the respondents to report potential discrepancies between Member States and/or Partner Countries. As detailed, all the available methods for data analysis complied with different data types:

- Principal Component Analysis (PCA): a table of quantitative data of n individuals with *p* variables; PCA represents similarities between rows and links between columns.
- Correspondence Factor Analysis (CFA): contingency table representing the crossing of two qualitative variables on a sample of size n.
- Multiple Correspondence Analysis (MCA): allows a CFA with *p* qualitative variables performed using a complete disjunctive table.
- Discriminant Factor Analysis (DFA): allows description of the links between the variable to be explained and the explanatory variables and partition of all individuals into disjunctive classes.
- ANOVA: analysis of variance was also implemented.

This quick methodological parenthesis on the different statistical techniques used being completed, the most significant results concerning Q2 (Trust in the Knowledge Ecosystem/Industry) (see Figure 3.3).

There are a limited number of countries in which the maximum score (10) was given (25 per cent). Thus, in France, Germany, Greece, Portugal and the UK, confidence towards the Knowledge Industry (KI) in general is higher compared to other Member States and/or Partner Countries. The set mean is 5.73, median is 6 and the standard deviation is 2.844196. The standard deviation is a measure of how spread out numbers are. Most scores are between 3.2 and 8.8. The set is not widely scattered, which implies that, despite certain disparities, there is one sort of consensus in Europe on the issue of 'conceptual trust in KI' (see Figure 3.2).

In this context, and if players are trusted as a whole, how can the apparent contradiction between these figures and the sense of lack of trust long echoed by the main actors in the field (Benneworth, 2009) be explained? Could it be that it is not the scientific stakeholders that are at stake, but some component

Figure 3.2: Trust distribution in the knowledge industry

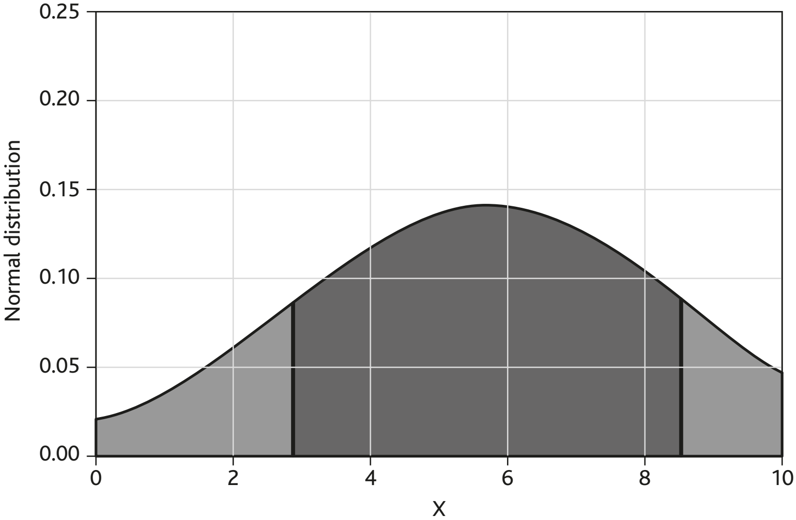
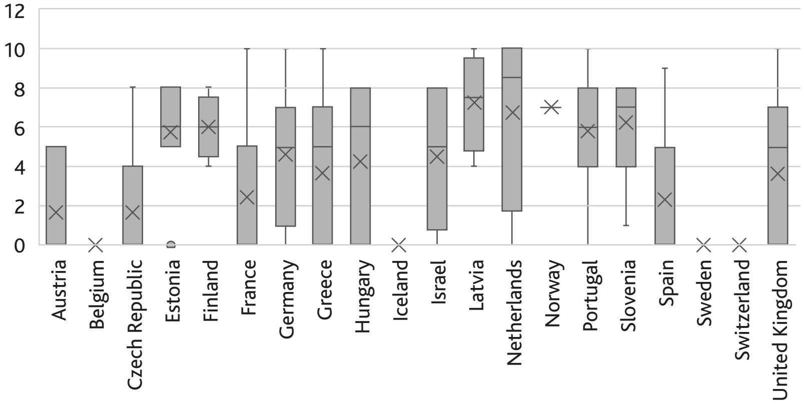
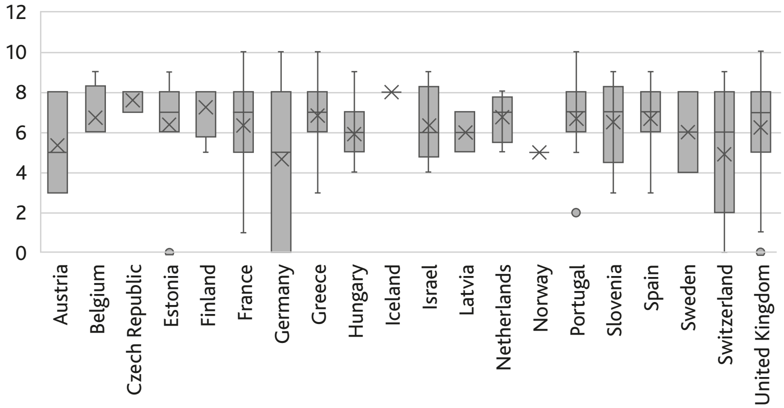


Figure 3.3: Trust in knowledge-industry players



of them? If so, which one? This hypothesis is going to be explored through the analysis of the third question in our survey. Nonetheless, this sense of an ecosystem on which citizens rely is somewhat confirmed by chart analysis showing a *negatively skewed* set (see Figure 3.3). The left tail is longer; the mass of the distribution is concentrated on the right of the figure and leans towards higher scores (above 6). The distribution is thus left-*skewed* (despite the fact that the curve itself appears to be skewed or leaning to the right).

Figure 3.4: Trust in the knowledge ecosystem

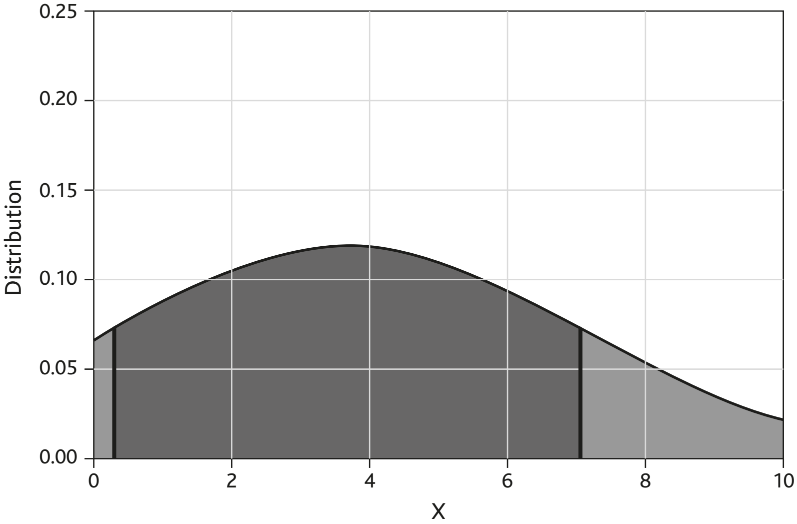


Furthermore, the mean (5.73) being slightly lower than the median (6), we have limited, though outgoing outliers in the low end, which implies that negative readings of the concept of trust in KE are more marginal (and analytically mostly identified in Switzerland, France, the UK and Germany) than the reverse.

In this respect, there is a certain bifurcation in these countries, as the people with the most and least confidence in the ecosystem reside there (see Figure 3.4). This observation, added to the fact that scores above 8 are scarce in the set, tends to indicate that this statement of confidence does not, for citizens, in any way mean a blank cheque for the scientific ecosystem. Analysing the volatility of the concerned population would be interesting to try to forecast a potential turnaround in the situation, but since this is a predictive analysis, it is outside the scope of the current investigation.

Concerning question 3, the most significant results are shown in Figure 3.4. The matter appears to be more disputed. Indeed, 39 per cent of the participants did not express an opinion about trust in given social entities/institutions (compared to only 12 per cent when the question was less detailed). This reluctance to answer when the issues increase in granularity has been observed for a long time (Tourangeau et al, 2000) and potentially reflects two sociological trends: the feeling that the respondent lacks sufficient knowledge of the field to make a commitment (but in this case, these rates should be found on a majority of questions, which is not the case) or the wish not to stigmatise this or that component of society (Lamont and Mizrachi, 2013). In this case, it seems that we are in the latter instance. In any event, the chart and statistical analysis shows that this questioning is more problematic for participants at the European level: candles are less homogeneous (see Figure 3.4), and analytically, the variance measuring

Figure 3.5: ANOVA analysis comparing trust in science with trust in the scientist



sample spread values is much higher than that expressed in the previous query (11.417289 vs 8.0839839). There is clearly less consensus on this topic than on the previous one. Median is higher – and in larger proportion than in question 2 – than the mean (median = 5, mean = 3.690331), which indicates again that the outliers are at the high end of the distribution, in spite of the important amount of non-respondents which skew the distribution to the right. This right skew is an oddity linked to the fact that the lower bounds are extremely low relative to the rest of the data. In other words, the propensity of survey participants to avoid answering is emblematic of a situation explained earlier.

In the end, we compared the potential interactions between the two questions. Broadly speaking, we asked ourselves whether the hypothesis of inference from one question to the other could be verified, that is, whether the answer on trust in research influenced the answer on research stakeholders. It appears not, as the null hypothesis about a subject was rejected and confirmed by the ANOVA analysis (Figure 3.5). The null hypothesis can be thought of as a *nullifiable* hypothesis. That means one can nullify it or reject it. What happens if it is rejected? It gets replaced with the 'alternate hypothesis', which postulates about what might be true about a given situation. In our context of interrelation, the null hypothesis can be worded as follows: 'Does trust in the ecosystem automatically imply trust in operators of this ecosystem?'. Logically, if this were the case, the answers to the two questions should be statistically correlated. However, this correlation does not exist. We can therefore say, in all objectivity, that *it is not because* *a population has conceptual confidence in the knowledge ecosystem that it will have the same confidence in the actors of the said ecosystem.*

The reasons for this disenchantment have been abundantly analysed, but the novelty lies in the fact that the problem of confidence in the ecosystem of knowledge is clearly not a question limited to individuals (contrary to the 'rotten apples' theory) but rather to the operators of the latter (Bonanno, 2015). Therefore the operators of knowledge would have to undergo metamorphosis, and not the researchers backed by these operators – without denying cases of fraudsters that remain merely low-signal incidents, even if the financial impact is far from negligible. In this context, would it make sense to move from witch-hunting to governance and organisational-failure hunting as it is increasingly clear – as demonstrated by 'Dieselgate'[1] – that *institutional deviations lead to personal deviations and not the other way round*?

What component of the ecosystem receives the most positive and negative feedback from citizens? Hence, which level of trust is placed in the various players making up the knowledge economy?

The lack of confidence in the research operators being empirically established, trying to identify which operators are concerned is now necessary. As already mentioned, we have identified six components of stakeholders in the research process. These six components are significant for what is today understood as research actors in the broadest sense. This was confirmed in the first question of this study. These six groups have been ranked to answer this question.

Pairwise comparisons show that we may strongly reject the null hypothesis ($p < 0.01$) only for the pair of countries in medium grey and conclude that only a limited number of countries differ in their views of the level of trust directed towards industrial research. However, the grey graduation indicates that these proportions are not all of the same intensity. Light grey thus indicates a very strong presumption against the null hypothesis, while medium grey indicates a weak presumption. The dark grey indicates, as mentioned earlier, the absence of presumption against the null hypothesis (Table 3.1).

It is, therefore, safe to say that the following findings show a relative consensus at the European level. Furthermore, the statistics also show a high rate of non-response, supporting the idea of a certain difficulty for participants to take a stand when pointing the finger at this or that stakeholder as being responsible for a situation.

This non-response rate could be problematic, leading to a bias in the readability of the results, so we have weighted the results to reduce the latter risk.

As we do not have precise information on the participants outside their country of origin (this information was not collected to avoid a risk of

Table 3.1: Absence of presumption against the null hypothesis

Austria	Belgium	Czech Republic	Estonia	Finland	France	Germany	Greece	Hungary	Iceland
Austria	1	0.955785	0.987289	0.541464	0.691739	1.97466e-06	0.146483	0.635401	0.24885
Belgium	0.955785	1	0.935864	0.459431	0.625448	2.68012e-08	0.078548	0.535042	0.174217
Czech Republic	0.987289	0.935864	1	0.484104	0.664226	1.16967e-09	0.0650663	0.567383	0.170418
Estonia	0.541464	0.459431	0.484104	1	0.849991	9.75785e-10	0.270123	0.743523	0.495095
Finland	0.691739	0.625448	0.664226	0.849991	1	1.0412e-06	0.283664	0.973951	0.442525
France	1.97466e-06	2.68012e-08	1.16967e-09	9.75785e-10	1.0412e-06	1	1.03096e-148	2.25984e-41	5.83064e-11
Germany	0.146483	0.078548	0.0650663	0.270123	0.283664	1.03096e-148	1	0.00172319	0.768243
Greece	0.635401	0.535042	0.567383	0.743523	0.973951	2.25984e-41	0.00172319	1	0.179562
Hungary	0.24885	0.174217	0.170418	0.495095	0.442525	5.83064e-11	0.768243	0.179562	1
Iceland	0.65347	0.670255	0.628493	0.37941	0.462654	0.00105663	0.174548	0.428163	0.224246
Israel	0.43004	0.352373	0.366921	0.805948	0.692571	9.06773e-08	0.490703	0.538357	0.703441
Latvia	0.30003	0.244706	0.253154	0.53803	0.476737	0.000940691	0.99267	0.351886	0.883932
Netherlands	0.141227	0.0991838	0.0974038	0.262332	0.245577	0.00110169	0.572777	0.112824	0.523147
Norway	0.617554	0.579878	0.606059	0.884567	0.806574	0.0288477	0.791636	0.77425	0.867185
Portugal	0.557107	0.450122	0.470292	0.851167	0.93208	2.01992e-66	0.000175683	0.772589	0.214819
Slovenia	0.248388	0.186958	0.188725	0.471371	0.420608	2.02851e-05	0.996226	0.244485	0.864716
Spain	0.633725	0.532886	0.564907	0.743748	0.975038	2.34103e-43	0.00130616	0.997435	0.177312
Sweden	0.634713	0.582531	0.613929	0.987717	0.879869	0.00101946	0.565361	0.83866	0.679558
Switzerland	0.00174489	0.000344928	0.000153493	0.000727876	0.00289554	0.00719374	3.37729e-05	1.45145e-07	0.00218404
United Kingdom	0.492508	0.381318	0.391128	0.95129	0.849855	3.59468e-172	2.95681e-10	0.532066	0.249147

(continued)

Table 3.1: Absence of presumption against the null hypothesis (continued)

Israel	Latvia	Netherlands	Norway	Portugal	Slovenia	Spain	Sweden	Switzerland	United Kingdom	
0.65347	0.43004	0.30003	0.141227	0.617554	0.557107	0.248388	0.633725	0.634713	0.00174489	0.492508
0.670255	0.352373	0.244706	0.0991838	0.579878	0.450122	0.186958	0.532886	0.582531	0.000344928	0.381318
0.628493	0.366921	0.253154	0.0974038	0.606059	0.470292	0.188725	0.564907	0.613929	0.000153493	0.391128
0.37941	0.805948	0.53803	0.262332	0.884567	0.851167	0.471371	0.743748	0.987717	0.000727876	0.95129
0.462654	0.692571	0.476737	0.245577	0.806574	0.93208	0.420608	0.975038	0.879869	0.00289554	0.849855
0.00105663	9.06773e-08	0.000940691	0.00110169	0.0288477	2.01992e-66	2.02851e-05	2.34103e-43	0.00101946	0.00719374	3.59468e-172
0.174548	0.490703	0.99267	0.572777	0.791636	0.000175683	0.996226	0.00130616	0.565361	3.37729e-05	2.95681e-10
0.428163	0.538357	0.351886	0.112824	0.77425	0.772589	0.244485	0.997435	0.83866	1.45145e-07	0.532066
0.224246	0.703441	0.883932	0.523147	0.867185	0.214819	0.864716	0.177312	0.679558	0.00218404	0.249147
1	0.319032	0.237349	0.141952	0.438809	0.390632	0.214012	0.427354	0.436958	0.014993	0.360206
0.319032	1	0.683495	0.380602	0.986309	0.620438	0.637776	0.537805	0.878974	0.00334067	0.697685
0.237349	0.683495	1	0.716124	0.815283	0.39759	0.996511	0.35146	0.651276	0.0697676	0.439559
0.141952	0.380602	0.716124	1	0.624359	0.131459	0.675374	0.1121	0.425404	0.12206	0.149029
0.438809	0.986309	0.815283	0.624359	1	0.819572	0.807744	0.774656	0.907139	0.167774	0.858512
0.390632	0.620438	0.39759	0.131459	0.819572	1	0.284879	0.771274	0.903564	5.27339e-08	0.705011
0.214012	0.637776	0.996511	0.675374	0.807744	0.284879	1	0.243451	0.6246	0.02692	0.323267
0.427354	0.537805	0.35146	0.1121	0.774656	0.771274	0.243451	1	0.839241	1.17877e-07	0.524223
0.436958	0.878974	0.651276	0.425404	0.907139	0.903564	0.6246	0.839241	1	0.0382852	0.959885
0.014993	0.00334067	0.0697676	0.12206	0.167774	5.27339e-08	0.02692	1.17877e-07	0.0382852	1	1.46329e-08
0.360206	0.697685	0.439559	0.149029	0.858512	0.705011	0.323267	0.524223	0.959885	1.46329e-08	1

reverse engineering, a major non-compliance with the GDPR [General Data Protection Regulation]). We consider that, precisely in view of this volume of non-response, which manifests itself only in this question, there could be a risk of systematism of certain answers.

To avoid our samples not accurately reflecting the general population, we have chosen, rather than accept a poor match between the sample and the population, to use weights to bring the two more closely into line. This is known as ‘non-response weighting’.

The reweighting methods consist of dealing with non-response by modifying the survey weighting of respondents, to fill in the absence of certain answers. Each individual k is then assigned an answer probability c_k, and if this is known for all individuals, then unbiased estimators are available using the respondents’ responses weighted by survey weights divided by the probabilities of answer. In the case of estimating a total Y and a simple random sample design of n individuals among N, we obtain the following unbiased estimator:

Equation 1 reweighting estimator

$$\breve{Y} = \sum_{k \in R} \frac{Y_k}{\pi_k \times c_K} = \sum_{k \in R} \frac{Y_k}{\frac{n}{N} \times c_K} = \sum_{k \in R} w_k Y_k$$

where R is the set of survey respondents, π_k is the probability of inclusion of the individual k in the sample (equal here to n/N) and w_k is the modified weight of respondent k. However, we assumed that the distribution in the responding population was the same as in the non-responding population. Of course, this is debatable from a methodological point of view, but the lack of precise information (see earlier regarding regulatory constraints on data processing) leads us to this hypothesis. It is therefore the intensity, not the distribution, of this population that the analysis has impacted.

At the same time, identifying these components is deemed essential to build the trust and to attribute the responsibilities to the people. So, three questions were formulated. As the results from the previous phase of the study made it clear that it was not the science that creates mistrust but the different stakeholders that make it up, the questions targeted those stakeholders. The questions included in the research are: (1) Does trust in the field of science imply trust in the different players concerned with the subject matter? (2) What component of the ecosystem receives the most positive feedback from citizens? Hence, which level of trust is placed in the various players making up the knowledge economy? (3) Which component of the ecosystem is more responsible for the lack of trust? Thus the perimeter of the lack of trust was identified.

Deriving insight from the survey questionnaires, the discussions on these questions are held. It is concluded that if a population has conceptual confidence in the knowledge ecosystem, it does not imply that it will have the same confidence in the actors of the said ecosystem. This point has effectively distinguished between the concepts of science and the actors dealing with it as proposed earlier. Second, no knowledge ecosystem stakeholders score either positively or negatively, except for influencers (that is, agents having the ability to influence end users of a product or service by promoting or recommending the items on social media and beyond). As such, no actors are trusted or distrusted except influencers who are clearly distrusted. Third, academic research collects the most trust markers. However, it is also considered relatively responsible for the problems of trust raised by the knowledge economy. Fourth, industrial research received a relatively good confidence score from civil society. Fifth, researchers (243 responses) and public research and innovation policymakers including ministries and research organisations (283 responses) are primarily responsible for the lack of confidence in the knowledge ecosystem. Further, respondents give a good confidence score to the end users of the knowledge economy's outputs (patients, and so on), which pleads for the deployment of citizen science as a future important part of this ecosystem. After the second round, the lack of trust was placed on private financiers (464 cumulative responses). Finally, after the two questioning rounds, 42 per cent of the population found it difficult to attribute responsibility for a lack of confidence in the knowledge ecosystem, indicating a problem in understanding the granularity of its components.

Conclusion

Public trust in research and policy is a key theme of this volume. Understanding the issues of civil society regarding lack of trust towards stakeholders in the knowledge economy, the remediation process is presented. The process focuses on identification of the potential methodologies and tools for rebuilding trust and confidence in the knowledge economy and ecosystem. The process proposed is considered useful by the consensus between the participants of the study. The first phase of study already made clear the significant expectation for transparency in the development of policies relating to the knowledge ecosystem. Policymakers should take this into account in their work. How they develop policy and the evidence they depend on can help or hinder the development of trust. Additionally, objective tools are needed to assess and make this degree of transparency visible. In fact, the public's approach seems not to be focused on insincere discourses claiming to better governance in the knowledge economy. They are more focused on a tool strengthening confidence in the knowledge economy: a tool providing a greater visibility by tracking down and reducing the misuse of research results, enhancing the social legitimacy of the research/innovation and making effective the data sourcing. Citizens ask to become actors in R&I (research and innovation) advances and therefore wish to understand them before adopting them: evoking the progress of a research project in the form of a beautiful narrative is largely insufficient to build trust, and pedagogical marketing of innovation – and public policies in this area – need to be built. The format of the tool of trust should preferably be set by recommendations and not solely by relying on documents.

Furthermore, interactivity and participatory trust building is greatly needed. Stakeholders considered that the economic modelling should be established *ex post* and not *ex ante*. They also desire to see the tool become part of a certain universality, in line with what people understand today of the knowledge society. Finally, they support the idea that machine learning-based recommendation tools and indicators could be integrated into the appraisal processes of R&I proposals to promote operational objectivity with respect to their strategic positioning. In a nutshell, the interests of civil society towards the knowledge economy – what can be named the societal expectations of civil society toward the knowledge economy – were found to be of general interest. In other words, *knowledge can be taken as a common good that creates meaning rather than economic value.*

Note

[1] The Volkswagen emissions scandal, also known as Dieselgate or Emissionsgate, designates a fraud uncovered by the United States Environmental Protection Agency (EPA), which led in 2015 to the issuance of a formal notice of violation of the Clean Air Act to the Volkswagen Group. The agency had found that the automaker had intentionally programmed diesel engines to activate their emissions controls only during laboratory emissions testing, which caused the vehicles to artificially meet legal emissions standards during regulatory testing, while they emitted up to 40 times more in real-world driving.

References

Araiza, W.D. (1997) 'Democracy, distrust, and the public trust: Process-based constitutional theory, the public trust doctrine, and the search for a substantive environmental value', *UCLA Law Review*, 45: 385.

Ashmawy, I.K.I. (2018) 'Maintaining ethical behaviour in universities adopting the integrity approach', in B.M. de Waal and P. Ravesteijn (eds) *Proceedings of the 14th European Conference on Management, Leadership and Governance*, Reading: ACPI, p 10.

Bateson, G. (1941) 'IV. The frustration-aggression hypothesis and culture', *Psychological Review*, 48(4): 350–5.

Bekelman, J.E., Li, Y. and Gross, C.P. (2003) 'Scope and impact of financial conflicts of interest in biomedical research: a systematic review', *JAMA*, 289(4): 454–65.

Benneworth, P. (2009) *The Challenges for 21st Century Science: A Review of the Evidence Base Surrounding the Value of Public Engagement by Scientists*, Universiteit Twente: Center for Higher Education Policy Studies.

Ben-Yehuda, N. and Oliver-Lumerman, A. (2017) *Fraud and Misconduct in Research: Detection, Investigation, and Organizational Response*, Ann Arbor: University of Michigan Press.

Berardelli, J. (2020) '10 common myths about climate change – and what science really says', 27 February, CBS News, available from: https://www.cbsnews.com/news/climate-change-myths-what-science-really-says/ [accessed 26 January 2022].

Besley, T. and Velasco, A. (2020) 'Politicians can't hide behind scientists forever – even in a pandemic', *#LSEThinks* [blog], 6 May, available from: https://blogs.lse.ac.uk/covid19/2020/05/06/politicians-cant-hide-behind-scientists-forever-even-in-a-pandemic/ [accessed 26 January 2022].

Bird, S.J. (2014) 'Socially responsible science is more than "good science"', *Journal of Microbiology & Biology Education*, 15(2): 169–72.

Bonanno, E.R. (2015) 'An evidential review of police misconduct: officer versus organization', *2015 Undergraduate Awards*, available from: https://ir.lib.uwo.ca/ungradawards_2015/9/ [accessed 8 April 2022].

Bradley, R.S. (2011) *Global Warming and Political Intimidation: How Politicians Cracked Down on Scientists as the Earth Heated Up*, Amherst, MA: University of Massachusetts Press.

Brewer, P.R. and Ley, B.L. (2013) 'Whose science do you believe? Explaining trust in sources of scientific information about the environment', *Science Communication*, 35(1): 115–37.

Brinkmann, J. and Lentz, P. (2006) 'Understanding insurance customer dishonesty: outline of a moral-sociological approach', *Journal of Business Ethics*, 66(2/3): 177–95.

Brion, S., Mo, R. and Lount, R.B., Jr (2019) 'Dynamic influences of power on trust: changes in power affect trust in others', *Journal of Trust Research*, 9(1): 6–27.

Brown, A. (2013) 'Understanding pharmaceutical research manipulation in the context of accounting manipulation', *Journal of Law, Medicine & Ethics*, 41(3): 611–19.

Brown, P. (2003) 'The opportunity trap: education and employment in a global economy', *European Educational Research Journal*, 2(1): 141–79.

Camargo, K., Jr and Grant, R. (2015) 'Public health, science, and policy debate: being right is not enough', *American Journal of Public Health*, 105(2): 232–5.

Campbell, J.A. (1975) 'The polemical Mr. Darwin', *Quarterly Journal of Speech*, 61(4): 375–90.

Campbell, J.L. (2019) 'Self-responsibility gone bad: institutions and the 2008 financial crisis', *American Behavioral Scientist*, 63(1): 10–26.

Carucci, R. (2016) 'Why ethical people make unethical choices', *Harvard Business Review*, available from: https://hbr.org/2016/12/why-ethical-people-make-unethical-choices [accessed 26 January 2022].

Cavestro, W., Durieux, C. and Monchatre, S. (2007) *Travail et reconnaissance des compétences*, Paris: Economica.

Cochran, W.G. and Carroll, S.P. (1953) 'A sampling investigation of the efficiency of weighting inversely as the estimated variance', *Biometrics*, 9(4): 447–59.

Cook-Sather, A. (2002) 'Authorizing students' perspectives: toward trust, dialogue, and change in education', *Educational Researcher*, 31(4): 3–14.

Clarke, J.A. (2015) 'Against immutability', *Yale Law Journal*, 125: 2.

d'Ancona, M. (2017) *Post-Truth: The New War on Truth and How to Fight Back*, London: Random House.

Edelman (2019) *Edelman Trust Barometer: Global Report*, available from: https://www.edelman.com/trust/2019-trust-barometer [accessed 8 April 2022].

EUROSTAT (2020) Population by educational attainment level, sex and age (%) - main indicators, available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Share_of_the_population_by_educational_attainment_level_and_selected_age_groups,_2020_(%25).png [accessed 8 April 2022].

Fensham, P.J. (2014) 'Scepticism and trust: two counterpoint essentials in science education for complex socio-scientific issues', *Cultural Studies of Science Education*, 9(3): 649–61.

Fox, O. and Stoett, P. (2016) 'Citizen participation in the UN sustainable development goals consultation process: toward global democratic governance?', *Global Governance*, 22(4): 555–74.

Frericks, P. and Höppner, J. (2019) 'The turn toward self-responsibility in current societies: differences, challenges, and perspectives', *American Behavioral Scientist*, 63(1): 3–9.

Gans-Combe, C., Faucheux, S., Kuszla, C., Petousi, V. and Tina, G. (2019) *DEFORM Main Findings: A Policy Report*, available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3389414 [accessed 26 January 2022].

Geddes, B. (1990) 'How the cases you choose affect the answers you get: selection bias in comparative politics', *Political Analysis*, 2: 131–50.

Glover, S. (2018) 'Who do they think we are? Politicians imagine the nation', *Griffith Review*, 61, available from: https://www.griffithreview.com/articles/who-do-they-think-we-are-politicians-imagine-australia/ [accessed 26 January 2022].

Hardy, C., Phillips, N. and Lawrence, T. (1998) 'Distinguishing trust and power in interorganizational relations: forms and facades of trust', in C. Lane and R. Bachman (eds) *Trust Within and Between Organizations: Conceptual Issues and Empirical Applications*, Oxford: Oxford University Press, pp 64–87.

Hopf, H., Krief, A., Mehta, G. and Matlin, S.A. (2019) ‘Fake science and the knowledge crisis: ignorance can be fatal’, *Royal Society Open Science*, 6(5): art 190161, available from: https://doi.org/10.1098/rsos.190161 [accessed 26 January 2022].

Lamont, M. and Mizrachi, N. (eds) (2013) *Responses to Stigmatization in Comparative Perspective*, London: Routledge.

Langer, L. Tripney, J. and Gough, D. (2016) *The Science of Using Science: Researching the Use of Research Evidence in Decision-Making, Final Report*, London: EPPI-Centre Social Science Research Unit UCL Institute of Education University College London, available from: http://www.alliance4usefulevidence.org/assets/Science-of-Using-Science-Final-Report-2016.pdf [accessed 17 August 2019].

Lewicki, R.J., McAllister, D.J. and Bies, R.J. (1998) ‘Trust and distrust: new relationships and realities’, *Academy of Management Review*, 23(3): 438–58.

Millstone, E. and Van Zwanenberg, P. (2000) ‘A crisis of trust: for science, scientists or for institutions?’, *Nature Medicine*, 6(12): 1307–8.

McTighe, T. (2019) ‘What is considered a conflict of interest or what to disclose in lectures and publications?’, *Reconstructive Review*, 9(1), available from: https://doi.org/10.15438/rr.9.1.221 [accessed 26 January 2022].

Mueller, T. (2019) *Crisis of Conscience: Whistleblowing in an Age of Fraud*, London: Atlantic Books.

Nelkin, D. (1994) ‘Forbidden research: limits to inquiry in the social sciences’, in E. Erwin, S. Gendin and L. Kleiman (eds) *Ethical Issues in Scientific Research: An Anthology*, London: Routledge, pp 355–70.

Nguyen, A. and Catalan, D. (2020) ‘Digital mis/disinformation and public engagement with health and science controversies: fresh perspectives from Covid-19’, *Media and Communication*, 8(2): 323–8.

Posner, S.M., McKenzie, E. and Ricketts, T.H. (2016) ‘Policy impacts of ecosystem services knowledge’, *Proceedings of the National Academy of Sciences*, 113(7): 1760–5.

Puusa, A. and Tolvanen, U. (2006) ‘Organizational identity and trust’, *Electronic Journal of Business Ethics and Organization Studies*, 11(2): 29–33.

Ramli, A.H. (2019) ‘Patient trust on The Hospital Service Delivery System’, *Business and Entrepreneurial Review*, 16(1): 17–30.

Robert Bosch Stifftung Science Barometer (2017) A Representative Survey of German Citizens on Science and Research, available from: https://www.bosch-stiftung.de/en/project/science-barometer-representative-survey-german-citizens-science-and-research [accessed 8 April 2022].

Scheufele, D.A. and Krause, N.M. (2019) ‘Science audiences, misinformation, and fake news’, *Proceedings of the National Academy of Sciences*, 116(16): 7662–9.

Science Foundation Ireland (2015) *Science Foundation Ireland—Science in Ireland Barometer: An Analysis of the Irish Public's Perceptions and Awareness of STEM in Society*, available from: https://www.sfi.ie/resources/SFI-Science-in-Ireland-Barometer.pdf [accessed 8 April 2022].

Simon, H.A. (1986) 'Rationality in psychology and economics', *Journal of Business*, 59(4, Pt 2): S209–S224.

Sturgis, P. and Smith, P. (2010) 'Assessing the validity of generalized trust questions: what kind of trust are we measuring?', *International Journal of Public Opinion Research*, 22(1): 74–92.

Tourangeau, R., Rips, L.J. and Rasinski, K. (2000) *The Psychology of Survey Response*, Cambridge: Cambridge University Press.

University of Zurich Department of Communication and Media Research – IKMZ – Westfälische Wilhelms-Universität Münster – Institut für Kommunikation (2019) Science Barometer Switzerland, available at https://wissenschaftsbarometer.ch/en/science-barometer-switzerland/ [accessed 8 April 2022].

Van Noorden, R. (2014) 'Online collaboration: scientists and the social network', *Nature*, 512(7513): 126–9.

Veldkamp, C.L.S., Hartgerink, C.H., van Assen, M.A. and Wicherts, J.M. (2017) 'Who believes in the storybook image of the scientist?', *Accountability in Research*, 24(3): 127–51.

Waszak, P.M., Kasprzycka-Waszak, W. and Kubanek, A. (2018) 'The spread of medical fake news in social media–the pilot quantitative study', *Health Policy and Technology*, 7(2): 115–18.

Welsh, R., Glenna, L., Lacy, W. and Biscotti, D. (2008) 'Close enough but not too far: assessing the effects of university–industry research relationships and the rise of academic capitalism', *Research Policy*, 37(10): 1854–64.

Zemba, Y., Young, M.J. and Morris, M.W. (2006) 'Blaming leaders for organizational accidents: proxy logic in collective-versus individual-agency cultures', *Organizational Behavior and Human Decision Processes*, 101(1): 36–51.

Ziegler, R. (2017) 'Trusting science in an age of distrust', *EuroScientist*, 13 December, available from: https://www.euroscientist.com/trusting-science-age-distrust/ [accessed 4 December 2018].

4

Modes of influence: what can we learn from international codes of ethics for health-related research?

Helen Busby

Introduction

Many international ethics codes and guidelines exist for health-related research.[1] The Declaration of Helsinki, first promulgated in 1964, is recognised as a foundational statement in international research ethics. This milestone Declaration might seem inevitable in retrospect, following as it did some years after the Nuremberg Code was devised in the aftermath of the Second World War, yet historians have underlined that the road to Helsinki was not a straight one: proposals to develop an international medical research ethics code were greeted at first with considerable ambivalence and resistance by the World Medical Association (WMA; Bonah and Schmaltz, 2020). Despite its troubled beginnings, the Declaration formulated and asserted the rights of research subjects in clinical studies for future generations: historians, ethicists and clinicians, among others, have given extensive attention to its place in the architecture of post-war clinical research.[2]

The domain of health-related research has expanded greatly since the formative work was undertaken for the early Declaration. Following the Alma-Ata Declaration of 1978, a broad approach to health promotion emerged as key to the attainment of the World Health Organization's 'Health for All' strategy.[3] Public health approaches were widely advocated, and multidisciplinary approaches to global health challenges gradually came to the fore. Added to the emphasis on a holistic approach to well-being that characterised this influential movement, a more recent 'One Health' movement mandated the bringing together of different disciplines to research health across different sectors – such as animal health, food and environmental safety – alongside human health (Lebov et al, 2017). Even for specific diseases that have dominated research endeavours in global health – such as malaria, TB and other endemic diseases – a broad perspective was now understood to be necessary. To this we can add the increasingly sophisticated contributions of epidemiology and the social sciences to international health research (Dankwa-Mullan et al, 2010). Finally, the embedding of the Sustainable Development Goals in the architecture of the United Nations and its health agencies further underlined the breadth of approach needed to achieve the desired health improvements (Buse and Hawkes, 2015). Thus, the research agenda in global health has moved well beyond clinical research.

Helsinki (as it will be referred to here for brevity) was amended no fewer than seven times over the subsequent decades, in addition to the several notes of clarification that were issued on specific points. While earlier revisions focused on extending and updating the text, substantial changes were introduced in subsequent years. In response to widespread concerns arising from the conduct of clinical trials in countries with low resources, a new requirement was introduced into the 2000 version to emphasise the need for research to be responsive to the health needs of study populations. Expanded in 2008, this clause sought to address the transition of participants and their healthcare after the end of a research study, at which point they could be placed in a highly vulnerable position.[4] More generally, this version marked a shift from a focus on individual research subjects to include a broader emphasis on equity and justice. The debate crystallised around HIV trials undertaken in countries that could offer only limited treatment options for AIDS (Schuklenk, 2001). The trialling of HIV antiretroviral drugs without due consideration of what would come next, after the researchers had left the study population behind, had come to be seen as highly problematic (Lurie and Wolf, 1997). However, as critics of this requirement observed from the outset, researchers may not be able to obtain funding for the provision of care after the end of trials and are rarely in a position to shape public policy in host countries. Most were in agreement that greater attention should be given to the provision of successful treatments to the communities and countries where they had been trialled. Yet tension and disagreement raged about the extent to which the Declaration should require specific solutions to this problem or advocate for justice in broad terms. Meanwhile, alternative approaches to the problem of inequalities in international drugs trials were being formulated, including the Cordoba Declaration from the Latin American and Caribbean Bioethics Network, which promoted a more radical approach that ultimately influenced the UNESCO bioethics committee report on benefit sharing.[5]

When the 2013 version of Helsinki emerged, it formulated post-trial obligations for research sponsors and host country governments, as well as for researchers (Mastroleo, 2016). By this time, however, while many countries had referenced earlier versions of Helsinki in their domestic laws, most simply retained these original references. The stepping back from the later versions was more overt among the US regulators, with the Food and Drug Administration (FDA) having by the mid-2000s signalled its intention to recognise only the 1989 version.[6] But in Europe too, various versions of Helsinki were referenced in domestic laws, as well in EU laws and policies.[7] Ultimately, the fraught process of revising the Declaration came to be associated with the fracturing of a fragile consensus around international trial ethics. Nevertheless, even as challenges to some of the provisions of the Declaration raged, an allegiance to the principles it set out remained essentially universal.

The CIOMS guidelines on research ethics

We now turn to a set of guidelines for ethics in health-related research whose development was facilitated by policymakers in international organisations, for whom there was an imperative for such a code to be developed for application in broader settings including the Global South. The Council for International Organizations of Medical Sciences (CIOMS), established in 1949 by WHO and UNESCO to facilitate international exchange of views on medical science, had been functioning for some time within the nexus of international organisations. The development and safe use of medicines had been a primary theme for the organisation's meetings, and many of its working groups focused on various aspects of pharmacovigilance. Following the adoption of the goal of Health for All, we began to see the emergence of a broader field of health-related research. Now CIOMS would be mobilised to take forward a policy-oriented set of ethics guidelines for international research. To a considerable extent, the CIOMS research ethics guidelines first emerged from dissatisfaction at an international level with lack of agreement between countries about how the Helsinki Declaration should be implemented – as well as need for more diverse stakeholders to have an input (Schmidt, 2020). Officially, the guidelines are presented in terms of applying the Helsinki principles to research in low- and middle-income countries. However, as will be indicated later, they are characterised by an engagement with broader dimensions of justice and fairness in relation to the global health agenda.

The CIOMS guidelines were first published in 1982 under the title Proposed International Ethical Guidelines for Biomedical Research Involving Human Subjects.[8] From the outset, the guidelines aimed to provide 'internationally vetted ethical principles and detailed commentary on how universal ethical principles should be applied, with particular attention to conducting research in low-resource settings' (CIOMS, 2016: viii). From the early 1990s, the CIOMS working group recognised the substantive concerns that were emerging about the conduct of clinical research in low-resource settings, and addressed these as a central concern. Whereas earlier versions were directed at clinicians and focused on biomedical research, later editions addressed the wider research community of researchers, funders and sponsors, and encompassed approaches such as the use of data in health-related research. While still concerned with the 'classic activities that fall under health-related research', the most recent guidelines evidence an engagement with a wide range of issues such as capacity building, community engagement, vulnerability, and research in disaster settings (CIOMS, 2016: ix). Among the significant innovations in 2016 was the emphasis on *social value* alongside scientific value as necessary criteria for ethical research.[9] It was no longer considered sufficient justification for scientists to demonstrate that they wanted to further their understanding and contribute to knowledge in the scientific community: proposals undergoing ethics review must now be evaluated in terms of their contribution to social good as well. The concept of social value as a threshold criterion for ethics evaluation has generated extensive debate, and is not uncomplicated to apply in practice (Rid and Shah, 2017; Wenner, 2017). However, this anchoring of ethics review to wider social contexts is a significant achievement in terms of broadening ethics evaluation to include a consideration of priorities in health research (Barsdorf and Millum, 2017; Lutge et al, 2017).

Although it is beyond the scope of this chapter to map all the many international codes and frameworks of health research ethics, two other such documents that have considerable currency will be mentioned here.

We turn next to the Good Clinical Practice Guidelines (GCP; ICH, 2016), which provide guidance for the conduct of pharmaceutical studies. These guidelines were first issued in the form of a Note for guidance on Good Clinical Practice, under an agreement between the EU, the US and Japan under the auspices of the International Conference on Harmonisation of Technical Requirements for Human Use (ICH, 2016).[10] The GCP, as its name suggests, aims to define the standards and organisational arrangements for the conduct of clinical pharmaceutical trials. It is significant that review by an ethics committee is stated to be one of the criteria for good conduct, as is properly obtained informed consent. However, GCP provides a 'thinner' description of ethical responsibilities than CIOMS or Helsinki, and indeed this is consistent with the ICH's goal of achieving equivalence between pharmaceutical trials being conducted in different countries, so as to facilitate mutual recognition of research findings and, ultimately, of medicinal products. While it achieved its status gradually with the allegiance of professional bodies, GCP is now recognised to be *the* code for pharma researchers and trialists internationally and as such is widely used to inform professional training. In addition, GCP was adopted as guidance by regulatory agencies and eventually achieved a quasi-regulatory standing through being referenced in key laws and regulations in the fields. Nevertheless, GCP is formally a standard-setting document. Given its tight parameters and specificity, the code describes procedures rather than articulating moral principles: as such, it is less influential beyond the professional domain of pharmaceutical research than Helsinki and CIOMS.

Finally, the Oviedo Convention or, to give it its full name, the Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine, must be acknowledged as central to the canon. This framework document was drafted by bioethics experts for the Council of Europe, and signed and ratified by some Council members and by a number of non-member countries.[11] The document sets out broad, general principles, based on the minimum ethical standards accepted across the membership at the time. The scope of the document is extensive, ranging as it does across developments in medical science including research, and the principles set out echo those of Helsinki, emphasising human dignity, primacy of the person, and equity. Correspondingly less attention is paid to the specificities of research. Given its philosophical sweep, the Oviedo framework does not provide guidance for the day-to-day situations faced by researchers and research stakeholders. Nor is there recourse for individuals to claim that their rights under the Convention have been breached, as the Convention works at the state level. Nevertheless, the Convention is seen (especially by lawyers) as a unique achievement in the field of bioethics – being binding on the states that have signed it – and is considered to 'represent a historical step in the elaboration of a common European biomedical law' (Andorno, 2005: 135). However, the absence of a consensus and a fuller set of signatories and ratifications seems to undermine this claim, and even within its own terms, the Convention must be considered only a partial success.

Even bearing in mind that the description given is highly selective in relation to the many normative documents in health-related research, it serves to emphasise the extent to which ethics guidance has become codified in this field. We should not, of course, confuse the proliferation of ethics guidelines with an absence of harm in health-related research. As early as 1999, Arthur Kleinman wrote that codified ethics norms 'risk irrelevance' when applied without a deep understanding of the context of human experience in such research (Kleinman, 1999: 72). Ironically, there is little research about how ethics guidelines impact on or shape research design and conduct in the field. Indeed, it has been convincingly stated that, to a considerable extent, 'practice overwhelms ethics' under the pressure of commercial imperatives in the clinical trials industry (Petryna, 2009: 31). It has to be accepted then that such guidance is only a resource for the various actors in the field. To the extent that ethics codes are meaningful and influential for individual actors, they may help to shape a culture of research integrity. However, it is difficult to locate any research exploring this empirically. In the next section we shall begin to evaluate the mechanisms by which these codes have achieved some influence.

Theorising modes of influence

The ethics codes and declarations referred to are sometimes seen as pillars of research ethics regulation in the health field.[12] There is no doubt that Helsinki, having been pre-eminent in the formative decades of the post–Second World War years, became effective in legal terms through its incorporation into many national and international laws and codes. However, there is also a wider political dimension to the Declaration: with Helsinki, we can discern a threshold after which the self-regulation of research by medical professionals without reference to others' interests became unacceptable. It has to be conceded that this was a slow-moving paradigm shift that took place over some decades, but still, a momentous one. Nevertheless, controversies have characterised its progress over time, and the status of the Declaration itself as the primary point of reference for research ethics was, in practice, gradually relinquished by key stakeholders.

By contrast, the CIOMS ethics guidelines drew inspiration from community engagement approaches and work being undertaken in low-resourced countries, as well as from human rights. They became – and remain – a point of reference for applying ethics principles, notably in ethics review at WHO HQ and regional offices. Furthermore, the influence of these guidelines may be discerned in the many research partnerships in place with WHO, which include other UN organisations such as UNICEF and the plethora of multilateral organisations and public–private partnerships through which major private donors channel funds, as well as more traditional arrangements with governments, ministries of health and collaborating academic organisations: when members of this extensive network partner with WHO to fund, sponsor or collaborate on research, the CIOMS guidelines function as a significant guiding tool (Ward et al, 2018). Given that these guidelines are primarily used in the evaluation of research at an organisational level, the concept of stewardship will be useful here. According to Laurie et al, regulatory/ethics stewardship operates both at national and at supranational levels where 'ethics stewards [act] with a stated mandate of protection first then the promotion of ethically sound, scientifically robust research after due deliberation' (Laurie et al, 2018: 342). They observe that such stewards include the managers and administrators of ethics committees, as well as their more senior members, who help researchers to navigate the many ethical requirements that they face and, sometimes, to reconcile them. Being a key point of reference within international organisations, the CIOMS guidelines have achieved significant influence on the evaluation of research in a globalised world.

In the meantime, guidelines for good clinical practice aimed in the first instance to influence professional practice in the field so as to facilitate the exchange and mutual recognition of findings. The GCP was eventually referenced in relevant laws in many countries, and referred to in European law including the Clinical Trial Directive (2001/20/EC) and subsequent Regulation (536/2014).[13] On the other hand, GCP is not so widely known about or used by other stakeholders precisely because it does not engage with wider social issues. Our fourth code, the Oviedo Convention, being an international treaty, is an exception among the guidelines already discussed. However, it is the additional protocols to Oviedo – especially those addressing ethical problems in more specific terms in relation to genetics, medically assisted reproduction and cloning and the recommendation on research using biological materials of human origin – that provide more useful commentary on ethical problems encountered by researchers.[14]

Each of these codes and declarations has had successes in terms of influencing policies and laws, yet none is hegemonic. Thinking about the mechanism by which they have achieved influence, we can see Helsinki in terms of 'soft law'. Soft law is a term used to describe the process by which normative declarations, codes and guidelines which – albeit not themselves regulations or treaties – become embedded in national and international laws and policies.[15] This can be extended to think of the influence of such declarations on public policy more broadly.[16] CIOMS guidelines were targeted at the process of ethics review at the international level, and were also adopted by some national ethics committees alongside other guidance. Thus, they found their mode of influence to some extent through stewards within these organisations. Third, while the GCP was eventually referenced in both national and EU laws, it had already become the basis for training for clinical investigators and their associates in pharmaceutical studies. This underlines the significance of training in influencing ethical practice in research. Finally, Oviedo has the status of international research regulation, and exerts influence primarily through governments and courts. Leaving aside the problem that not all the countries involved in its drafting have signed or ratified it, it is seen as symbolically important, even posited as the jewel of international regulation. However, because of its breadth and legal tone, its guidance may be less likely to be used directly by researchers and stakeholder organisations. While there are differences in the provisions made by the various codes/guidelines, efforts have been made in recent years to reconcile these, especially so for CIOMS and Helsinki. A feature of this landscape today is that each of the guidelines makes reference to some or all of the others, which has added to their collective influence across the diverse sectors of health-related research.

Production, representativeness and legitimacy

The themes that have been traced suggest that in future, international codes for research ethics – rather than being conceived of in static terms – might seek to achieve cascades of influence. We have seen that embedding research ethics guidance in hard law (as in Oviedo) was the exception rather than the rule: other mechanisms of influence have included 'soft law', dissemination through professional training, and application of the guidelines through quasi-regulatory processes and via ethics stewardship in key organisations.

While it can be powerful, ownership of a code by a single profession may be associated with problems in terms of its wider relevance and perceived legitimacy. This was exemplified by the process through which the Declaration of Helsinki was decided and revised: medical perspectives dominated the discussions throughout – albeit that for the latest revision, the draft was opened to public consultation. The WMA's dominance of the 'discursive space' throughout several decades made it difficult for non-medical stakeholders, such as policymakers, patient organisations and other non-governmental organisations (NGOs), to influence its redrafting (Schmidt, 2020: 121). Furthermore, its recommendations mainly targeted physicians, leaving little scope for the other emerging players in the fields of health research.

Certainly, the drawing up of wide-reaching moral codes by a narrow constituency was not unusual for its time. Indeed, the GCP was convened by professional bodies, and the text for the Oviedo drafted by bioethics experts on behalf of ministers. There is little on record about the involvement of non-experts in the development of these guidelines. For the CIOMS guidelines too, the extent to which non-professionals were involved has been extremely limited.[17] From these histories, it is clear that there has been limited representation of regional perspectives in the drawing up of the established guidelines. Organisations such as WHO, WMA and CIOMS have tended to devise transnational norms based on common perspectives managed by a centralised secretariat in consultation with professional organisations.

Discussion

The argument that is being developed here is that which organisation 'owns' the document is not as important for its legitimacy, but rather how it is produced. Taking a critical perspective, it can be said that strong control from the centre may in fact limit the influence and reach of these codes in the context of a contemporary research agenda characterised by interdisciplinarity, pluralism and (some) lay involvement.

The concept of mode 2 knowledge production, influential in science and society, is relevant here: in ethics as in science, the 'context speaks back', to coin a phrase from Nowotny et al (2011: 50), and this has to be explored and understood. There has been critical engagement with the CIOMS guidelines in particular, and proposals have been made for their application and interpretation in different regions, localities and contexts.[18] Bandewar, writing from the viewpoint of bioethics in India, puts it as follows:

> It is noteworthy that a number of countries draw on international ethics guidance such as CIOMS and DoH despite their known limitations. One reason seems to be the lack of resources to develop guidance documents sensitive to the local context. However, a few ethics codes developed locally by ethnic groups, such as the code developed by the San people of South Africa, question the dominance of international guidance. (Bandewar, 2017: 140)

Global ethics guidelines have been described – and decried – as sustaining a dynamic of ethical imperialism. More specifically, Israel describes this process as 'exporting principlism' (Israel, 2017: 7). Such guidelines aim to mitigate the 'ethical variability' that can allow serious harms to be imposed on research subjects in diverse countries (Petryna, 2009: 10). Yet the intensity of ethical scrutiny associated with the application of the guidelines may undermine researchers' freedom of manoeuvre to address complex ethical issues as they arise in the field (Posel and Ross, 2015).

How, then, to sum up the status of this paradoxical genre? Drawing on perspectives from sociology and social anthropology, global ethics guidelines can be viewed as policy instruments through which international organisations aim to influence diverse national, regional and local ethics bodies and communities (Feldman, 2011). However, the notion of global codes produced at the centre is fraying as we move to more dynamic relationships between the centre and other regions of the world – and other actors. The story of health-related research ethics points to this admittedly imperfect genre having had a profound influence on research culture in diverse fields. However, these codes emanated from elite professional organisations and transnational institutions and today are subject to criticism on precisely these grounds. Whereas established guidelines such as those produced by WMA, WHO and CIOMS tend to emphasise transnational values, other groups developing ethics guidance in response have emphasised a reflexive process and engagement with regional priorities as important dimensions. Initiatives that are international rather than transnational will surely be pertinent to the relevance and perceived legitimacy of ethical guidelines in future.

Could ethics guidelines and codes for non-health research provide alternative routes to fairer research and justice including in the Global South? An example of an international approach is the Global Code of Conduct for Research in Resource-Poor Settings produced by the Trust EU project is timely in this context. This code addresses community priorities and perspectives, while drawing on the concept of benefit sharing which has been under discussion for some years in the international institutions, and local perspectives from other regions.[19]

The following lessons might be drawn from the history of health-related research and applied to the process of developing and revising guidelines in other fields: these initiatives should endeavour to balance the protection of research subjects with promotion of research, and to clarify the source of their authority (Smith and Weinstock, 2018). It will be important for any new code on non-health-related research to build capacity to connect with diverse institutional structures and communities beyond 'the centre' from the outset. Training and opportunities for feedback should go hand in hand with the publication of the document itself. In addition, revisions can be an opportunity to open up the discussion and ownership further, and this process should be carefully planned.

To these suggestions for process, we can add a comment on the text of such a document itself: it is useful if guidelines have commentary as well as principles. This enables those who will use the document to refer to explanations of the underlying rationale for the proposed principles and rules. As has been observed, the publication of literature reviews and bibliographies that inform ethics guidelines would be helpful for those wishing to engage with them in more depth (Bandewar, 2017). At the same time, we can see from Helsinki especially that principles outlast rules. New codes for other research fields should seek to find a 'sweet spot' between high principles and specificity. Finally, such a code should be devised to be usable by diverse groups: if it is good to think from the point of view of researchers, evaluators and ethics 'stewards', it has more chance of achieving traction.

Conclusion

It is not necessary for a research ethics code to be a pillar of international law to achieve change. A better model for today's world may be to think in terms of achieving a cascade of influence, through research organisations, deliberative institutions and NGOs, and wider communities. While partnership with professional organisations and international institutions may be necessary to achieve influence today's world, authority lies not in Geneva or in Brussels, but in multiple movements and places. It is through dialogue with diverse communities of doers, thinkers and users of research – including ethics committees, community organisations and activists, as well as researchers – that ethics codes will achieve change.

Notes

[1] The ethics codes, guidelines and laws discussed in the chapter are all widely referred to by organisations concerned with health research at the international level. ('Codes' and 'guidelines' are arguably similar in this context, and these terms will be used interchangeably

in this chapter.) Drawing on contemporary anthropological approaches to the study of public policy, these can be seen as forming part of the '*dispositif*' or 'ensemble of practices, institutions, architectural arrangements, laws, administrative measures, scientific statements, philosophical propositions and morality that frame a disciplinary space' (Foucault, 1980; Shore and Wright, 2011: 11). From this perspective, it will be useful to explore and analyse how these various normative documents frame ways of doing research.

[2] See Schmidt et al (2020) for recent historical perspectives on the Declaration of Helsinki; and Lederer (2004) for a detailed analysis of the drafting of the early versions of the Declaration.

[3] The Declaration of Alma-Ata (1978) was made at an International Conference on Primary Health Care at Alma-Ata in the USSR, and the Health for All strategy was subsequently adopted by WHO: https://www.who.int/publications/almaata_declaration_en.pdf [accessed 7 September 2020].

[4] See Thiers et al (2008) for a contemporaneous view of trends in the globalisation of clinical trials. Among the responses to these concerns was the European Group on Ethics (2003) Opinion Nr 17 on ethical aspects of clinical research in developing countries. The issue of post-trial access remains live: see Iunes et al (2019).

[5] See Report of the International Bioethics Committee on the Principle of the Sharing of Benefits, at: https://unesdoc.unesco.org/ark:/48223/pf0000233230

[6] In 2008, the FDA in the US published its own 'Final rule' *on Human subject protection; foreign clinical studies not conducted under an investigational new drug application*, which referred to the Good Clinical Practice code (GCP) as the point of reference for the ethics of such studies. See Moreno (2020) for more on the 'American Stamp' on research ethics.

[7] For example, the 'Clinical Trials Directive' 2001/20/EC refers to the 1996 version of the Helsinki Declaration, and the subsequent Regulation EU No 536/2014 to the 2008 version.

[8] The most recent version of the guidelines together with an overview of their publication history is provided at: https://cioms.ch/working_groups/bioethics/

[9] Guideline 1: Scientific and social value and respect for rights in the CIOMS (2016) guidelines, available from: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

[10] The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) was created in April 1990 at a meeting in Brussels. Since 2015, in recognition of the globalisation of research, it has expanded its scope to include other countries and regions.

[11] Among the Council members that did not sign are Germany, Ireland, Malta and the UK, objecting to significant clauses as either too restrictive (as was the case for the UK) or unduly permissive (the concern expressed in Germany). Others including Italy, the Netherlands, Poland and Sweden signed but did not subsequently ratify the Convention, in other words did not implement its provisions in national law. The current status of signatures and ratifications can be viewed at: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/164/signatures [accessed: 18 August 2020].

[12] Sprumont writes of three pillars of modern research regulation: human rights legislation, pharmaceutical product regulation and guidance on professional standards (Sprumont, 2020). It is acknowledged that these have evolved over time; thus Helsinki is characterised as only later having expanded from a document concerned with professional standards into the broader terrain of human rights (Sprumont, 2020: 243).

[13] The EU's 'Good Clinical Practice Directive' (2005/28/EC) further reinforced reference to the GCP standards.

[14] ETS No 168 (1998) Additional Protocol on the Prohibition of Cloning Human Beings; ETS No 186 (2002) Additional Protocol on Human Rights and Biomedicine Concerning Transplantation of Organs and Tissues of Human Origin; CETS No 203 2008) Additional Protocol Concerning Genetic Testing for Health Purposes; and Recommendation CM/Rec(2016) 6 of the Committee of Ministers to Member States on research on biological materials of human origin.

[15] For a succinct definition of soft law versus hard law, see: https://www.ecchr.eu/en/glossary/hard-law-soft-law/

[16] See Sekalala (2018) for a consideration of soft law and bioethics in global health.

[17] Based on CIOMS documentation, the revision of the most recent guidelines was undertaken by a working group of ten members and one chair, with the input of four advisors. As Schuklenk points out, the majority were from the Global North, and whole regions of the world (for example, China, and Central America) went unrepresented, as did countries with extensive experience of hosting trials, notably South Africa. Furthermore, one individual on the working group is said to have represented the perspective of research participants. A modest consultation was undertaken to inform the revision process. However, in Schuklenk's view the claim that the principles in the document are regarded as universal is seriously undermined by their 'demonstrably problematic production process' (Schuklenk, 2017: 172).

[18] See, for example, Aarons (2019) for a critique of CIOMS guideline 20, which makes recommendations for pre-evaluation of proposals to be mobilised in emergency or disaster situations: informed by the situation in the Caribbean, Aarons proposes alternative plans for such situations. See also Barugahare and Kutyabami (2019), who question the absence of East African voices in the revision of the CIOMS guidelines, but also observe that, given their generality, there is scope for local and regional interpretation and application of these.

[19] On international perspectives on benefit sharing, see especially Article 15 – Sharing of benefits – and Article 21 – Transnational practices of the UNESCO (2005) Universal Declaration on Bioethics and Human Rights. For a pluralist and multicentred perspective, see the Trust EU project at: http://trust-project.eu/. A grassroots perspective is offered by The San Code of Ethics (available on the Trust EU website).

References

Aarons, D. (2019) 'Addressing the challenge for expedient ethical review of research in disasters and disease outbreaks', *Bioethics*, 33(3): 343–6.

Andorno, R. (2005) 'The Oviedo Convention: a European legal framework at the intersection of human rights and health law', *Journal of International Biotechnology Law*, 2(1): 133–43.

Bandewar, V.S. (2017) 'CIOMS 2016 (editorial)', *Indian Journal of Medical Ethics*, 2(3): 138–40.

Barsdorf, N. and Millum, J. (2017) 'The social value of health research and the worst off', *Bioethics*, 31(2): 105–15.

Barugahare, J. and Kutyabami, P. (2019) 'Nature and history of the CIOMS International Ethical Guidelines and implications for local implementation: a perspective from East Africa', *Developing World Bioethics*, 20(4): 175–83.

Bonah, C. and Schmaltz, F. (2020) 'From Nuremberg to Helsinki: the preparation of Helsinki in the light of the prosecution of medical war crimes at the Struthof Medical Trials, France, 1952–4', in U. Schmidt, A. Frewer and D. Sprumont (eds) *Ethical Research: The Declaration of Helsinki, and the Past, Present, and Future of Human Experimentation*, New York: Oxford University Press, pp 69–100.

Buse, K. and Hawkes, S. (2015) 'Health in the sustainable development goals: ready for a paradigm shift?', *Globalization and Health*, 11: art 13, available from: https://doi.org/10.1186/s12992-015-0098-8 [accessed 28 January 2022].

CIOMS (Council for International Organizations of Medical Sciences) (2016) *International Ethical Guidelines for Health-related Research Involving Humans* (4th edn), Geneva: CIOMS.

Dankwa-Mullan, I., Rhee, K.B., Stoff, D.M., Pohlhaus, J.R., Sy, F.S., Stinson, N., Jr and Ruffin, J. (2010) 'Moving toward paradigm-shifting research in health disparities through translational, transformational, and transdisciplinary approaches', *American Journal of Public Health*, 100(S1): S19–S24.

European Group on Ethics (2003) *Opinion Nr 17 on Ethical Aspects of Clinical Research in Developing Countries*, Luxembourg: Office for Official Publications of the European Communities.

Feldman, G. (2011) 'Illuminating the apparatus: steps towards a nonlocal ethnography of global governance', in C. Shore, S. Wright and D. Però (eds) *Policy Worlds: Anthropology and the Analysis of Contemporary Power*, New York: Berghahn, pp 32–49.

Foucault, M. (1980) 'The confession of the flesh' [interview, 1977], in C. Gordon (ed) *Power/Knowledge Selected Interviews and Other Writings, 1972–1977*, New York: Pantheon, pp 194–228.

ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) (2016) *Guideline for Good Clinical Practice*, E6(R2), Geneva: ICH.

Israel, M. (2017) 'Ethical imperialism? Exporting research ethics to the Global South', in R. Iphofen and M. Tolich (eds) *The SAGE Handbook of Qualitative Research Ethics*, London: Sage, pp 89–102.

Iunes, R., Uribe, M.V., Torres, J.B., Garcia, M.M., Alvares-Teodoro, J., de Assis Acurcio, F. and Guerra, A.A., Jr (2019) 'Who should pay for the continuity of post-trial health care treatments?', *International Journal for Equity in Health*, 18: art 26, available from: https://doi.org/10.1186/s12939-019-0919-0 [accessed 28 January 2022].

Kleinman, A. (1999) 'Moral experience and ethical reflection: can ethnography reconcile them? A quandary for "the new bioethics"', *Daedalus*, 128(4): 69–97.

Laurie, G., Dove, S., Ganguli-Mitra, A., Fletcher, I., McMillan, C., Sethi, N. and Sorbie, A. (2018) 'Charting regulatory stewardship in health research: making the invisible visible', *Cambridge Quarterly of Healthcare Ethics*, 27(2): 333–47.

Lebov, J., Grieger, K., Womack, D., Zaccaro, D., Whitehead, N., Kowalcyk, B. and MacDonald, P.D.M. (2017) 'A framework for One Health research', *One Health*, 3: 44–50.

Lederer, S. (2004) 'Research without borders: the origins of the Declaration of Helsinki', in V. Roelcke and G. Maio (eds) *Twentieth Century Ethics of Human Subjects Research: Historical Perspectives on Values, Practices, and Regulations*, Stuttgart: Franz Steiner, pp 199–217.

Lurie, P. and Wolf, S.M. (1997) 'Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries', *New England Journal of Medicine*, 337(12): 853–6.

Lutge, E., Slack, C. and Wassenar, D. (2017) 'Defining and negotiating the social value of research in public health facilities: perceptions of stakeholders in a research-active province of South Africa', *Bioethics*, 31(2): 128–35.

Mastroleo, I. (2016) 'Post-trial obligations in the Declaration of Helsinki 2013: classification, reconstruction and interpretation', *Developing World Bioethics*, 16(2): 80–90.

Moreno, J.D. (2020) 'The Declaration of Helsinki and the "American Stamp"', in U. Schmidt, A. Frewer and D. Sprumont (eds) *Ethical Research: The Declaration of Helsinki, and the Past, Present, and Future of Human Experimentation*, New York: Oxford University Press, pp 351–65.

Nowotny, H., Scott, P. and Gibbons, M. (2011) *Re-thinking Science: Knowledge and the Public in an Age of Uncertainty*, Cambridge: Polity Press.

Petryna, A. (2009) *When Experiments Travel: Clinical Trials and the Global Search for Human Subjects*, Princeton, NJ: Princeton University Press.

Posel, D. and Ross, F. (2015) 'Opening up the quandaries of research ethics: beyond the formalities of institutional ethical review', in D. Posel and F. Ross (eds) *Ethical Quandaries in Social Research*, Cape Town: HSRC Press, pp 1–26.

Rid, A. and Shah, S.K. (2017) 'Substantiating the social value requirement for research: an introduction', *Bioethics*, 31(2): 72–6.

Schmidt, U. (2020) 'In the absence of alternatives: the origins and success of the Declaration of Helsinki, 1947–82', in U. Schmidt, A. Frewer, and D. Sprumont (eds) *Ethical Research: The Declaration of Helsinki, and the Past, Present, and Future of Human Experimentation*, New York: Oxford University Press, pp 101–30.

Schmidt, U., Frewer, A. and Sprumont, D. (eds) (2020) *Ethical Research: The Declaration of Helsinki, and the Past, Present, and Future of Human Experimentation*, New York: Oxford University Press.

Schuklenk, U. (2001) 'Helsinki Declaration revisions', *Indian Journal of Medical Ethics*, 9(1), available from: https://ijme.in/articles/helsinki-declaration-revisions/?galley=html

Schuklenk, U. (2017) 'Revised CIOMS research ethics guidance: on the importance of process for credibility', *Indian Journal of Medical Ethics*, 2(3): 169–72.

Sekalala, S. (2018) *Soft Law and Global Health Problems: Lessons from Responses to HIV/AIDS, Malaria and Tuberculosis*, Cambridge: Cambridge University Press.

Shore, C. and Wright, S. (2011) 'Conceptualising policy: technologies of governance and the politics of visibility', in C. Shore, S. Wright, and D. Però (eds) *Policy Worlds: Anthropology and the Analysis of Contemporary Power*, New York: Berghahn, pp 1–25.

Smith, M. and Weinstock, D. (2018) 'Political legitimacy and research ethics', *Bioethics*, 33(3): 312–18.

Sprumont, D. (2020) 'Research ethics regulation: rules versus responsibility', in U. Schmidt, A. Frewer, and D. Sprumont (eds) *Ethical Research: The Declaration of Helsinki, and the Past, Present, and Future of Human Experimentation*, New York: Oxford University Press, pp 241–81.

Thiers, F.A., Sinskey, A. and Berndt, E. (2008) 'Trends in the globalisation of clinical trials', *Nature Reviews Drug Discovery*, 7(1): 13–14.

Ward, C.L., Shaw, D., Sprumont, D., Sankoh, O., Tanner, M. and Elger, B. (2018) 'Good collaborative practice: reforming capacity building governance of international health research partnerships', *Globalization and Health*, 14: art 1, available from: https://doi.org/10.1186/s12992-017-0319-4

Wenner, D.M. (2017) 'The social value of knowledge and the responsiveness requirement for international research', *Bioethics*, 31(2): 97–104.

5

Interdisciplinary perspectives on ethics and integrity in Europe: acknowledging differences to foster mutual understanding

Eleni Spyrakou, Panagiotis Kavouras, Vassilis Markakis,
Matias Barberis Rami and Costas A. Charitidis

Introduction

Any discussion about a new framework on research ethics and research integrity (RE&RI) should take into consideration the existing ethics guidelines and codes in the various fields of science. In fact, there are several problems that need to be considered when creating a proposed framework. For example: already existing diverse codes, guidelines and frameworks; interdisciplinary differences; varying regulations; various review systems; institutional infrastructures that do not support good practice of research; innovations that challenge existing ethics; and new entrants to the field raising 'old' problems. A framework on RE&RI should be comprehensive and based on agreed values, principles and standards. Agreement is also needed by all stakeholders in constructing and applying such a framework. Full dissemination of the framework is necessary, including training and continuing professional development. Furthermore, due to social changes and innovations in technologies and research methods, such a framework should be regularly updated.

The PRO-RES project (aiming to PROmote ethics and integrity in non-medical RESearch) that stimulated this volume held workshops and interviews with relevant stakeholders, in particular researchers, science/technology experts, regulators, funders/research councils and policymakers which showed that researchers, plus other people with different roles and types of involvement in scientific research, face difficulties in orienting themselves with respect to the variety of different codes, guidelines and existing frameworks, as well as ethical review and evaluation systems, and different institutions. The aforementioned workshops, through the participants' expertise, covered a wide range of topics, disciplines and research fields, including: ethics and research funding organisations (RFOs); technological innovation (cutting-edge research like nano, biotech and so on); privacy, surveillance and covert research; ethics in the context of intelligent decision support, autonomous systems, social media and bioinformatics, in short *data science*; research ethics in disaster and conflict settings; ethics in responsible research and innovation (RRI); ethics in finance and economics research; ethics, social mining and explainable artificial intelligence; ethics in policy advice (with focus on economics and finance research); and ethics, integrity and qualitative research methods. Natural sciences, engineering, social sciences and humanities were all represented in the workshops. The workshops' aggregated conclusions have been reported in the PRO-RES project's deliverable D2.1 'Thematic priorities report' which was compiled based on the individual, concluding reports of the workshops.[1] Given the pluralism of represented scientific disciplines and the various backgrounds of the participants coming from different research environments, these workshops provided the appropriate occasions for thematising notable issues related to RE&RI that can be regarded as interdisciplinary and of particular interest for a variety of stakeholders. These thematised issues are the subject of this chapter.

Institutional structures and constraints

Within the European research environment, national regulations tend to differ and, at times, can be mutually inconsistent even among EU Member States. What is more, even within the same country there are very different attitudes by disciplinary associations towards research ethics, ranging from the non-recognition of the issue to very detailed codes. In the past, ethics in the conduct of research had conventionally been seen merely as a matter of the virtue of the individual researcher rather than institutional structure. This has changed over the years and nowadays it is recognised that Research Performing Organisations (RPOs) and RFOs, plus national and international organisations and think tanks, share equally with individual researchers the responsibility of enabling and promoting the conduct of responsible research, the proper use of research results and the implementation of good practices. However, the poor reflection and awareness among researchers of the motivations, interests and stakes involved in their research activity, including its wider ethical implications for the common good, remains an important issue related to the established research culture in many countries and research environments. This poor awareness of the importance of ethical issues, whether they are ignored or just considered an onerous bureaucratic obligation, is the result, not only of the individual researcher's lack of interest or training on RE&RI issues but, also, of the structural weaknesses that research institutions have developed over the years. Although, the importance of the assessment of research projects for RE &RI is more and more now recognised, various past developments have led to the current situation characterised by complexity of regulations and procedures and, to a great extent, typical, conventional compliance with ethics standards.

Despite the fact that the pursuit of knowledge and genuine curiosity have always been the main drivers of research endeavours, any research nowadays is a complex enterprise involving different agents, with various interests, motives, obligations and commitments, often contradictory. Hence, ethics assessments are crucial and, simultaneously, it is necessary to keep a balance between freedom of science and ethical principles. However, there are many constraints that explain the acknowledged complexity and need to be taken into consideration. For instance, certain orientations of the current institutional infrastructure of research organisations, such as the pressing drive to publication in high-impact-factor journals, can induce researchers to make mistakes or forfeit their integrity. Furthermore, there seems to be an excess of rules and prescriptions that, if applied too literally, can obstruct research activities. For example, in some institutions there are excessively strict or inappropriate procedures discouraging research in difficult areas or making important social and humanities research virtually impossible. Regulating research often has additional functions hanging on to it, which are not always appropriate or within scope, for example applications being rejected supposedly due to ethics concerns when the real decision is a matter of censorship or reputation management (Hedgecoe, 2016; Carey, 2019).

The participants of the PRO-RES series of workshops identified the following issues as some of the most significant ones regarding the institutional structures and constraints:

1. The question of defining and categorising research and its purpose, which is sometimes unclear. Different disciplines have different standards and possibly different ethical requirements.
2. The issue of cultural differences in ethical perception.
3. The issue of the role and necessity for approval of the ethics committees for the research. Proceeding with a research project without ethics approval is common in some disciplines, for instance software engineering. In addition, it is not legally required. There are some advantages to not regulating the process of ethics in research too much. Outsourcing the ethics responsibility to the ethics board is a risk but it is often a common practice. There is, also, a risk of creating a 'nanny state', which is overregulated with too many rules, codes and so on, to follow.
4. Peer-review processes as the core academic mechanism could become an opportunity for reviewing ethical issues also.
5. Ethics approval is considered as a 'yes' and 'no' outcome, but it should be a process which helps to refine ethical considerations.
6. It is important to make a distinction between academic and professional researchers. In economics and finance, for instance, much research is

carried out by associations, private companies, think tanks and other entities. The question is whether the same restrictions apply to them. There is, also, the risk of displacement, with researchers moving to certain institutions based on what they can do there.

7. Research ethics assessment has already become loaded with other functions, such as the census function, liability function and reputation management function.
8. A lack of effective training at all levels, including supervisors and senior academics or managers, can perpetuate bad institutional practices. If there is no national agency supervising research institutions, organisations are tempted to cover up poor ethics or misconduct and to think that this is the best way to protect their interests. This is aggravated by the tendency to adopt quantitative rather than qualitative measures in performance evaluation because they are quicker and cheaper to administer.
9. It is extremely important to have dedicated staff that researchers can talk to, in confidence, about questions/concerns/issues affecting integrity or ethics. A formal channel for complaints, leading to official investigations, is desirable but it is not sufficient. This needs to be supplemented by confidential counsellors and ombudspersons who can advise and support complainants and whistle-blowers.
10. A flexible and transparent approach avoiding rigid structures is advisable. This should make the ethics decision-making process as explicit as that of the scientific process. To that direction, tools and resources should be provided to promote research ethics and support researchers in taking full responsibility for their work. It is, therefore, recommended not to provide a checklist to be ticked slavishly or pre-compiled formats, that is, based on a mainly formal approach, in order to favour raising the awareness of researchers about their responsibility.
11. Emerging technologies and interdisciplinary research outcomes call for mechanisms and procedures foreseeing and anticipating future ethical challenges.

Part of the institutional structures and constraints is the way that ethics appraisal procedures are set and the way that research ethics committees (RECs) and institutional review boards (IRBs) work, which will be defined in the following section.

Existing ethics appraisal procedures and the ways RECs/IRBs work

More and more accepted is the idea that all stakeholders involved in research and related to its products, whether directly, such as researchers, RPOs and RFOs, or indirectly, such as publishers, policymakers, policy advisors and think tanks using scientific results, need to have established ethics appraisal procedures and experts conducting this kind of assessment. This acceptance derives from a broader discussion of, and agreement on, the responsibility of scientists, academia, industry and policymakers not just to pursue the 'no harm' perspective but, also, to bring to the core ethical considerations a richer interpretation of responsible research, confronting issues of societal relevance. For this reason, an exploratory approach to ethics that would proactively look for new ethically relevant questions pertaining to science is being broadly adopted. Introducing the notion of broader benefit of research in the evaluation process together with the institutionalisation of ethics and responsibility in a co-creation perspective is becoming the main tendency among institutions related to scientific research. In this way, ethics appraisal procedures can become integrative and not considered 'just' paperwork.

On various occasions within the PRO-RES workshops, it was highlighted how the application of standard ethical review procedures, not originally designed for certain disciplinary fields, can produce paradoxes that jeopardise the pursuit of those same ethical aims. For this reason, one of the main objectives of a proper RE&RI framework should be to maintain the balance between promoting general ethical principles applicable to all types of research and across different disciplines, while maintaining those elements that recognise the specific identity of each discipline. In particular, one main recurring issue is the suitability of the ethics appraisal procedures applied in medical research for broader application in non-medical research. Ethics appraisal procedures in non-medical research are predominantly based on a medical model, for example, the risk assessment process in social sciences draws on biomedical ethics. Beyond social sciences, biomedical ethics also informs other ethics guidance across other scientific disciplines, like computer science. While this has been the norm, it raises a challenge as the biomedical ethics model seems not to fit always with qualitative research in the digital world. One common complaint among social scientists is that understanding people and contexts are regarded as trivial issues, while technical/technological research and biomedical research are generally perceived as being morally good. This perception legitimises biomedical research over non-medical research and has inadvertently engendered a reliance on the biomedical ethics review. Thus, it is important to focus our discussions on questioning this reliance on biomedical ethics models and exploring alternatives which might be better suited for non-medical, particularly social science research (see Chapter 2 in this book).

However, in exploring an alternative model that might be consistent with non-medical research, it is pertinent to separate the 'why' of the research from the 'how'. That is, the justification for non-medical research needs to be clearly articulated, but this is different from how the research will be conducted. Hence there is a need for reflexivity to provide clear accounts of the process. Though qualitative research is considered subjective, this transparency improves its reliability and the integrity of the data. Also, we can still learn from the biomedical approach to conflict of interests to protect the integrity of qualitative research.

Furthermore, the existing ethics appraisal procedures and, in general, the governance/ethics frameworks applied by RPOs and RFOs need to face the challenge of collaborative research among researchers coming from different scientific fields and disciplines, as well as of the collaborative research among researchers coming from different countries and research environments/cultures. These ethics appraisal procedures are not always updated so as to deal with interdisciplinary research projects, for instance to deal with the cross-disciplinarity between digital and social science research, which raises challenges when seeking ethical guidance from institutional ethics councils/committees. What is more, inconsistency of ethics considerations within and across countries has significant implications for collaborative research. These are challenges that members of RECs and IRBs are called to resolve. Hence it is important that RECs and IRBs be staffed by experts coming from as many different scientific fields as possible, while having received proper training in dealing with these kinds of challenges, whether they have personal experience in interdisciplinary and international research collaborations or not. Thus, training on RE&RI seems to be crucial not only for researchers, but also for the experts conducting ethics assessments. Further to the establishment of committees and procedures to safeguard scientific integrity, 'it is equally important to foster an academic culture that understands and values integrity. Education and training are vital in this respect' (Deketelaere et al, 2019).

The collection, analysis and handling of data

Approaching the issue of data requires an understanding of the methodologies behind data collection, analysis and further handling. The variety of methodologies used in research poses different questions in terms of integrity, responsibility and ethics.

Regarding data collection, the main ethical issues that may arise are mostly regarding the General Data Protection Regulation (GDPR) and informed consent. These topics will be developed in depth in this chapter. However, the case of digital data is key in terms of collection. Along with the advancement of technology and the large quantity of data digitally available, further questions arise: first, is there a thin boundary about what is digital and non-digital? Once this question is answered, a key consideration may be the issue of how a researcher can recruit participants ethically in digital environments. How do we handle data in a responsible way? The point here is undoubtedly the privacy of data and how these are collected and used.

The second round of discussions dealt with data analysis. The theories and methodologies chosen are important to understand the ethical dimensions, as they may create high risk of biases. For example, while applying a specific theoretical framework, are we just trying to represent a group or advocate for them? The transparency of the research process requires the researchers involved to justify the decisions made along the whole development of the research, including theory. This is particularly crucial for qualitative research. What can qualitative researchers propose as an alternative to regulation via institutions as a basis for public trust? Much qualitative research can only survive on the basis of instantaneous ethical decision-making in the field and on strong assurances of the confidentiality of data. These are not easily reconciled with demands for institutional rather than personal accountability and for data transparency as the means of assuring integrity. What would the world lose if qualitative research was shut down by ethics regulation? What sort of regulatory framework would allow qualitative research to continue?

During the PRO-RES workshop on 'Ethics, Integrity and Qualitative Methods' the contributions made to qualitative research by arts-based methods was highlighted. More particularly, the arts-based methods discussed included the visual arts, creative writing, performative arts, textile arts, fashion, photography and film. In these discussions it was acknowledged that creative practice can overcome the confines of more conventional question-and-answer techniques, open up experience and make the familiar strange for participants. However, there are also wider questions about the purpose of research, and the ethical obligations that researchers hold in relation to disseminating their findings. Research studies often generate the standard outputs of chapters and journal articles but their audience is restricted to academia. Arts-based creative mediums of dissemination are increasingly utilised by researchers, who have an ethical imperative to present their work to audiences beyond the confines of academia, connecting with diverse publics to increase understandings and negotiate social change.

Moreover, the interpretation of qualitative data when analysing it without any software (meaning doing it by hand) could introduce the personal biases of the researcher. Furthermore, even when software is being used (such as NVivo), biases can still be introduced. In these cases too, the interpretation process requires transparency, in-depth presentation of the methodologies and the strategies of analysis chosen, and a clear definition of how the research was done. Additionally, the fact that some data may come from different cultural backgrounds and sometimes are translated can lead to the issue of the reliability on the translation. A potential solution for these issues lies in the collaboration of researchers, as it may bring different perspectives.

The third point about data is handling. It is important to take care of data concerns throughout the whole research process, from data gathering through to publication. Indeed, some ethical issues come up when data is stored and eventually shared (even monetised) with external users whose purposes differ from those for which the data were initially collected. For instance, when talking about storage, who is the owner of the data? Is it the researcher, the participant or the institution? What level of privacy or publicity is allowed? Regarding digital data, it may be mentioned that digital data never die, therefore repositories may have clear rules about how data would be managed in the future. The European Commission strongly suggests that the archiving procedures should be part of the informed consent shared with the participant (European Commission, 2018b).

The issue of data is quite big and increasingly debated in the ethics field. While this section does not claim to be exhaustive, central issues regarding data management in research processes are reviewed. Yet some challenges provide room for discussion in academic debates: how to further involve the participants through different instances of the research process such as conferences, or with diffusion material? How the trust-building and emotions for qualitative research may be handled? Could the participant lead the research in order to give more legitimacy to data handling?

A key solution to address most of the issues regarding data handling is stronger collaborations between research participants and project stakeholders. The PRO-RES project intends to capitalise on experiences in this direction, for instance with cases such as the Social Science One project.[2] This is a collaborative project involving nine funders who partnered with Facebook to gain access to digital data available on its platform. Despite a tension generated around the privacy issue, this is a representative case about how to partner with technology companies for research purposes.

The PRO-RES framework proposes to handle the data issue including aspects beyond the privacy issue: it should be user-centric, provide examples of business models that are ethics-aware and less profit-centred (increasing awareness of people, introducing public incentives, adding sustainability and promoting 'ethics-by-design' as core resource for collaborative research). This set of guidelines should serve to address the issue of research in a technological era and how data is treated along the research process, avoiding researcher personal biases, methodological and theoretical uninformed decisions and business interests around data created or stored.

Informed consent and the protection of privacy

It is generally accepted that the process of informed consent and its effectiveness lies in specific principles such as disclosure, understanding and voluntariness (see: i-CONSENT, 2019). However, several challenges arise from these principles, deriving from different aspects of the concept of informed consent. It is evident that even when it is conceivable to collect it, the informed consent of participants does not always guarantee their safety and security and 'unreasonably places the burden on them to educate themselves in order to understand the risks associated with sharing their personal data' (JCHR, 2019: 12; Krane and McDonald, 2019). In many real-life situations participants cannot be aware of all the implications of their participation.

Another aspect deriving from informed consent procedures is the protection of privacy, which, at a European level, was highlighted in the framework of the GDPR that increases awareness among researchers about the importance of protecting personal data. However, the main ethical problems in privacy research start from the freedom of research and the independence of the researcher, from the relationships between researchers and public or private funding bodies, and from issues related to citizen involvement in research activities. Thus, and in order to respect peoples' rights and privacy, we need to treat participants as individuals, always taking into account their educational and cultural background and their set of attributes (i-CONSENT, 2019). Moreover, a better conceptualisation of the public sphere, in which observable activities of scientific and social import occur, is deemed necessary, as well as to define the boundaries between the public and the private sphere.

Accountability and transparency

The ALLEA (All European Academies) European Code of Conduct for Research Integrity[3] (ECoC), includes accountability in the list of the fundamental principles of research integrity on which good research practices are based, together with reliability, honesty and respect (ALLEA, 2017). Accountability refers to the responsibility – and the subsequent obligation to provide satisfactory reasons for the ways that research is conducted – that researchers and institutions bear 'for the research from idea to publication, for its management and organisation, for training, supervision and mentoring, and for its wider impacts' (ALLEA, 2017: 4). So, accountability does not have to do only with providing output data to be considered in research and explanations for outcomes, but it is equally connected to the whole research process, all stages of it, compiling 'a narrative research story' (Wagner, 2020). This responsibility, borne by researchers and institutions, is towards all direct and indirect recipients of research results, individually and collectively, to society at large and to future generations, since research is fundamental to a democratic society, a necessary part of a functioning and well-informed society, and provides the opportunity for various values to be exercised.

Accountability, properly pursued and applied, presupposes transparency, namely the open availability of all the information related to a research project's processes, data and sources of funding, and provided with clarity in public.[4] We could say that accountability is agent-centred; it is related to a

'Who?' question, whereas transparency reflects not only the agents involved in research, whether conducting it, funding it or enjoying its benefits but, also, the processes involved, the stages of development, the scope, the aims, the goals of research enterprises, and any quality or ethics assessment that has been done prior to the commencement of the research. So, transparency is related to a number of questions about research, like 'who conducted this research?', 'how was it conducted?', 'to what end?', 'who funded it?' and so on. To what extent and in what way can research be regarded as independent? Accountability and transparency contribute to the building of people's trust in science, in researchers and in research results.

The aforementioned definition of transparency is a general one. However, there are, also, exceptions, mitigating circumstances in demanding all relevant information to be always broadly open to everyone, as there are discipline – or research field – specific restrictions. For example, there are research cases in which a balance between openness and its abuse, between transparency and vulnerability to attacks is necessary. The right to know quite often comes into conflict with the right to privacy, when by 'privacy' we refer, not to funding arrangements, but to the participants of a research study and the protection of the person and their right to remain anonymous. Privacy has certain legal perspectives, and it is important to distinguish among what is private, what is semi-private or semi-public, what constitutes the public space, especially now that the digital world and the social media space occupy a significant portion of the social reality. Thus the challenge of anonymity and the question of informed consent are crucially related to the extent that the demand for transparency can be fully satisfied.

Another important dimension of transparency has to do with the need and the willingness of researchers to disclose even negative evidence – evidence that challenges fundamental beliefs or ideologies. For instance, in economics the various 'schools of thought' frame how evidence or research outputs are interpreted. Regardless of which policy outcomes are chosen, it is important to indicate the bases on which those choices are made. Until very recently there was no code of conduct for economists, and ethics issues have been, therefore, frequently questions of trust and ideology, leading to various problems, such as multiple involvements or conflicts of interest of researchers, and biased use of data, which may explain the crisis of trust towards economists. Nevertheless, there are several measures that can mitigate the problem: learning to recognise conflicts of interest through training, favouring an impact-based approach to models rather than a mathematical elegance-based approach, and favouring a global code of conduct for all researchers.

There is, also, the issue of publishing critical results, for example, air quality in buildings, and the related questions about who is responsible for the consequences of publishing such research results, who should publish them, who owns the results, and how consequent reputation issues affect the proper publication of such results. Furthermore, even poor research should be allowed to be made public – the most ethical way of dealing with that is to challenge it publicly. The use of open-source data is quite often susceptible to questions about data's quality, about the different practices of acquiring and presenting them.

An additional, central issue in the 'transparency discussion' has to do with the funding of research and the declaration of interests. As the participants of the PRO-RES workshops concluded, perhaps there is nothing wrong with declared interests – for example, in the US there is open declaration of such interests. The question should be how to make them more 'accountable' by agencies with alternative views challenging them. So there is a need to bring 'truthfulness' to the table and this acceptance is extended, also, to policymaking. Following a 'do not harm' maxim does not seem to be enough. Any policy experimentation has a potential for harm. There are always trade-off considerations – doing harm to one group to benefit another or to achieve other policy objectives. There is also a normative/political element in any policymaking perspective, often involving delicate or controversial issues. For example, on the necessity or not of austerity as implemented after the economic problems of 2008, many people might think that it is wrong; however advocating it is a normative, legitimate judgement. And such judgements can be supported and justified by various biases or ideologies. In the nature of economics, there are always conflicts of interest. Transparency about interests, coupled with normative codes of conduct/commitment and quality marks, may be the best way forward in those areas where it is not appropriate to ban certain positions by means of legislation.

Vulnerability regarding research areas and people involved

Researchers may deal with different types of vulnerability: from the conditions of the subjects of the studies, the potential impact in the researcher's safety, or simply the background or field research. A key concern on this topic is: who may be considered vulnerable? The ethics may go beyond these two fields separately, and consider both the people involved and their background. This is a situation that implies high levels of uncertainty and instability, which may require certain flexibility when discerning an ethical approach.

A lot of research is done in the military field, both in times of peace and conflict, making this a significant area of analysis. It is crucial, therefore, for missions conducted by humanitarian organisations to understand some ethical concerns: how research can be independent and impartial in highly politicised contexts? How can the life integrity of both researchers and local communities be safeguarded when treating sensitive topics that might be at the centre of the dispute? Even going to detailed issues, some questions regarding data use arise: would data protection be in line with saving lives, or may such an alignment have associated life costs? How can one safeguard the confidentiality of the data? How can one avoid abuses in data management?

It is important to understand and evaluate the risks of conducting research in critical scenarios, namely in dangerous, possibly life-threatening or precarious circumstances, in disaster and conflict settings, or working with vulnerable population. Three main problems were reported in the PRO-RES workshops: first, there may be high exposure for both the researcher and the population involved when something opposite to the beliefs of the community is disseminated as a result of a study. Second, it may sometimes be difficult to involve people in research, even though this is highly needed. A potential contribution to the solution of this problem may come from investigative journalism and the fact that many journalists have turned to do more social-type research, to do field research in order to gather good evidence to support their reports, and, in that way, to attract more participants in research and to circulate important information to broader audiences. Third, the connection between ethical research and lawful research may touch the issue of confidentiality: is the researcher supposed to follow the law by, for example, asking certain populations for documentation of their identity? Different solutions have been discussed to face these problems. As a first idea, there may be an *ex ante* evaluation regarding the risks associated with the research in vulnerable contexts or with vulnerable populations and, consequently, a decision whether the research is feasible or not may be made. In positive cases, specific conditions and safeguards might be checked beforehand. Additionally, researchers may avoid being personally involved in context-related issues, when they conduct research in the field and they form a clearer view on what is really happening in a certain environment or context, as the objectivity of their research will be jeopardised.

Following the directives of the European Commission, a question to address is: 'Can the research results be obtained by involving another, less vulnerable, group?' The best way to address the issue is by 'Explain[ing] how will you ensure that the individuals you involve will not be stigmatised, retraumatised or otherwise harmed through their participation in your research' (European Commission, 2018a).

Some challenges in this field are the change of personnel, the loss of research-memory, keeping track of ethics standards, data management and the impact on vulnerable populations who are the object of various types of research on a regular basis, such as refugees, asylum seekers and migrants (Block et al, 2013; Habib, 2019; Czaika and Godin, 2021). Within this issue, the creation of an institutional memory turns out to be a feasible solution that can be developed by two actions: on one side, cooperation between stakeholders and strengthening the network of researchers working on humanitarian or conflict areas; on the other side, the involvement of local communities in the research ethics process, ensuring it is iterative and locally relevant.

Data protection remains an open challenge for the vulnerability issue: how is it possible to guarantee the control and protection of sensitive data in different cultural backgrounds or hostile environments? Indeed, information coming from the vulnerable population, when correctly used, could lead to real changes in policymaking.

Open access and open science

After many years of debate regarding the open access (OA) to scientific publications and its usefulness, the scientific community seems to have reached a critical point. Several recent developments such as Plan S (as of 2021, all scholarly publications of research results must be published in OA journals, on OA platforms, or made immediately available through OA repositories) show that OA could be the reference point in scientific publishing within the next few years, with many publishers who had been reluctant to abandon the subscription business model showing an acceptance of OA (Aspesi and Brand, 2020).

Two primary reasons why open science is important are: integrity, as it opens publication up to peer review; and it also allows reuse of the data, which retrospectively benefits the society. Since billions are spent annually on academic and government research, related to grant administration, talent management, data and image manipulation, and impact measurement, promoting open science will, also, have economic impact on all stakeholders involved. The term 'open science' encompasses information-based solutions, such as research data management, facilitating collaboration, improving research integrity and evolving systems of evaluation (Hersh, 2019).

However, problems remain in terms of what is meant by 'open' and what is being opened. Open science and open data include concern about how to reconcile conventional assurances to qualitative research participants that the data they contribute, in whatever form, will be handled in ways that protect their identity, and the growing demand that all data should be deposited in ways that allow unconstrained public access. This is thought to be an important safeguard against research misconduct, by allowing independent verification of analyses and conclusions.

Another issue is that despite the benefits of OA, it can lead to predatory challenges due to institutional pressure to publish in OA and fees that act as a barrier. This might result in some institutions and young career researchers being left out. What is needed is *open science* that benefits the researcher without intermediaries, and recognises quality over quantity.

Publishing procedures, review systems and dissemination practices

Peer-review processes

The peer-review process, according to the participants in the PRO-RES workshops, should follow a more holistic approach when assessing a research study. Currently, peer review is a process of assessing (1) the soundness of the methodology with which a research question or a hypothesis was examined, (2) the reliability of data presented answering the research question or verifying the hypothesis and (3) the robustness of the analysis of the research data. This approach has been the cornerstone to safeguard the accountability of science, as an internal procedure; meaning a procedure within which the work of researchers is being checked by their peers, other researchers. This, however, leaves many issues, like the ethical issues raised by a scientific study and its societal relevance.

The PRO-RES workshop participants communicated that during peer review incipient ethical issues need to be addressed or, at least, the ethical scrutiny through which the research methodology has been put should be assessed, particularly before the research commenced. We should point out that this opinion was not communicated as a way to replace existing procedures of ethical assessment of proposed funded research, active on a European level, through the different funding mechanisms of the European Commission. The participants expressed their concern that additional *post facto* safeguards should also be put in place.

Seen from this perspective, the authors believe that this additional level of ethical assessment can fit into the existing review procedure via the following ways:

- *for journals* with the conventional review system the procedure of ethical assessment should be clearly presented in detail as part of the manuscript, that is, being a standard annex or as supplemental material, similar to the way underlying data are presented in some journals;
- *for preprints* to be eligible for OA publication, under the condition that they clearly present in detail the ethical assessment of the presented research;
- *for registered reports* to be accepted, similarly to journals and preprints, they should clearly present in detail the ethical assessment of the proposed research.

The importance of publishing negative results

The participants were involved in discussions related to the drive to publish a lot, especially in the biomedical sector, and to publish only success stories: papers containing research that confirms the alleged initial hypotheses. The discussions stressed that a change in the philosophy of scientific or scholarly journals is necessary, so as to be more open to publishing negative results, since such publications would be extremely helpful for other researchers and systematic reviewers. It would be of special value if high-ranked journals become advocates of publishing negative results. However, this change in philosophy should be supported also from funding entities, since there are several cases in which the researcher himself avoids publishing negative results, in order to be in line with what has been promised at the proposal stage.

In a broader sense, failure analysis is critical. The scientific community needs acknowledgement of failure, and understanding of good failure, to help innovate, since failure is not something to criticise but a great opportunity to learn, by avoiding repetition of mistakes. To that point, a Journal of Ethical Dilemmas and Failures could be useful. In line with this opinion, a recent *Nature* publication suggests that after a paper is made public, the authors should collaborate with independent replicators, in order to design a replication experiment that both agree will be meaningful (Nosek and Errington, 2020). The authors name this initiative 'precommitment' and go as far as to suggest that this process could be documented using a Registered Report for the sake of transparency and efficiency.

Science journals and publishing

The challenges imposed by OA and open science are significant. The main concerns raised were related to the opportunities created for predatory publishers, journals that receive fees for OA publications and charge customers a subscription fee for the same publication (the so-called double-dipping), and making full costs of science communication fall on public funds rather than being shared by private or corporate users of that knowledge. In particular, the latter point has been widely recognised and more generally described as a lack of an open data sustainable financial model, as it is developing as a complex sociotechnical system (Kitchin, 2020). Specifically, the main challenge is how open data projects are funded sustainably in the absence of a direct revenue stream that is bound to be covered by direct government subvention.

The ethics of the editors is very central in these discussions. One point is the automatic rejection without peer review, for example, when an author is suspected of plagiarism. This decision is part of the initial screening procedure that several scientific journals apply. However, since plagiarism detection is done automatically via software, this decision is not invariably correct. In addition, there are also ethical and integrity issues for editors when accepting a paper without peer review. This is mostly for editorial/introductory pages, but may, in principle, happen with any kind of paper. There is a need to have some explicit and openly communicated criteria that editors should use in rejecting a paper such as manuscript composition, number of figures, wordcount, aggregation or citations. Emphasis has been put onto the need to train the future journal editors and reviewers, since young researchers do not have the needed experience in evaluating papers.

Taking into account that international hyper-competition in publishing has led to a lower quality level of publications (Ioannidis, 2018), publication ethics and the interests of editors and publishers do not necessarily coincide, according to the participants' views. The important question posed was whether a proper framework or a tool guiding editors in proper ways to make decisions exists. In addition, there exists an environment of uncertainty about which are the most important criteria to determine the value of a contribution. Some editors pay attention to the authors' status, while others emphasise the absolute quality of the paper or the interest of its findings. It is acknowledged, however, that there is an inherent difficulty to strike the right balance between maintaining freedom within a common framework.

The education of young researchers on RE & RI and the creation of a new research culture

There is a strong linkage between the quality of the research produced by the educational institutions and the commitment of their staff to research integrity. Taking into account the given interdisciplinary and the established international collaborations, the educational institutions need to collaborate in a global level in order to develop and maintain a strong research and educational culture that will be based on research integrity, respecting the scientific community, as well as the wider public (Lerouge and Hol, 2020).

That said, it is crucial to educate and train young researchers towards an 'RRI vigilant' mentality of research conduct, within the proper scientific working environment. This is actually a prerequisite that has to be promoted from the stage of the undergraduate studies, and continue throughout all levels of studies and research. More specifically, we need to define how to motivate young researchers to conduct their experiments while taking ethical implications into consideration. More and more European countries are establishing a rule for a minimum number of published papers in order for a researcher to obtain a PhD, which means that the pressure to publish is getting more intense for young researchers. Ethics must not be sacrificed in favour of fast publication (without cross-checking) (Conroy, 2020; Robishaw et al, 2020; Raykov, 2020).

Types of incentivisation supporting this research culture

The participants at the workshops delved into the question of how to strike the right balance between aspirational and regulatory approaches to foster adherence to an ethical research practice. As a necessary addition to increase the impact of ethical research, it is necessary to incentivise evidence-based policymaking by incorporating reliable science into policy. To reach such an end, science actors should be engaged in promoting an effective use of scientific results in support of EU policymaking. The drive to be inclusive in bringing in different voices, besides scientists, when providing knowledge for policymaking, must not result in diluting expertise and being less rigorous in assessing evidence. An inclusive approach should mitigate the fact that even in evidence-based policy, there is not one single, one-directional option based on the evidence provided.

While science, like all institutions, creates a particular distribution of rewards and incentives, some of these have become perverse, reflecting a publish-or-perish culture, created by the metrics used by funders and universities. If all we look at is, for instance, a researcher's h-index, this will perpetuate the problem. National funders are discussing ways to revise the reward system in science to address the underlying issues. This process is under way not only in Europe, but also in countries where incentivisation is being revised (Mallapaty, 2020).

The issue of incentives was also dealt with in the initiative of the Mutual Learning Exercise for Research Integrity that was organised by the European Commission (Hermerén et al, 2019). In this context, the discussion focused on issues like creating a common understanding of how the notion of incentives is being perceived in different European countries, and the necessity to align incentives in EU Member States. These incentives may focus on the individual level (symbolic or formal credit systems that openly recognise the contribution of a researcher to research integrity) or at an organisational level (international RPO ranking systems to include institutional measures to foster and safeguard research integrity).

The participants went beyond the context of RPOs, by discussing incentivisation to acknowledge excellence in research integrity in scientific journals. For reviewers there is nothing official, apart from disseminating a list of reviewers in a periodic fashion or by acknowledging the peer-review excellence of an individual researcher. When unethical behaviour occurs, publishers unofficially ban reviewers. A better approach would be to reward good behaviour, rather than punish bad. However, also in the case of scientific journals, there is a need for specific standards, in order to establish what counts as good ethical behaviour.

Challenges of research in international contexts

Research has been increasingly adopting an international perspective, with increased collaboration between institutions, countries and different national and regional stakeholders. However, ethical issues can be envisioned in certain circumstances, such as humanitarian missions, funding of research or the existence of diverse ethical bodies at national levels.

As discussed in the section on vulnerability, earlier in this chapter, humanitarian missions are a main area of reference to understand problems and challenges in research ethics in international contexts. International organisations such as the International Committee of the Red Cross or Médecins Sans Frontières have a long tradition of working with vulnerable populations, providing huge advantages for conducting research. However, there is a gap in understanding the differences between research ethics and humanitarian ethics, even though humanitarian ethics are based on international law with an ethical background (for example, principles of humanitarian intervention).

One recommendation resulting from the PRO-RES workshops is related to the context variable: the cultural and social circumstances within which certain research projects are carried out. International conflict areas are quite unstable and uncertain, with local people involved in research frequently saturated by researchers, and researchers' integrity that might be affected by the circumstances they face in the field. Therefore, the involvement of local communities along with the creation of a network of researchers in crisis scenarios may be useful to improve the way research is done.

A second issue that appears in international research contexts is related to funding and mobility. Some researchers who are unsuccessful in getting funding due to previous violations of codes of conduct or guidelines may try again in a different country using their funding schemes. This is evidence of a 'play with the system' attitude of some researchers and it needs international cooperation to address. Some possible solutions on these issues could be, for instance, the exchange of information among stakeholders, for instance funders in different countries, or national organisations supervising the system and applying background checks while taking into consideration any privacy requirements.

The last issue regarding international research context is the diversity of research ethics bodies across countries. While it is important to have a local approach to an international challenge, and to understand the diversity of contexts where research is carried on, the development of an internationally recognised framework remains a challenge. PRO-RES objectives are aligned with contributing to the harmonisation of the existing guidelines and the anticipation of the complexities of the international and dynamic research practices.

Challenges emerging with innovation and the latest technological developments

Ethics principles and the 'sustainability' of doing no harm

One of the points most intensively discussed was the fact that the application of standard ethics review procedures contains sections not originally designed for certain disciplinary fields. This can produce paradoxes that result in jeopardising the pursuit of the ethical aims of the review itself. Although 'first do no harm' is a prime research ethics principle, it is hard to achieve. Doing harm is inevitable in some disciplines, since economics in particular requires making choices and therefore brings up opportunity costs, which means someone or some group will necessarily be affected and therefore possibly adversely affected.

Substantive discussion was devoted to how the notion of responsibility could broaden the agenda of ethics assessment from its somewhat narrow focus on what was called the 'first do no harm' perspective derived from the biomedical context. Fast evolving technologies, like artificial intelligence (AI), or fast emerging technologies, like organoid-related research, put a strain on ethics practices. Such concerns appeared in the discussions, calling for these new ethical problems to be answered with an updated ethical arsenal; not just by echoing outdated ethics assessment practices.

However, new challenges are not necessarily produced by novel technologies. For example, the problem of climate change was a recurrent theme in the workshop discussions. Climate change means that economics and finance have to be rewritten, following an evidence-based policy supported primarily by physics, as well as social sciences, political science and other scientific fields. Physics provides the necessary evidence proving, first of all, that climate change is real, that it is already happening, and giving the time frame of the development of the phenomenon. Climate change, however, as a dynamic and complex phenomenon, has various impacts on the natural environment and society requiring scientific consensus and interdisciplinary approach. Globally, we are unprepared for a climate shock. At the time of this writing, the prediction is for zero economic growth in ten years and widening disparity in wealth and opportunities. So there is a need to mandate carbon offsetting in health services, in education, in transport. This is another example of why ethics procedures must be in a constant readiness to cope with such fundamental changes and stress the importance of impact assessment.

Such concerns are further supported by the fact that policy experiments rarely produce definitive 'truths'. Even the outcomes of policy experiments might be interpreted differently. Ethical risk and impact assessment find it difficult to foresee the consequences of new technologies. Additional concerns were raised with regard to the blurring of boundaries between research for academic purposes and for higher impact on society (like from non-governmental organisations), which poses new issues of regulation. For new, original, directions of research in particular, it could be difficult to specify benefits. A relevant debate is how to bring about considerations of a broader interpretation of responsible research, confronting issues of societal and social relevance. One suggestion is to adopt an exploratory approach to ethics that would proactively look for new ethically relevant questions pertaining to science.

Technology assessment

Another important topic is the issues raised when funding applications of debatable uses, for example, sex robots, atomic bombs, designer babies or spy-drone research. It might not make sense to have discussions about whether it is ethical to fund this research or not, based on the possible non-ethical implications of the developed technology. The workshops' participants argued that if profits can be made with a technology, it will be developed, if not by publicly funded research, then by industry. The limits of what we can do have been set, already, in many fields, much further beyond what is desirable. Development choices for new technologies are ethical in nature, but they are made more in innovative industry than in universities.

The real debate may be a parliamentary one, about whether or not to forbid the use of some new technology. This is an ethical debate and has a lot to do with technology and society, but in its essence, not to do with research. We cannot know a priori the applications of a new technology. For example, 3D printing was developed as a laboratory demonstration in 1984, whereas the internet storm on the '3D guns' arose sometime in 2012. This means that it took a while for ethical concerns to arise. To use the apocryphal response of question put to the founder of laser technology, 'What good is your invention?', the answer was 'What good is a baby?' For new things, at the research level where ethics review applies, the answer is not obvious due to delays between lab and market.

Another upcoming technological breakthrough that would pose ethical issues is machine learning. The problem is that the people using it are not the developers, and they just try to combine different tools developed by other people, for their own purposes. Many parts of these issues are like black boxes since not even the actual developers understand how they work. This lack of understanding about how it works propagates to the end user ending up with some AI that cannot actually be supervised by people. So this is an ethical issue about whether to trust AI tools or not. The new technical solutions can make obsolete some procedural safeguards that had previously been perfectly adequate.

Conclusion

This chapter's aim has been to present in a concise way the main issues that fall under the broad umbrella of RE&RI, which reflect the general problems and considerations that are being faced by individual researchers, students, RPOs and RFOs, and REC and IRB members when performing their roles in research projects. These issues' origins and history of development, their transformation over the years, their correlation and possible solutions constitute continuous debates and a dynamic environment of interaction among various stakeholders. In fact, the analysis of this broad umbrella of RE&RI and the examination of the developments within it could be considered as a meta-research field and render RE&RI a stand-alone field of study and expertise.

Throughout the PRO-RES project's development, all partners involved reconfirmed that a special place is reserved for policymakers and their advisors among the stakeholders involved in this dynamic environment. Dealing with these issues without taking them into consideration leaves the general picture incomplete. Good independent research provides useful information and proper evidence for the formulation of policies of general interest. At the same time, the main goal is to restore and enhance society's trust in science, and this cannot succeed without the participation of policymakers and their advisors in using scientific results to make and justify their decisions. Ethical commitment does not end with the various ethics reviews at the formulation and planning stage of research, but it follows research projects, results and products up to the final stages of dissemination, exploitation, implementation, endorsement and compliance. It is, therefore, necessary to incentivise evidence-based policymaking by incorporating reliable science into policy.

The PRO-RES project's guidance framework encourages policymakers and their advisors to seek evidence for their decisions from research that has been conducted ethically, responsibly and with integrity, providing them with useful tools for this (the Accord, the Toolbox and Resources). While recognising the actual difficulty in involving policymakers in the process of developing an RE&RI framework useful to them and to their purposes and concerns, and acknowledging the fact that policy advice and policymaking are complex procedures involving different, often contradicting aims, interests and challenges, the ambition of awareness raising towards RE&RI and the idea of transparency as a guiding principle remain our key notions.

Notes

[1] See: https://prores-project.eu/

[2] See: https://socialscience.one/

[3] The ECoC is recognised by the European Commission as the reference document for research integrity for all EU-funded research projects and as a model for organisations and researchers across Europe.

[4] See the Glossary of Terms and Concepts provided in the PRO-RES project's website http://prores-project.eu/glossary-of-terms-and-concepts/ [accessed 30 January 2022].

References

ALLEA (All European Academies) (2017) *The European Code of Conduct for Research Integrity*, available from: https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf [accessed 30 January 2022].

Aspesi, C. and Brand, A. (2020) 'In pursuit of open science, open access is not enough', *Science*, 368(6491): 574–7.

Block, K., Warr, D., Gibbs, L. and Riggs, E. (2013) Addressing ethical and methodological challenges in research with refugee-background young people: reflections from the field, *Journal of Refugee Studies*, 26(1): 69–87.

Carey, M. (2019) 'The tyranny of ethics? Political challenges and tensions when applying ethical governance to qualitative social work research', *Ethics and Social Welfare*, 13(2): 150–62.

Conroy, G. (2020) 'Is research integrity training a waste of time?', *Nature Index*, 12 February, available from: https://www.natureindex.com/news-blog/is-research-integrity-training-a-waste-of-time [accessed 30 January 2022].

Czaika, M. and Godin, M. (2021) Disentangling the migration-development nexus using QCA*, Migration and Development*, ahead of print, available from: https://doi.org/10.1080/21632324.2020.1866878 [accessed 30 January 2022].

Deketelaere, K., Maes, K. and Valkenaers, B. (2019) 'Universities and the Future of Europe', LERU briefing paper for the EU institutions, September, available from: https://www.leru.org/files/LERU-Briefing-Paper-Universities-and-the-future-of-Europe.pdf [accessed 7 July 2021].

European Commission (2018a) Ethics in Social Sciences and Humanities, October, available from: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020_ethics-soc-science-humanities_en.pdf [accessed 31 May 2021].

European Commission (2018b) Ethics and Data Protection, 14 November, available from: https://ec.europa.eu/info/sites/info/files/5._h2020_ethics_and_data_protection_0.pdf [accessed 31 May 2021].

Habib, R.R. (2019) Ethical, methodological, and contextual challenges in research in conflict settings: the case of Syrian refugee children in Lebanon, *Conflict and Health*, 13: art 29, available from: https://doi.org/10.1186/s13031-019-0215-z [accessed 30 January 2022].

Hedgecoe, A. (2016) 'Reputational risk, academic freedom and research ethics review', *Sociology*, 50(3): 486–501.

Hermerén, G., Zwart, H., Marušić, A. and Fanelli, D. (eds) (2019) Mutual Learning Exercise (MLE) on Research Integrity, Luxembourg: Publications Office of the European Union.

Hersh, G. (2019) ‘Collaboration is the key to open access and open science’, *Times Higher Education*, 21 August, available from: https://www.timeshighereducation.com/opinion/collaboration-key-open-access-and-open-science [accessed 25 May 2021]

i-CONSENT (2019) Gaps and Challenges to the Informed Consent Process by i-CONSENT, 7 June, available from: https://i-consentproject.eu/gaps-and-challenges-to-the-informed-consent-process-by-i-consent/ [accessed 30 January 2022].

Ioannidis, J.P.A., Klavans, R. and Boyack, K.W. (2018) ‘Thousands of scientists publish a paper every five days’, *Nature*, 561(7722): 167–9.

JCHR (UK House of Commons and House of Lords Joint Committee on Human Rights) (2019) *The Right to Privacy (Article 8) and the Digital Revolution*, November, available from: https://publications.parliament.uk/pa/jt201919/jtselect/jtrights/122/122.pdf [accessed 30 May 2021].

Kitchin, R. (2020) ‘Four critiques of open data initiatives’, Impact of Social Science blog, LSE, available from: https://blogs.lse.ac.uk/impactofsocialsciences/2013/11/27/four-critiques-of-open-data-initiatives/ [accessed 24 July 2020].

Krane, O. and McDonald, J. (2019) ‘Data protection and privacy: is the consent model broken?’, Experts Insights blog, Charles Russell Speechlys, 4 December, available from: https://www.charlesrussellspeechlys.com/en/news-and-insights/insights/commercial/2019/data-protection-and-privacy-is-the-consent-model-broken/ [accessed 30 January 2022].

Lerouge, I. and Hol, T. (2020) ‘Towards a research integrity culture at universities: from recommendations to implementation’, League of European Research Universities (LERU) Publications, January, available from: https://www.leru.org/publications/towards-a-research-integrity-culture-at-universities-from-recommendations-to-implementation [accessed 30 January 2022].

Mallapaty, S. (2020) ‘China bans cash rewards for publishing’ *Nature*, 579(7797): 18.

Nosek, B.A. and Errington, T.M. (2020) ‘The best time to argue about what a replication means? Before you do it’, *Nature*, 583(7817): 518–20.

Raykov, M. (2020) ‘Education researchers’ perceptions of and experiences with the research ethics application process in Europe and beyond’, *European Educational Research Journal*, 19(1): 10–29.

Robishaw, J.D., DeMets, D.L., Wood, S.K., Boiselle, P.M. and Hennekens, C.H. (2020) ‘Establishing and maintaining research integrity at academic institutions: challenges and opportunities’, in *The American Journal of Medicine*, 133(3): E87–E90.

Wagner, B. (2020) ‘Accountability by design in technology research’, *Computer Law & Security Review*, 37: art 105398, available from: https://doi.org/10.1016/j.clsr.2020.105398 [accessed 30 January 2022].

6

An ethical framework for think tanks: easier drafted than done?

Fabian Zuleeg

Introduction

There are no clear definitions of what constitutes a think tank; any organisation can choose to use this label. In its broadest sense, it is an organisation that aims to influence policy through research, analysis and policy advice. The non-academic research sector contains a wide range of different actors,[1] not only think tanks but also foundations (including party-political ones), independent research institutes, trade associations, research bodies linked to unions or employer organisations, consultancies, research departments of private companies or non-governmental organisations (NGOs) and so on. Not only is there a plethora of different organisations, some actors in the field use this lack of defined structures to influence policy without disclosing potential conflicts of interest, for example, acting through or even setting up a 'think tank' that only serves a particular interest, being neither transparent nor independent.

Think tanks[2] are generally rather small, both in terms of overall budget and in terms of the staff they employ. Many think tanks are self-financing NGOs; their funding model usually relies on diverse sources of generally short-term funding. However, some think tanks receive significant public sector/governmental core funding, at times explicitly being affiliated with, or part of, government or other structures, such as political parties.

Arguably, research does not lie at the core of think tank activity, but rather the focus is on analysis and policy advice, based on evidence, which can include research studies but also relies on a wide range of other sources. Much of the activity of think tanks takes place behind closed doors, with results often being unpublished. Monitoring and enforcing ethical conduct of such activities is far more challenging than setting up a research project in the public eye with clear ethical guidelines, an ethical review process and (public) project funding that is conditional on ethical conduct. Having this different focus implies that there is a need to create the right incentives for think tanks to adhere to ethical guidelines, in line with the marketplace in which they operate.

Why ethical conduct in think tanks matters

Whether think tanks conduct their activities in line with ethical principles matters to society. They can fulfil an important function in democracy, challenging policymakers to implement innovative solutions, based on evidence. The organisations in this sector can connect a range of different stakeholders and provide a platform for evidence-based debate. Non-academic research, analysis and policy advice can be of critical importance for decision-makers, being able to respond quickly and innovatively, specifically focused on policy and more easily integrated in the policy process. But if acting unethically, non-academic research organisations can exercise a negative influence on democracy, manipulating policymaking and public opinion.

But ethical conduct also matters to the think tanks themselves and the individuals working for them. In addition to a personal commitment to such principles, for a think tank's independence and derived credibility, it is also crucial both in terms of effectiveness and of legitimacy (Lux, 2021). Given that the explicit aim of think tanks is to influence (democratic) policy decisions, legitimacy derived from underlying commitment to ethical principles provides the basis for intervening in the decision-making process. In essence, think tanks that do not adhere to ethical principles lack the democratic legitimacy for their activities.

Barriers arising from the nature of the market

One reason that the distinction between academic and non-academic research matters is because organisations operating in the different segments of the research field face a very different marketplace and, hence, very different incentives. Table 6.1 is an attempt to chart the difference between evidence-based analysis and advice organisations (with a particular focus on think tanks) as compared to more traditional academic research organisations.

Of course, the distinction is nowhere as clear-cut in reality and there might well be hybrid forms, such as think tanks housed in universities. Presenting this as a strict dichotomy would be misleading. Rather it should be understood as a spectrum, with non-academic research organisations tending towards one end and academic research to the other. In both segments, there will be exceptions to the general structures sketched out in Table 6.1. It is important to emphasise that this is not a categorisation that distinguishes between ethical or non-ethical behaviour but rather an attempt to characterise the structural differences between the different actors. In both parts of the research spectrum, there is a need to follow legal requirements, for example the General Data Protection Regulation (GDPR).[3] There is

Table 6.1: Differences between think tanks and traditional research organisations

Traditional research organiasations/higher education research	**Think tanks and specific research institutes**
Commitment to scientific 'truth'	Commitment to a mission, goal/objective, impact and/or potentially a political/ ideological direction
Primary research, mostly based on historical data, generation of knowledge	Interpretation & analysis of existing, broad range of evidence, forward looking, use/ application of research results
Objective, current state of knowledge	Normative, recommendations
Provision of public goods	Working for/with stakeholders to meet specific interests and needs, public interest rather than public good
Peer review, critical assessment by other academics/academic institutions	Assessment by 'clients' & judged relevant through uptake by decision-makers. Rarely: think tank peer review or criticism
Base funding through dedicated research funding but with conditions determined by the communal funding bodies (governments, EU), long term, strong incentive to keep to research guidelines	Diverse funding with distinct interests, short term, weak incentive to keep to research guidelines, at times precarious balance between strategic content and interest of stakeholders providing short term funds
Defined organisational form, not for profit	Diverse organisational forms, variety of governance models, some not-for-profit, others not
Clearly definable conflicts of interest	Competing interests, no clear cut hierarchy of right vs wrong
Defined sector with governance structures, official status	Undefined sector, no common organisation, blurry delineation, self-identification (sometimes with intention to deceive)
Clear roles/relationship to institution of individuals (employment). Clear cut institutional association of individuals	Individuals can be objects/subjects, researchers, funders, founders, decision-makers, governance at the same time. Multiple forms of work relationships/ arrangements
Providing research environment/ collaboration	Providing a platform for debate
Research design to encourage the production of objective findings, transparency and replicability. Public, open access	Research & analysis that simultaneously aims to influence (e.g. interviews with decision–makers, taskforces, workshops), often behind closed doors and 'one-off'. Informal, below the radar, private information
Dispassionate provision of research results	Proactive involvement in the political debate

Table 6.1: Differences between think tanks and traditional research organisations (continued)

Traditional research organiasations/higher education research	Think tanks and specific research institutes
Restricted number of core funders. Some endowments. Increasing pressure to create spin-offs and return on investment	Wide variety of funders. Year-to-year survival, typically no/few endowments
Uniform type of research projects, guided by a common framework (determined by funders)	Different type of projects with different rules/guidelines
Longer term, structured, predictable	Fast, uncertain, risky, unproven/unprovable

also a need to maintain research integrity, combating plagiarism, falsification and fabrication. But arguably in the non-academic sector, misbehaviour is more difficult to detect and address as more of the activity takes place below the radar.

There will be 'bad actors' in either part of the spectrum, suffering (undisclosed) conflicts of interest, making biased use of data, manipulating evidence and deliberately misleading decision-makers and the public, as well as funders. However, arguably the non-academic sector is less 'regulated' (that is, less bound to a particular set of ethical rules and less dependent on funding attached to those rules) and it can be more difficult to evidence bias when compared to activities carried out by scientific method. In addition, organisations in the non-academic sector often lack the organisational infrastructure to effectively enforce ethical guidelines, such as ethical review processes, in part due to a smaller scale and due to the costs involved, without in most cases offering clear benefits, such as additional funding.

The additional challenge of COVID-19

Given the COVID-19 pandemic and its political, economic and social aftermath, the need and demand for think-tank advice and analysis has further increased, not least since the situation requires rapid responses to complex, new and interconnected policy challenges, in a world characterised by endemic uncertainty.

However, at the same time the crisis is likely to have a negative impact on funding for think tanks, given that providing support to such organisations in times of crisis is often seen as a relatively easy target for reducing expenditure by private firms, but also by cash-strapped public authorities. This increases the financial vulnerability of think tanks, making it more difficult to resist the influence of vested interests if these underpin the financial model of the organisation.

Implications of structural differences

These structural differences imply that academic and non-academic research organisations tend to operate differently, driven by a different set of incentives. While for academic organisations following the principle enshrined in an ethical research framework is core to their operations, this is not the case for non-academic organisations. Indeed, non-academic organisations often face funders that are not focused on ethical principles and might even not want the organisation to follow strict ethical guidelines but rather, for instance, achieve maximum effectiveness in policy influence. It follows that any ethical framework needs to be adapted to the needs of non-academic organisations and that its implementation needs to be incentivised for this sector.

This is reinforced by the competition within the non-academic research sector. At best, following an ethical research framework provides marginal benefits, for example potentially better access to public research funding, but at worst it can be a competitive disadvantage; in essence, if there is no significant sanctioning of misbehaviour, acting in a non-ethical way can be a competitive advantage, by, for example, being able to access funding for advocacy and lobbying or being more effective in changing policy by using covert and underhand methods.

Simply attempting to impose a framework will lead to displacement: for example, it might simply lead to one part of the sector no longer operating under such a framework, promising their clients more, precisely because they are not bound by such rules. It could thus become a competitive advantage not to follow the rules. There is a need to reverse incentives and to create an ethical framework that benefits those who are willing to abide by it.

Cornerstones of an ethical framework for non-academic research organisations

The structural differences between academic and non-academic research providers raise the question of what guidelines are appropriate. In other words, which operational principles should guide the work of a think tank to underpin its independence, credibility and legitimacy?

- Commitment to existing guidelines, for example, those on research conduct or on opinion polling (to all elements or only some)?
- Commitment to not deliberately mislead or misrepresent the evidence: 'the end does not justify the means'?[4]
- Clarity of purpose and interests?[5]
- Non-dependence on individual sources of funding/diverse funding sources, clear rules tied to funding arrangements?
- Governance and transparency to safeguard independence?

For non-academic research organisations, and in particular think tanks, independence lies at the heart of their credibility: they must ensure and demonstrate their independence, including a clear demonstration that the funding they receive might influence their agenda (such as which topics are being examined) but not the results of their analysis (the conclusions and recommendations). To ensure this, think tanks need to have adequate governance structures and be transparent about their operations.

A further crucial element is a multi-stakeholder approach. Think tanks should act as a bridge between a wide range of actors in the policy process, including those that might struggle to engage effectively without the facilitation of think tanks. Debates need to take into account the different viewpoints and ensure that policy recommendations are developed by taking into consideration different and, at times, also opposing views.

At the same time, think tanks need to be mindful of equality and diversity concerns, ensuring that there is not a bias in their operations driven by a lack of participation of certain groups, as part of the governance structures, management, senior staff or speakers at events.

Having different views and groups represented in the political/policy process is essential. It is part of democratic debate to have different views and conflicting solutions; indeed, having a process of thesis being confronted by antithesis, leading to synthesis, is core to the development of better policies. Different opinions, and advocacy of them, are a sign of a healthy democratic debate and a wide range of political opinion has to be legitimate in the absence of universal truth. Any ethical framework has to ensure that it addresses the abuse of freedom of speech but does not constrain the right itself or the broader, opinion-driven political debate. This requires an approach that is not rigid and legalistic but incentivises ethical principles, through better governance and greater transparency (see Box 6.1).

Box 6.1: Transparency and governance

Transparency is the minimum degree of disclosure of activities, governance arrangements, financing and people and organisations involved in the work of the think tank, which are open to all for verification. Transparency should follow the following characteristics:

- Comprehensive (historic data, timely (that is, not years out of date) activities, financial and governance transparency, details on all financial support (including source) above a relatively modest threshold, clearly showing all people directly involved in the think tank (staff, experts, governance)).
- Accessible (easy to find and access, including obvious links from the homepage, information available in both local and international languages, contacts for further information).

- Transparency by default (proactivity in making information available, clear reasoning for exceptions – for example, commercial confidentiality, proprietorial information, all information required to assess good governance, as set out earlier, is freely available).

Governance is how the think tank is managed, how decisions are taken both day-to-day and long term, the organisational form including statutes and by-laws and how the monitoring/oversight function is performed. Good governance should include the following elements:

- Financial governance (avoiding single/limited number funder dependency, plurality of types of funding, external auditing, financial management principles; for example, multiple people involved in payment).
- Structural governance (appropriate legal form, no direct power relation to government, parties or other stakeholders, clear roles and functions of governance bodies, including who appoints whom, implicit 'owners' (founders and so on)).
- Management (role of executive director or equivalent and how are they appointed/overseen, succession planning for key personnel, who can commit think tank resources, separate roles of income earners and pro bono overseers; for example in a non-remunerated governing board, who speaks for the think tank and/or can come up with think tank positions).
- Policies and principles (explicit mission statement, clear commitment to transparency, good governance, independence, quality management procedures, not-for-profit and so on, policies on data protection, copyright, environmental, gender/equality, commitment to evidence-based working, involvement of multiple stakeholders, external reporting).

Greater transparency tends to enforce better governance by ensuring that there is public scrutiny. Good governance is not only doing the right thing but being seen to do the right thing. But better governance and greater transparency are not cost-free. At the very least, they involve managerial and administrative efforts and can slow down decision-making; they can also impact negatively on financial sustainability (for example, by ruling out certain sources of funding), continuation (for example, by reducing the incentives for key people to remain involved) and impact (for example, by limiting the ability to work behind the scenes).

Yet, multi-stakeholder working, transparency and good governance, and ultimately independence, are not optional extras that can be discarded when inconvenient. A think tank that cannot deliver a high standard of governance and transparency might have to consider whether the think tank can fulfil its functions if not adhering to these principles. Think tanks should strive towards these high standards of governance and transparency and should be asked to demonstrate their adherence when interacting with funders. In a world where they take a stronger political role, they must be 'holier-than-thou', providing as few points of attack as possible to their new political opponents, who can otherwise challenge their legitimacy.

An ethical framework for academic and non-academic research

The European Commission–financed PRO-RES project (aiming to PROmote ethics and integrity in non-medical RESearch) has been working on the creation of an ethical research framework that is applicable to both academic and non-academic research.[6] At the heart of the project is the STEP Accord, which is envisaged as a statement of principle that academic and non-academic organisations can endorse and sign up to (see Box 6.2). The Accord is underpinned by principles and rationale that set out in greater detail what this Accord means when applied to research activities (see Box 6.3).

Box 6.2: The STEP (Scientific, Trustworthy and Ethical evidence for Policy) Accord

As signatories to the Accord:

- We recognise that an underpinning by high quality research, analysis and evidence, including policy appraisals and evaluations, is a pre-condition for evidence-based policy-/decision-making, and hence rational policy actions and effective outcomes.
- As individuals and institutions involved in commissioning, funding, sponsoring or conducting research, collecting or using evidence for policymaking, we aim to be as transparent as possible on how the high quality of that evidence is assured and will flag up any potential conflicts of interest.
- We agree that to a reasonable degree the independence and integrity of individuals responsible for the conducting and/or gathering of research evidence and its use in policymaking must be respected and supported in ways that ensure the evidence they produce is neither biased nor misleading.
- We will communicate, employ and/or apply only high quality evidence, research or enquiry, in other words evidence that has been undertaken, gathered, collated and analysed using sound, robust and ethical methods appropriate to the task.
- We will ensure that the commissioning, funding, management, conduct, dissemination and governance of research meet high standards of ethics and integrity.[7]

Box 6.3: The principles and rationale behind the Accord

The following points explain the rationale behind the Accord and supply links to supportive resources that will help in seeking to promote ethics and integrity in the evidence produced in all non-medical research:[8]

- Under a commitment to evidence-based policy, all evidence should be based as far as possible on ethically sound research and analysis.
- There are many forms of research and evidence. They include not just formal research projects and programmes, but a range of actions relating to investigation, collation, discovery, exploration, practice, and disciplinary development. Every kind of research and analysis needs to be done ethically.
- Research should be beneficent (or at least non-maleficent) in its aims, its substantive focus, in the process of research, and its application.
- Ethical issues can arise at every stage of research: conception, development, proposal, process, conclusion and dissemination. It follows that ethical consideration cannot be a single-stage process; it has to be continuous.
- Researchers and analysts have to be aware of, and sensitive to, the ethical dimensions of their work. That awareness depends on engagement in ethical discourse as an integral aspect of engagement in research and analysis. Ethical conduct cannot adequately be guaranteed by a fixed number of pre-set rules.
- All researchers and analysts should aim to develop a culture of ethical research, based on continuous discursive engagement. To achieve this, there has to be engagement of everyone responsible for the process, including researchers, stakeholders, peers and the users of research.
- Research and policy advice should not be based on pre-formed prejudicial ideologies or biased political or financial interests.
- Conflicts of interest should ideally be avoided in the production of research evidence and in the provision of policy advice. If this is not possible, all conflicts of interest should be openly disclosed.
- Whenever possible, all sources of information used to formulate evidence should be acknowledged, with exceptions being well-justified and, if feasible, noted (for instance in the case of confidential information or views).
- In order to produce high quality evidence, research must be methodologically robust.
- Only research that has also been conducted ethically and with integrity can be considered 'high quality'.
- All research should be funded, managed, conducted and disseminated ethically and with integrity.
- The processes and institutions involved in the selection of evidence, including research, to inform policy should be independent, open and transparent.
- The effectiveness and impact of all policies should be honestly and transparently assessed or evaluated using high quality research methods.[9]

The crucial challenge is how to implement, enforce and/or encourage these principles. Given the structural issues set out earlier, an ethical framework cannot simply be imposed from above. It has to be developed by the organisations themselves and rely on individual integrity, and the adherence to such a framework has to be incentivised rather than mandated. In addition, the implementation mechanism has to be designed in such a way that it rewards those organisations that adhere, while creating disadvantages for the 'bad actors'.

A 'European Alliance of Independent Think Tanks'

One route to implementation is the right kind of funding to improve think tanks' sustainability – for individual think tanks and for capacity-building in the sector as a whole. This funding should be conditional on the think tanks' commitment to the principles set out in Boxes 6.2 and 6.3, as well as a shared understanding of the common mission and purpose or public interest – the defence of an open, democratic, progressive and pluralistic society. Such funding needs to be structurally supportive rather than project-based to enable think tanks to improve their organisational capacity, alongside a recognition that strict adherence to these principles has financial sustainability consequences.

A fund of funds for institutional support of think tanks could be an important tool to improve sustainability of think tanks, where funders can, on a regional basis, channel funding. If set up with clear criteria for funding and high standards of transparency and governance, it could act as a 'cushion' against those that claim that this funding represents a hidden interest. However, a significant part of think tank funding should remain competitive to ensure that the marketplace can incentivise the pursuit of effectiveness, efficiency and competition. The current dependence on corporate or government funding should not be replaced by a dependence on a single source of funding, even if this is driven by philanthropic motives. All funding carries implications and diversity of funding is a crucial safeguard.

Where project funding is provided, a recognition is needed that for independent think tanks any project funding should include a high proportion of funding for the operation of the think tank, more akin to commercial funding models than academic funding models. Any difficult reporting, monitoring and evaluation requirements should be fully funded from the outset.

The establishment of a European Think Tank Centre or European Alliance of Independent Think Tanks could be an important step to build capacity for the sector. Such an organisation needs to be fully funded, with a transparent structure, including clear criteria for membership based on the principles of the STEP Accord. Such an organisation could lobby on behalf of the think-tank sector, for example with the EU institutions, not only for funding but also to protect think tanks where their operation is threatened by commercial or governmental interference. It could provide information on the European think-tank sector and its situation in different countries, as well as potentially developing a methodology and rankings of think-tank impact.

A European Alliance of Independent Think Tanks could provide a forum for exchange of experience and to develop joint activities and common networks. It could help to provide a focus for support for the transnationalisation of think tanks, in particular the creation of a network with an independent identity. In situations where a think tank can no longer operate in its home country, parts of its operations could be hosted by other network partners, providing an institutional framework for continued operations.

Such an alliance could also develop projects of scale that require cross-border think-tank cooperation, for example, on the future of European democracy. This could include a project on using information and communication technology (ICT; liquid democracy) to trigger informed debate by citizens on EU policies with direct policy impact (through the integration of such a system into the activities of a think tank). Such projects of scale could provide important input to develop cross-cutting and pan-European policies, such as the Conference on the Future of Europe.

Conclusion

Policymakers use research from academic and non-academic sources, including think tanks. To promote the use of ethical evidence, and incentivise the implementation of a common ethics framework for think tanks, this chapter has argued for the establishment of a European Alliance of Independent Think Tanks. This alliance would also be a necessary first step to further develop the ethical framework for this sector, building on the PRO-RES work, for example drawing out guidelines for projects that are not primarily based around research. In particular, it could:

- provide the forum and be the ‘pen’ for the further elaboration of the ethical framework, enabling a process by which the framework is co-created within the sector and thus has a much greater buy-in;
- develop guidelines for think tanks, on how to implement a variety of cross-cutting concerns, including for example on data protection;
- after developing an ethical framework, based on good governance and transparency, monitor how such a framework is being used, for example in funding decisions or whether adherence provides (privileged) access to decision-makers;

- develop an independent global hallmark/quality label for think tanks that adhere to certain principles, especially in terms of independence, governance, transparency and multi-stakeholder engagement; and
- develop a code of conduct for individuals that is recognised by the organisations in the sector and could form part of the employment contract of individuals.

At the same time, such an alliance could channel support to build capacity for those think tanks willing to abide by a common framework. This could include:

- providing organisational support (base funding);
- running training programmes for all think tank staff, including on transparency and good governance; and
- professionalising the management of think tanks through trainings focused on GDPR, ICT, communications, financial management systems, quality management and so on.

Further developing and implementing an ethical framework within the sector and underpinning this with support would maximise usage and impact of the framework. Capacity-building would strengthen organisations and the researchers working for them. Implementation would be incentivised through conditional funding, public scrutiny and by turning 'good behaviour' into a marketing tool that would turn ethical behaviour into a competitive advantage for non-academic research institutions. In the end, developing an ethical framework has to be embedded into the development of the sector as a whole. If it inherently makes sense to the organisations and researchers in the sector, both in terms of content and in terms of the incentives to implement it, there will be buy-in by individuals and organisations. If not, little will be achieved.

Acknowledgements

This chapter is based on a previous discussion paper: Fabian Zuleeg, 'An Ethical Framework for Think Tanks: Easier Drafted Than Done?', discussion paper, Europe's Political Economy Programme, PRO-RES (Promoting Ethics and Integrity in Non-Medical Research) Project, 2 September 2020, Brussels: European Policy Centre: https://www.epc.eu/content/PDF/2020/An_ethical_framework_for_think_tanks_v3.pdf

Notes

[1] This also includes market research/opinion polling organisations. However, for the purpose of this chapter, these have been excluded from the discussion due to the well-defined ethical framework that exists for carrying out such research.

[2] This chapter focuses on European think tanks, which differ from think tanks in other parts of the world, for example in terms of size or endowments.

[3] There is an argument that the legal framework and its implementation might require tightening up, to more effectively deter breaches, for example in data protection. However, care needs to be taken to distinguish between those that engage in criminal behaviour for commercial or political gain as against those that simply struggle to effectively implement the legal framework, for example due to complexity, lack of funding or uncertainty.

[4] Business conduct/ethics, for example, https://www.ifa.org.uk/technical-resources/ethics

[5] This is different from transparency as explored later. It is ensuring there is clarity on the 'politics' (objectives, mission, interests) of the think tank, for example, through the provision of a (binding) mission statement.

[6] This framework is not yet final, as at this writing, and will be adjusted according to comments and thoughts from different groups of stakeholders (including think tanks).

[7] See: https://prores-project.eu/framework/

[8] Further detail is available from: https://prores-project.eu/

[9] See: https://prores-project.eu/framework/

Reference

Lux, N. (2021) 'Außenpolitische Think-Tanks in Krisenzeiten, Zu Selbstverständnis und Relevanz von Denkfabriken: Debattenbeiträge seit 2017' [Foreign policy think tanks in times of crisis], *SWP-Zeitschriftenschau 2021/ZS*, 1, available from: https://www.swp-berlin.org/en/publication/foreign-policy-think-tanks-in-times-of-crisis [in English], [accessed 31 January 2022].

PART II

Ethical evidence in specific policy domains

7

Ethical research in public policy

Paul Spicker

Introduction

The main areas of public policy research might be thought of as:

- the study of the policy process, focusing on the factors and influences which shape public policy;
- public management or public administration, which are mainly concerned with how public organisations are run;
- policy analysis, which is about examining policy and finding out what it actually does; and
- policy advocacy, making the case for particular policies. This is often the province of 'think tanks' and lobbyists, but their work is based in research, too.

Some studies are mainly concerned with how policymaking works – research into policy – while others look for evidence that that can be used directly in the service of specific policies – research for policy. Peter John, working in the first paradigm, describes policy research in these terms:

> Research on public policy seeks to explain how decision makers … produce public actions that are intended to have an impact outside the political system. … Researchers in public policy aim to explain how public decision making works, why societies get the policies they do, and why policy outputs and outcomes differ from place to place and across time. (John, 2012: 1)

My own work has mainly been in the field of policy analysis, focusing on problem-based research intended to inform policies undertaken by specific agencies. I first came to this work because I was interested in poverty and inequality, not because I really wanted to study policy. It is probably true to say that many of the people who come to policy studies do so because they are committed to the study of specific issues – equalities, the environment, gender and human development are examples – and gathering insights that might advance the study of the policy process is not the main issue. It is much more important to find out what policy actually does.

It is possible to discuss policy as a 'black box', looking only at what goes in and what comes out – that is a useful starting place for evaluations – but characteristically, to understand what is actually happening, a policy analysis will also need to consider issues relating to the process of implementation. Process matters. One of the most basic truisms of policy studies is this: there is no such thing as perfect implementation. The lengthy subtitle of Pressman and Wildavsky's *Implementation* tells a story:

> how great expectations in Washington are dashed in Oakland; or, why it's amazing that federal programs work at all, this being a saga of the Economic Development Association as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes. (Pressman and Wildavsky, 1992)

Wishing and hoping that policies might work is not enough to make things happen – that is one of the reasons why policymakers need policy analysts, to tell them what is actually happening. The authors are unequivocal: their contribution is fundamentally an ethical activity. They are right to think so.

Policy analysis as an ethical activity

Policy research in general, and policy analysis in particular, is research for a purpose. Studies of policy are guided by a distinctive range of normative considerations; those considerations go to the heart of the subject.

There are three core moral principles or, more accurately, three clusters of related principles, which have to be recognised. The first, which is liable to be underestimated in a 'pure' academic environment, is about public service. There are some social scientists who think that the purpose of their work must be to produce generalised insights that are 'value-free' or neutral, stripped of any kind of moral judgement: that used to be said of sociology (for example Seubert, 1991), and it is still a common claim made for economics, in despite of the public roles that economists commonly occupy (Mongin, 2006: 258–9). The position of people from diverse fields working in public administration and policy analysis is somewhat different. The American Society for Public Administration recommends a range of principles, including among others advancing the public interest, upholding the law, strengthening social equity and promoting democratic participation (ASPA, 2014). Under the third of these heads, for example, they enjoin public servants to 'Treat all persons with fairness, justice, and equality and respect individual differences, rights, and freedoms. Promote affirmative action and other initiatives to reduce unfairness, injustice, and inequality in society'.

Any list of general principles of the sort is going to be indicative rather than comprehensive. Typically the moral duties of someone working in public services will include:

- requirements for the agency to act ethically in relation to the community which it serves. This commonly includes duties relating to rights, fairness, regard to due process and duties of care;
- the responsibilities of public servants to the agency, including contractual obligations, fulfilment of designated roles, legal obligations and accountability;
- individual requirements for ethical conduct, such as professional competence, integrity and respect for persons; and
- ethical constraints not to abuse one's power or position: for example, impartiality and setting aside personal interests.

Policy research cannot sensibly be disentangled from all this. Sometimes public administrators will initiate or commission research; sometimes researchers will themselves be administrators in their own right, who need to find things out about policy as part of their job. It is difficult to see how a researcher could work with an agency committed to principles of fairness, respect and the public welfare and yet still claim to be maintaining a morally neutral stance. Policy analysis can be 'outsourced', but it cannot be treated as a moral void; the duties associated with public service pass as part of the engagement.

The second set of principles relate to the 'public' nature of the activity – the public sphere. The guidance on ethical research from learned societies has so often concentrated on issues of privacy and the rights of participants in research that it has been liable to overlook, or even to lay aside, the public character of much research activity. In a separate contribution to the PRO-RES agenda, I have argued that a substantial range of issues in research fall clearly into the public arena: secondary analysis, the use of published material, the process of generalisation itself (Spicker, 2021). Most social scientists have no compunction about using material in the public domain which is derived from sources that may themselves be based in unethical conduct – the product of scandals, public enquiries, criminal trials, along with some questionable classic studies in sociology and psychology. Sharing material in the public domain is fundamental to learning, intellectual development and the advancement of science – and the claim to privacy in such circumstances is often morally offensive.

In the specific context of policy research, however, I need to go further. There are examples of policy analysis that are confidential or private, but the most basic assumption to make about most activities relating to public agencies is that they should be considered to be public, unless there are powerful reasons to the contrary. Examples might include:

- examining laws and regulations
- reporting legal cases
- using publicly collected data as a basis for policy analysis
- examining processes of accountability, financial probity or service delivery, and
- testing whether or not the actions of public agencies are done in conformity with their policy.

The public nature of this kind of activity is rooted in the sources of the information: if the information has legitimately been made public, it is legitimate to use it. There have been circumstances where governments have refused to release information they hold – sometimes on the basis that it belongs to individuals, even if that information is aggregated and anonymised, sometimes on the basis of commercial confidentiality. There has been a disturbing trend in discussions of research ethics to treat public information as if it was, in some way, private and confidential, and that the people referred to should be protected from the consequences of public criticism. That view that has been particularly influential in relation to criminology, where many researchers seem to think that research participants have a right to keep their criminal activities secret (for example Loxley et al, 1997; Pearson, 2009; and see Israel, 2004). They are mistaken. All crime is a public act, because the definition of a crime requires there to be a legal process which makes it so. The Social Research Association, drawing on one of the (rather vague) exemptions in the US Common Rule, suggests that the consent of participants may be required if they seem likely to incur criminal or civil liability as a result of what is found (SRA, 2003: 33).That seems to me to have things backwards – it is just at that point that the information is, and ought to be, considered public. There may be some uncertainty about whether a crime has been committed and, if so, who is responsible for it; some laws are illegitimate, for example because they breach human rights, and there may be room to question whether certain activities are appropriately considered to be criminal; but there should be none about treating a crime as a public issue.

The scope of the public domain is wide: there are many activities which can be considered, in an open society, to be legitimate subjects for public scrutiny. The Canadian Tri-Boards formerly advised that: ‘REBs (research ethics boards) should recognize that certain types of research – particularly biographies, artistic criticism or public policy research – may legitimately have a negative effect on organizations or on public figures in, for example, politics, the arts or business’ (Government of Canada, 2003). They were right. Their current emphasis on ‘critical inquiry’ does not grasp the nettle in the same way. We need to understand that good critical research can legitimately have negative effects.

This leads into a third set of principles: a commitment to democratic governance. Joshua Cohen argues that the open exchange of views is central to our understanding of democracy (Cohen, 1997). Democracy is government by discussion – sometimes called 'deliberative democracy' – a form of government where decisions are accountable and open to scrutiny, discussion and argument. The argument for openness, transparency and the public use of public information leads us to a third set of principles, based on the relationship between the public forum and the character of democratic governance. 'Democracy' has other meanings, of course; the term refers to a broad set of concepts. Some are based in political ideals, referring grandly to concepts such as 'the sovereignty of the people', citizenship or self-determination. Some are primarily institutional, concerned with issues such as majority voting, elections and legal redress. And some define democracy in terms of the process of decision-making – a focus which is central to policy studies. Services and agencies are 'public', not just because some have their origins in governmental or state action, but because they develop their activities for public purposes, generally in line with established norms that make it possible to identify why they are legitimate (Spicker, 2009). Their course of action can generally be traced to a source of authority – possibly an external source, such as a legislative framework, the circumstances their foundation or the terms on which they are funded; but it is no less possible that the source is internal, such as a constitutional document and a set of corporate decisions which are compatible with those terms of engagement. There is a 'golden thread' running through services in government, charity and independent public services (Goodin, 2003). This is sometimes boiled down to a quasi-legal emphasis on 'constitutionalism', but it is more basic – and older – than that term suggests; it is an important element in the principle of 'the rule of law', the idea that authority depends, not on the say-so of leaders, soldiers and charismatic individuals, but on a framework of rules (Hart, 1961) – a tissue of conventional understandings as to what is, and what is not, consistent with that framework. It is a central feature of most Western democracies that the exercise of authority is legitimate only when it is done in accordance with the law. That principle is protected by a structure of accountability, where the exercise of power has to be justified, and might always be subject to a higher authority.

Policy research matters because it informs the process of deliberation, and policy advocacy is itself part of the democratic process. In general, too, policy research plays a vital critical role: the primary role of the analyst, Wildavsky argues, is to speak truth to power (Wildavsky, 1993). Majchrzak argues that the process of policy research is both moral and political: moral because it is imbued with values, political because it has to be negotiated in political forums in order to be effective and worthwhile (Majchrzak, 1984). Policy research comes with a sense of public duty, and failure to engage with the moral issues is in itself a moral failure.

Human subjects in policy research

It is common in discussions of research ethics to frame the criteria in terms of research with 'human subjects' (for instance, National Commission for the Protection of Human Subjects, 1979; COUHES, 2019), but that does not fit well with policy research. The main subject of policy research is policy, rather than individual human beings, and the thing that is being investigated is often – not always – how the policy is done, and what effects it has. Policy research necessarily relies, to some degree, on interaction with human respondents; information will typically be gathered both from people who operate policy, and people affected by it, such as service users – many people working in the field would see engagement with the people who are most affected as a moral imperative in itself. In my own work, as the director of a small, jobbing research unit, I have undertaken a range of projects that were focused on the responses of the people affected: for example, the experiences of people in poverty, the views of schoolchildren about the design of their schools, the safety of minority ethnic groups. The third-mentioned project asked participants for open-ended, qualitative accounts of how they were treated by public services. I learned, for example, how calls to police services for urgent help had been ignored; how little was being done in schools about racist bullying; or how Gipsy Travellers were received when trying to get medical care. The point of doing this is not better to understand what sort of people they were – the human beings were not the subject – but to find out more about the operation of public services. (The response of the police force, as it turned out, was expressed as a particular matter of concern, and was subsequently the subject of some criticism by their inspectorate.)

If this is not about human beings, who is it about? It is just as likely, possibly more so, that the focus of policy research will fall on an organisation. I've undertaken evaluations, for example, of anti-poverty policy, the coordination and local delivery of medical training, an employability project for homeless people or allocations by a social housing provider. The central question in such examples is how an organisation (or, in the case of anti-poverty policy, a group of organisations) is performing. Organisations are collective entities, not human beings. They are a social reality, rather than a physical one. Our day-to-day lives, businesses, utilities, schools and contacts with government are pretty much unavoidable. Organisational behaviour is pervasive in contemporary, developed societies: if you have eaten today, or used an electrical appliance, or received information or mail, it is almost certain that you will have had contact with an organisation that made that possible. The organisational focus is not unique to policy research. It is a routine part of a range of academic fields – business studies and management, political science, social policy, and law among them. Other fields of study call at times for research on issues affecting organisations in their specific area of interest: for example, in health studies, social work and education. Each of these fields has further ethical dimensions to consider in their own right, commonly covering not just the impact on the people and organisations directly engaged in the research, but the broader implications of research, both as a process and as a means of developing the way that such activities are understood.

Organisations make their decisions as collective units: in any formally constituted organisation, there has to be some kind of structure or process that makes it possible to distinguish collective decisions from individual behaviour. Many organisations are 'corporate': they have a legal personality, so that they can register and exercise rights that otherwise would be confined to human beings – for example, employing people, holding a bank account, or buying and selling services. People within organisations occupy specific roles, and their roles shape their choices and course of action. When workers within an organisation participate in research, they generally do so in their capacity as employees or officials, not as individual human beings. And, in so far as decisions are made by the organisation, they are not made by those officials or employees. If the organisation has given consent to participation in research, it is unusual for its employees or officials to have any choice about it. The constraints typically reflect contractual relationships – employees are subject to direction; the obligations associated with the roles that people occupy; and legal requirements, such as Freedom of Information legislation. That means that the common strictures that have been applied to 'research with human subjects', about issues such as voluntary participation and consent, make little sense in this context. In many if not most cases, they simply do not apply.

Research into 'human subjects' generally seeks to avoid harm to people participating in research studies. The focus on the organisation, rather than the individual respondents, means that there is often a possibility of harm to specific people – most usually, the employees of that organisation. If I engage in the evaluation of a policy, of an organisation or of the use of funds, there must be the possibility of a negative finding or judgement – that something has not worked, that it might have failed to achieve its objectives, that it might be discontinued, that another approach may be more effective. Some of these evaluations are commissioned or required by funders – it is not uncommon for projects in the UK to be given time-limited funding for three years. The process of evaluation puts people's jobs and prospects on the line. Make no mistake: one of the consequences of conducting an evaluation is that some people might be hurt. It goes with the territory. If policy researchers and analysts cannot live with that, they need to find another line of work.

Policy analysis has been, in some circumstances, treated as exempt from ethical scrutiny. The so-called 'Common Rule' applied by US federal government agencies generally treats the analysis of government-funded research that are investigating governmental as something beyond ethical scrutiny. The exemption extends to research

> designed to study, evaluate, improve or otherwise examine public benefit and service programs, including procedures for obtaining public benefit or service programs, including procedures for obtaining benefits or services under those programs, possible changes in or alternatives to those programs or procedures, or possible changes in methods of levels of payment for benefits of services under those programs. (DHHS, 2018)

US law has a track record of dealing with thorny moral issues by throwing a blanket over them and telling them to go away – the refusal to legislate about the press and the absence of a legal framework relating to religion are examples – so it is not that surprising that in a legalistic set of rules, the moral issues should simply be ducked. It is understandable too, I suppose, that US government departments do not want to pay for an evaluation, only to discover that the researcher is unable to provide one because of an ethical concern that no one should ever have a service withdrawn or denied because they would then be worse off. However, the lack of engagement with ethical problems is disquieting. There are strong reservations to make about any approach that treats a sensitive, threatening, potentially harmful process as being beyond ethical scrutiny. It is precisely in those circumstances that ethical consideration matters most.

Ethical policy research

The values that guide policy research might be understood, at least initially, in terms of a framework of deontological principles. The first dimension to consider is beneficence. In the context of research, beneficence is often assumed to apply to participants and respondents; sometimes, it is reduced only to a principle of non-maleficence (do no harm). In the context of public policy, the demand for beneficence extends far more generally, to the potential consequences of research – consequences for the public, for specific groups targeted by public services, for people who are vulnerable (that is, liable to be harmed disproportionately), and for those who are excluded or marginalised.

Second, there are democratic norms to respect, covering issues such as legitimate authority, transparency, deliberation and the rule of law. 'Accountability' is usually interpreted in terms of the structure of authority: the actions of officials are authorised by senior officials, who have in their own right been authorised through a set of formal structures, and in the context of government, it can be argued, ultimately by an electoral process. (In recent years, there has been a disturbing trend in the US and UK for senior officials to arrogate to themselves the role of 'leadership', in imitation of private entrepreneurs. The claim is vague, but also ethically questionable; it is probably inconsistent with democratic accountability [Cook, 1989; Spicker, 2012].)

Third, there is citizenship – 'the right to have rights'. In a democratic society, members of the public come in contact with public services, not as subjects or supplicants, but as citizens endowed with rights, entitlements and a guarantee of due process. There are exceptions: some services are delivered on a discretionary basis, and in some cases such as prisons citizens may be considered to have forfeited their rights; but there is still a presumption that some rights will remain inviolate, and a guarantee in others that legitimate interests will be safeguarded. There are often particular grounds for concern when citizens are excluded or disempowered; there is a strong, distinct moral argument for 'voice' and for measures to address disadvantage and exclusion.

Last, but not least, there are further principles governing process and procedure. In a democracy, every institution and agency depends on a structure of legitimate authority, established within a legal framework. There are likely to be a series of constraints – institutional practice, conformity with financial requirements, bureaucratic norms, legal rules and professional standards. Typically all of these will apply: there is not one form of accountability, but many. Those constraints will themselves be backed up by moral imperatives. A due respect for the rules goes beyond considerations of effectiveness or public benefit; they define the limits of what makes a policy legitimate. The rules demand respect even in circumstances where they appear to be counterproductive, and respect for those constraints has all the force of other moral duties.

Research in public policy is infused with a sense of its ethical force, but the kinds of ethical criteria that are raised by such research are somewhat different from the issues which are commonly considered when people write about 'research ethics'. There are two key lessons to draw. The first is that the primary focuses of moral concern in public policy research are not the same as those considered in existing codes of ethics. Most conventional guidance on research tends to focus on a relatively limited set of principles, considering the impact of research, the protection of research participants, respecting academic disciplines, and the research relationship. There are implicit, and sometimes explicit, assumptions, that the nature of research is essentially private, that it is being done with individual human subjects, and that the primary duty of researchers is to protect the interests of research participants. These issues are not negligible, but they fall a long way short of a full consideration of the ethical dimensions of policy research in practice. Research in public policy is far more commonly undertaken in areas of public interest; it relates to social and public roles; and the duties of the researcher are likely to fall, not only to agencies and institutions, but towards the people those institutions are supposed to serve – citizens, service users and, quite possibly, towards people who are excluded, who the services and agencies have previously failed to reach. Most discussions of research ethics are concerned with the research process – how information is obtained, and under what circumstances. In public policy research, by contrast, the political and moral context of the research, and its potential application, matter far more.

The second lesson is that the values that infuse public policy research are rather poorly served by existing procedures of ethical review. It does not help, of course, that the attention and emphasis given to ethical scrutiny is so firmly focused on other issues, but the gaps could not be closed simply by extending the remit and agenda of review bodies. More fundamental is that there is very little in the principles that guide ethical policy research – beneficence, democracy, citizenship and legitimacy – that could be adequately respected and safeguarded mainly by reviewing an initial proposal, or any single-stage process undertaken prior to the implementation of a research programme. The design of a research project cannot be the only time at which the ethical dimensions of policy research are considered: it matters critically how interactions during the research are guided by ethical principles, what the research actually reveals as it proceeds, and how the findings are processed and presented. Researchers have to be aware of the ethical dimensions of their activity. At every stage of the research – design, testing, engagement, gathering information, processing it, reporting and dissemination – those dimensions will need to be taken into account.

Conclusion

While the issues that are identified here have been expressed in terms of principles and duties, it does not follow that they are best served by being treated as (yet another) set of rules. They are too wide-ranging, and raise too many interwoven issues, for that. Researchers need to take account of the principles they need to consider; the duties that should guide their actions – and those they might incur during the process; and the things they should not do. They need to be sensitive to potential impacts, the possibility that the things they do may have bad effects; to risk, which is the damage they might do; and to vulnerability, which is the extent to which the people affected can recover from that damage. Policymakers and others using public policy research likewise need to consider the ethical issues applicable to this type of research. No one can anticipate every eventuality. The scope of ethical conduct is complex, many-faceted and open to further issues arising. Following the argument for virtue ethics, the engagement needs to offer the opportunity for a process of reflection on, and discussion of, ethical issues that may well in practice have no clear or authoritative resolution. It follows that ethical research calls for continuous, discursive engagement. It is the only effective way to ensure that ethical issues are given the time and space to be considered thoroughly.

References

ASPA (American Society for Public Administration) (2014) 'Code of ethics', available from: https://www.aspanet.org/ASPA/Code-of-Ethics/ASPA/Code-of-Ethics/Code-of-Ethics.aspx [accessed 31 January 2022].

Cohen, J. (1997) 'Deliberation and democratic legitimacy', in R.E. Goodin and P. Pettit (eds) *Contemporary Political Philosophy: An Anthology*, Oxford: Blackwell, pp 143–55.

Cook, B.J. (1989) 'Politics, political leadership, and public management', *Public Administration Review*, 58(3): 225–30.

COUHES (Committee on the Use of Humans as Experimental Subjects) (2019) 'Overview', available from: http://web.mit.edu/committees/couhes/index.shtml [accessed 25 January 2022].

DHHS (US Department of Health and Human Services) (2018) '2018 Requirements (2018 Common Rule), subpart A: Basic HHS Policy for protection of human subjects', Washington, DC: HHS Office for Human Research Protections, available from: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/revised-common-rule-regulatory-text/index.html [accessed 31 January 2022].

Goodin, R. (2003) 'Democratic accountability, the third sector and all', JF Kennedy School of Government, Harvard University, available from: https://cpl.hks.harvard.edu/publications/democratic-accountability-third-sector-and-all [accessed 17 March 2022].

Government of Canada (2003) *Tri-council Policy Statement: Ethical Conduct for Research Involving Humans*, section G, Ottawa: Government of Canada, available from: https://web.archive.org/web/20040208144509/www.pre.ethics.gc.ca/english/policystatement/policystatement.cfm [accessed 25 January 2022].

Hart, H.L.A. (1961) *The Concept of Law*, Oxford: Oxford University Press.

Israel, M. (2004) 'Strictly confidential?', *British Journal of Criminology*, 44(5): 715–40.

John, P. (2012) *Analysing Public Policy* (2nd edn), Abingdon: Routledge.

Loxley, W., Hawks, D. and Bevan, J. (1997) 'Protecting the interests of participants in research into illicit drug use', *Addiction*, 92(9): 1081–5.

Majchrzak, A. (1984) *Methods for Policy Research*, London: Sage.

Mongin, P. (2006) 'Value judgments and value neutrality in economics', *Economica*, 73(290): 257–86.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979) *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, available from: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html [accessed 31 January 2022].

Pearson, G. (2009) 'The researcher as hooligan', *International Journal of Social Research Methodology*, 12(3): 243–55.

Pressman, J. and Wildavsky, A. (1992) *Implementation* (3rd edn), Berkeley: University of California Press.

Seubert, V. (1991) 'Sociology and value neutrality', *The American Sociologist*, 22(3/4): 210–20.

Spicker, P. (2009) 'The nature of a public service', *International Journal of Public Administration*, 32(11): 970–91.

Spicker, P. (2012) '"Leadership": a perniciously vague concept', *International Journal of Public Sector Management*, 25(1): 34–47.

Spicker, P. (2021) 'Research is not a private matter', in R. Iphofen and D. O'Mathúna (eds) Ethical Issues in Covert, Security and Surveillance Research, Bingley, UK: Emerald, pp 29–40.

SRA (Social Research Association) (2003) *Ethical Guidelines*, available from: https://the-sra.org.uk/common/Uploaded%20files/ethical%20guidelines%202003.pdf [accessed 25 January 2022].

Wildavsky, A. (1993) *Speaking Truth to Power* (2nd edn), New Brunswick, NJ: Transaction Books.

8

Bitter harvests: ethical divides in the GM foods debate

Alan Simpson

Introduction

In September 2020, BBC Radio 4's *The Reunion* programme revisited some of the controversies surrounding Britain's original GM (genetically modified) foods debate in the 1990s (BBC, 2020). It is a programme that will turn out to be more prescient than nostalgic. As British ministers increasingly look for half-credible, post-Brexit trade deals, the GM debate will come back with a vengeance. All the ethical issues that went unresolved in earlier debates about the science (and the politics) will resurface.

As an active UK Member of Parliament (MP) at the time I was knee-deep in the original controversy surrounding the ethics, politics and efficacy of GM crops. Participating in the Radio 4 programme reminded me of just how fragile societies become when science and politics become detached from ethics. Activists get pitted against scientists, ecology against the advancement of knowledge, consumers against food suppliers. It ends up as a battle between binary choices rather than the pursuit of holistic and systemic solutions. In any new debate about food standards, environmental security and GM technologies, it's worth reflecting on lessons we ought to have learned from last time round.

The embodiment of evil

What disturbed me most about the original debate was the extent to which science itself became seen as the embodiment of evil. At the time, this was true, but only because the science of biotechnology became the plaything of corporate interests. Politicians failed to see how this this developed and how they too were being played. The biggest biotech companies seeking to manipulate the public debate wanted to own the food chain, not feed the planet. The race was on to deliver the corporate ownership of nature. 'Getting to market first' became the overriding corporate priority. Scientists as much as politicians were pawns in the process.

The ethics of genetic manipulation then became detached from its economic and ecological consequences. Research appraisals were lost in technical details rather than their ethical implications. Key presumptions about the basis of human existence on this planet were largely marginalised. As one unnamed farmer at a public meeting put it: "Agriculture used to be about learning how to dance with nature, now the pressure is on just to stomp all over it."

This was where industrial agriculture had taken us. Monoculture was overtaking agriculture. Small, sustainable, rotational farming was being defined as incapable of feeding the planet. Big agribusiness became the preferred answer, if only it could bump up the yields and get rid of the pests. Research justifying it became quantitative (in yields) not qualitative (in impacts).

Corporate giants fed the urgency and fragility of food security debates in order to build their new dominions. In doing so they ran the danger of defrauding both science and society. In the wake of the COVID-19 pandemic the same (disingenuous) arguments are likely to reappear.

Back to the future

The propagation of seed varieties is not new. Since humans began to farm the land farmers have saved, shared and propagated seed varieties. These now run into their millions. Seed varieties are stored in over 1,000 seed banks around the world. The largest of these – known as the Doomsday Vault – is in Svalbard, Norway. The vault contains more than 1,050,000 seed varieties, from staple crops like maize, wheat and rice, through to vegetables, herbs and wildflowers. We are not short of seed varieties, merely the wisdom of how to share, save and use them.

Historically, most seed propagation has been within species limits. What GM offered was the possibility of crossing species boundaries; taking part of the gene sequence of one species (say a fish) and putting it into the gene sequence of an entirely different entity (a staple crop or flower). The science was (and is) devilishly clever. Where it was weak to non-existent was in the ethical scrutiny of its consequences and liabilities.

Around the same time as the GM debates, Mercedes had launched a swish new vehicle called the A Class. The only trouble was that it had a habit of falling over when turning corners. The huge furore that followed triggered a complete product recall by the manufacturers. Anti-GM campaigners were quick to point out, that in nature there are no such recall mechanisms. As the COVID-19 pandemic has painfully demonstrated, once unleashed, the public debate becomes more about societal recovery than product recall. It makes the science of prior scrutiny and environmental security all the more important.

American biotech corporations had stolen a march on the GM debate long before seeking access to European markets. They convinced US regulators to take the technology out of the 'medical' appraisals process – usually involving extensive clinical trials – and evaluating it instead under the category of 'novel foods'. A revolutionary technology thus ended up being 'tested' as little more than a new variety of cake decoration. It left both researchers and policymakers in a double bind.

In ethical terms, debates about the evaluation and scrutiny of GM crops never recovered from being removed from the medical ambit. Nor were their merits judged against key elements of environmental law. These are enshrined in what is best known as the 'precautionary principle' in environmental science, helpfully summarised in the US National Institute of Health's journal as follows: 'The precautionary principle … has four central components: taking preventive action in the face of uncertainty; shifting the burden of proof to the proponents of an activity; exploring a wide range of alternatives to possibly harmful actions; and increasing public participation in decision making' (Kriebel et al, 2001: 871).

No elements of this were incorporated into the research ethics surrounding the appraisal of GM crops. As long as they were 'distinct, useful and stable' GM crops in the US were given the green light for production, distribution and (patented) use. Longer-term impacts on human and environmental health never got a look-in. Even 'stable' took on an ambiguous character as GM crops quickly developed their own vulnerability to new blights, declining yields and variable weather conditions. But by that time corporations were already hoping to own the entire farming game.

In the process, science itself was being reduced to the 'novel', health interests reduced to the marginal and markets structured in favour of the unaccountable. US companies had secured the right not even to label foods containing GM products. So much for informed consumer choice.

The world owes a huge debt to the environmental activists who, in the 1990s, ensured that a GM debate even took place. Some research scientists tried to join in, but those raising doubts about the state of the science were roundly denounced by others making the corporate case for convenient compliance. Those who did, quickly lost research funding. We now know that huge amounts of money were thrown into attempts to persuade governments to fast-track approval of 'patented' GM crops. Fortunately, the public decided to have nothing of it and governments were forced to follow suit. It was at this point that I became actively involved.

Frankenfoods

As an MP, I was part of a Labour government that was about to offer fast-track approval to the use of GM crops in Britain. The ministry responsible for food, farming and agriculture (DEFRA) attempted to grant licences for the use of such seeds, based largely on the US approvals. To prevent this, a grassroots environmental campaigning movement – Friends of the Earth – launched a High Court challenge which I (and two MPs from other parties) co-signed. The successful legal judgment forced the government to accept that more rigorous scrutiny was required.

At the same time, environmental activism was taking the whole question of food accountability (and informed consumer choice) to a different level. Supermarket protests about GM 'Frankenfoods' became commonplace. We organised constructive 'shop-ins', where volunteers filled shopping trolleys with food products, only to insist at the checkout that cashiers removed any item that might include 'any of those horrid GM ingredients'. The colossal queues and disruption that followed quickly saw supermarkets retreat from a position of indifference to GM labelling to one of active product separation. Seeing products sidelined and spurned, food producers rapidly took the same view.

Government enthusiasm for GM crop trials was dealt a further blow by activists turning up to remove the crops from test sites. I have to own up to a degree of disappointment about my part in this. As a protester, I had taken a supply of organic seeds along with me. I wanted to plant these in place of the GM crops I was pulling up. When arrested and charged, I hoped to have the opportunity to plead guilty, not to the disruption but to planting organic crops that would be safe for my children to eat. I even notified the Chief Constable about my intentions (and where his officers could apprehend me).

After listening patiently, the Chief made it abundantly clear that, in the hierarchy of crimes he had to deal with – murder, robbery, drugs networks, organised crime, violence and fraud – deploying officers to chase me round a field, just to provide the political platform for a day in court, did not rank highly among his priorities. I was told, in no uncertain terms, what to do with my offer of peaceful arrest. Other protesters were more successful.

There are two important points behind this melee of activities. The first is that it raised a public debate about the 'precautionary principle' in science; that the more radical the innovation, the more rigorously its consequences should be anticipated and tested before general release. The second point is the 'dancing with nature' question: where do policymakers draw the line between owning ecosystems and stewarding them? How then do you draw up relevant research/policy criteria that avoids the trap of short-term corporate profiteering? These questions are just as pertinent to the lives of farmers in the Global South as to those in the Industrial North.

The ethics of insecurity

When biotech corporations failed to convince Western societies about the desirability of GM foods, the lobbying emphasis changed tack. Their ploy was to trigger a sense of affluent guilt: richer nations may turn their backs on

GM crops, but how could policymakers deny the poorest on the planet the right to avoid starvation or malnutrition? The argument was both powerful and fraudulent. Three examples illustrate the point.

GM Golden Rice was heralded as the way to eradicate the blindness afflicting some 600,000 children living on vitamin A deficient diets in the poorest parts of the planet. How could affluent societies deny this to the poor? It fell to another non-governmental environmental organisation, Greenpeace, to point out that you would have to eat huge quantities of the Golden Rice to make good the vitamin deficiency; that malnutrition was about more than a single missing nutrient; and that (in any case) it would be cheaper and easier either to give vitamin A supplements to the poor or help them grow complementary crops (orange sweet potatoes or green leaf vegetables) already rich in vitamin A.

For their pains, Greenpeace were subjected to an industry-orchestrated campaign of abuse – including lobbying by Nobel Prize scientists – before the manifest shortcomings of the technology were scientifically exposed (Hilbeck and Herren, 2016). Subsequent research attempted to analyse the controversy surrounding the whole debate, recognising that pro- and anti-campaigners usually fell into camps that lacked any common interface (Kettenburg et al, 2018), a nightmare for policy evaluation. It largely demonstrated the chasm that separated the molecular from the secular in food security debates. As Hilbeck and Herren (2016) commented: 'Combating hunger and malnutrition one vitamin and mineral at a time is a failed ideology, no matter which vitamin or mineral one starts with, and which kind of delivery system one chooses. Malnourished people do not suffer from single-vitamin-deficiencies added up. They suffer from hunger, as in "lack of food".'

The same basic issues surrounded Monsanto's attempt to promote genetically modified Bt corn in Kenya. Without doubt, Kenyan farmers faced real problems of 30–50 per cent crop losses caused by the corn borer. Monsanto's Bt corn claimed to be blight resistant, offering a huge boost to farmers 'expected yields' (and incomes). This was the case pushed by both Monsanto and the Gates Foundation.

A great deal of subsequent criticism was levelled at Bt corn's damaging impact on other important species (including the monarch butterfly), but none of this dented Monsanto's confidence at the time. They organised a promotional tour in the UK involving some of the Kenyan farmers using Monsanto's crops. This might have gone smoothly had it not been for the fact that we had already gathered evidence about a different approach taken by farmers at the other end of the same plain in Kenya. Their answer had been to plant Napier grass between the rows of conventional corn.

Corn borers are irresistibly drawn to the more succulent stems of Napier grass in preference to the corn. The problem (for them) is that Napier grass releases a gum that traps and kills the corn borer. And when the grass itself dies it feeds its own nutrients back into the soil. What's more, those Kenyan farmers were free from the cost of buying Monsanto's proprietary pesticide. They were also able to save their own seeds, rather than becoming dependent on annual purchases of GM patented seeds. As soon as proposals emerged to bring over the second set of Kenyan farmers, the Monsanto's promotional tour was truncated.

At a policymaking level, however, it highlighted the difficulties officials and politicians faced in making informed, ethical decisions when confronted by widely divergent pressures; the urgency of food poverty crises pulling in one direction, the lobbying of transformative technology owners in the other.

The ownership of life

My final example of ethical dilemmas surrounding GM technologies involved a single farmer. When I brought Percy Schmeiser into the House of Commons to talk to MPs none of us (Schmeiser included) thought the issue would turn into a 20-year saga. But that's what it became.

Schmeiser is a Canadian farmer. He and his wife, Louise, had spent 50 years developing their own varieties of oilseed rape (canola), suited to the farming conditions they faced in Saskatchewan. The Schmeisers saved their seeds and had never grown GM crops. In 1998 they discovered some of Monsanto's GM oilseed rape growing vicariously in their fields and presumed it had blown in from trucks passing by. What followed was a long legal battle in which Monsanto sued the Schmeisers for $1 million for growing patented GM crops without a licence. It was similar to lawsuits Monsanto had taken out (or threatened) against farmers across different continents.

The case went all the way through to the Canadian Supreme Court. In the end, the Court came up with a Caucasian Chalk Circle verdict, in which both sides claimed victory. By a 5–4 majority the Court ruled that Monsanto had seed patenting rights that had been infringed. Then a 9–0 court judgment ruled that the Schmeisers owed Monsanto nothing, not even their legal costs.

This was 2004, and the story ought to have ended there. What followed, however, took it into realms of the surreal. For policymakers, it also raised ethical issues way beyond questions of the science, stability or ecological impact of GM crops themselves.

The $640 war

In late 2005, the Schmeisers found more of Monsanto's GM canola growing in their fields. They notified Monsanto, who duly turned up, tested the crops, confirmed the canola variety was theirs and agreed to remove the uninvited guests. It was at this point that the story became bizarre. As Schmeiser himself later explained:

> ‘They [Monsanto] said, first of all, we’d have to sign a release form. And in this release form it said my wife, myself or any member of our family could never, ever, take Monsanto to court again for the rest of our lives, no matter how much they contaminate us in the future on our land or on this farm …
>
> And the other thing in the release form, they said that our freedom of speech would be taken away. In other words we could never ever talk about what the terms of settlement were.’ (Schmeiser and Goodman, 2010)

Understandably, the Schmeisers declined the offer, saying they and their neighbours would remove the crops themselves and send Monsanto the bill. In the event, the bill came to $640 (plus $20 costs).

It was Monsanto’s response that then raised massive issues about food technology, ecological security and democratic accountability. The company refused to remove the GM canola without a signed agreement; asserting also that the Schmeisers had no right to remove plants that were the patented, legal property of Monsanto. It was a claim that sought to turn contamination into an intellectual property asset; the removal of which was to be made conditional on everlasting gagging orders.

Yet again, the Schmeisers were forced to seek redress (and release from threat and intimidation) through the courts. When, in 2008, the claim finally made it to court Monsanto’s lawyers unconditionally agreed to pay the $640. But this issue was never about the cash. As Percy Schmeiser put it:

> ‘I hope my battle with Monsanto is over. But I realise that as long as I bring awareness around the world about Monsanto’s patent – not only Monsanto’s patent, but Bayer, Syngenta, DuPont – what their patents do for the control of the future of our seed and our food supply, and that’s what it was all about. GMO’s were never meant to feed a hungry or starving world. They were meant to get control of farmers’ seed supply. That gives them control of the world food supply.’ (Schmeiser and Goodman, 2010)

I doubt if any of the scientists undertaking research into genetic manipulation have ever done so with such megalomaniac intent. But this is where society ended up. For policymakers and their advisors, the brilliance and reach of the science calls out for a different set of benchmarks against which use, accountability, appropriateness and ownership all need to be judged.

Conclusion

What the BBC's *Reunion* programme brought home was that (then and now) the debate about GM foods has always lacked a meaningful anchor point. Begin with the wrong questions and you never get close to the right answers. Even within the science of biotechnology itself, the questions differ dramatically as you move from medicine to nature.

In medicine, genetic manipulation may offer vast possibilities for tackling the most debilitating of genetically inherited conditions. It may be no less valuable in the treatment of currently challenging diseases or the development of new vaccines. Even here, though, the question of gene editing raises issues requiring a new framework of medical ethics, setting limits on the redesigning of human life (Hirsch et al, 2019).

But in food sciences the issues are (literally) more earthed. In all the upheavals that the post-COVID-19, climate-torn world is throwing at us, we need new benchmarks against which the next GM-food debates have to be measured. Most do not even begin in the sphere of biotechnology. Although drought, starvation and food shortages are endemic in many parts of the world, there is not a global shortage of food. For policymakers and advisors involved in food security deliberations, biotechnology issues come a long way behind a recognition that:

- currently, one third of all food produced, globally, goes to waste;
- nearly 1 billion people could be fed on less than 25 per cent of the food wasted annually in the US, Britain and Europe alone;
- 25 per cent of the world's fresh water is used in growing foods that never get eaten;
- in the poorest countries, more food is wasted by the lack of storage (or transportation) facilities than the absence of crop varieties;
- in developed nations, over 50 per cent of food waste is in the home; and that
- reducing food waste, localising food systems and switching to restorative agriculture is probably the most effective strategy for tackling climate change.

The COVID-19 pandemic has localised economics (and economic security) in ways we have yet to fully appreciate. The era of food economics, dominated by globalised markets and food multinationals, is being overtaken by more localised (and accountable) food systems. The entire debate about the place of GM crops within this framework needs to be revisited. For policymakers, all the key questions have changed, leading to key ethical guidelines:

- Food security and soil restoration considerations must come before intellectual property rights.
- Environmental impact (in the form of 'the polluter pays' liabilities) must take precedence over the intensification of production.
- Farmers' rights to save (and share) seeds ought to form the basis upon which we make existing seed banks the cornerstones of a new (ecological) global commonwealth.

Both research and policymaking should embrace a recognition that no-till, rotational agriculture offers the most direct way of taking carbon out of the atmosphere, restoring soil fertility and reversing desertification as it does so. These become the most direct answers to today's (and tomorrow's) food poverty. Against such benchmarks, GM crops have contributed little, and in the future may offer less. The most 'successful' of today's GM crop varieties cater for a small cross-section of intensive agricultural production; largely for export markets and/or feedstock. None deliver security to those living on marginal lands or in extreme food poverty. None offer answers to the wider climate crisis. All are structured around the interest of food multinationals, not food communities or sustainable farming.

The next round of food-policy deliberations requires an ethical framework that addresses these challenges: feeding the planet while taking carbon out of the air and returning it to the soil. The COVID-19 pandemic is forcing everyone to address food security in fundamentally different terms. The ethical basis underpinning such a rethink can no longer ignore warnings that industrial agriculture is only 50 harvests away from soil exhaustion. The roller coaster of climate change-related weather events may make this period even shorter.

This was the perspective that politicians failed to adequately address 20 years ago in our original debates about GM crops. To repeat the mistakes today would merely accelerate the coming crisis. Policymakers (and their advisors) may find a better starting point in the approach taken by the international Slow Food movement, whose annual Terra Madre (Mother Earth) gathering in Turin addresses food security through the medium of knowledge-sharing, durability of production and the direct experience of food/farming communities. It is an approach that puts the precautionary principle into practice.

The main limitation of Terra Madre is that its focus has been on the maintenance (and restoration) of food cultures. What we have to grasp is that climate change may already mean that regions, cultures and food communities must become the guardians of someone else's climate future: as our own climate conditions (and seasons) change, so the lessons of cultivation (and crop varieties) need to draw on more internationally comparable traditions from elsewhere. This is where access to global seed banks will become far more significant than the next round of (patented) GM crops.

It is unreasonable to expect Terra Madre to provide this framework, but international policymakers (and institutions) can.

Europe was right in its original decision to use the precautionary principle to limit access given to GM crops and foods. If farming is to operate within its own (annually reducing) carbon budgets – as it must – the same approach should be taken with GM foods. This is not to lock policymakers (and advisors) into an 'anti-meat' position, merely to arm them with questions that address food security and climate stability within a common policy framework. Restorative farming policies do both. Even in the short term, GM food technologies risk delivering neither.

As a *de minimus* position, the ethical basis of support given to any development/deployment of GM crops must be underpinned by strict 'producer liability' legislation. This would ensure that neither farmers nor consumers are exposed to repeated Percy Schmeiser–style legal battles. The liability shoe must be on the other foot. The scrutiny of GM technologies should be tested on the same basis as medical applications. Only then can policy formation draw adequately on scientific (and ecological) evidence on the issue. Only then will policymakers know if the evidence is reliable and whether it can be 'trusted'.

Scientists, science advisors and journalists have an obligation to ensure clarity and openness in debates and discussions surrounding the future use of GM technologies. Building public trust into the process is critical to an understanding of science's role in (sustainably) emancipating rather than enslaving society.

Instinctively, I suspect that policymakers will find all the most stabilising and socially inclusive answers in a re-engagement with approaches to restorative food and farming. Only rigorous and transparent science research – not dependent on or determined by corporate finance – will prove whether I'm right or not.

References

BBC (2020) *The Reunion: The GM Crops Debate*, Radio 4, first broadcast 6 September, available from: https://www.bbc.co.uk/programmes/m000mbpp [accessed 2 February 2022].

Hilbeck, A. and Herren, H. (2016) 'Millions spent, no one served: who is to blame for the failure of GMO Golden Rice?', *Independent Science News*, 10 August, available from: https://www.independentsciencenews.org/health/millions-spent-who-is-to-blame-failure-gmo-golden-rice/ [accessed 2 February 2022].

Hirsch, F., Iphofen, R. and Koporc, Z. (2019) 'Ethics assessment in research proposals adopting CRISPR technology', *Biochemia Medica*, 29(2): available from: https://doi.org/10.11613/BM.2019.020202 [accessed 2 February 2022].

Kettenburg, A.J., Hanspach, J., Abson, D.J. and Fischer, J. (2018) 'From disagreements to dialogue: unpacking the Golden Rice debate', *Sustainability Science*, 13(5): 1469–82.

Kriebel, D., Tickner, J., Epstein, P., Lemons, J., Levins, R., Loechler, E.L., Quinn, M., Rudel, R., Schettler, T. and Stoto, M. (2001) 'The precautionary principle in environmental science', Environmental Health Perspectives, 109(9): 871–6.

Schmeiser, P. and Goodman, A. (2010) 'Percy Schmeiser vs Monsanto: the story of a Canadian farmer's fight to defend the rights of farmers and the future of seeds', *Democracy Now*, 17 September, available from: https://www.democracynow.org/2010/9/17/percy_schmeiser_vs_monsanto_the_story [accessed 2 February 2022].

9

Ethics in smart information systems

Francesca Pratesi, Roberto Trasarti and Fosca Giannotti

Introduction

Big data analytics and social mining raise a number of ethical issues: indeed, the scale and ease with which analytics can be conducted today completely change the ethical framework (Uria-Recio, 2018). We can now do things that were impossible a few years ago, and existing ethical and legal frameworks cannot prescribe what we should do. Artificial intelligence (AI) is becoming a disruptive technology, and resources for innovation are currently dominated by giant tech companies.

In recent years, we have witnessed different initiatives in Europe aimed at providing environments and infrastructures to share research data and technologies, in accordance with the principles of Open Research Data and Open Science. The general idea behind these initiatives is to provide ecosystems for enhancing scientific collaborations among researchers and practitioners, even those from different disciplines. Examples of recent initiatives in different research fields are: EOSC Pilot (2018) and SoBigData (2015; Forgó et al, 2020) (social sciences), SeaDataCloud (Sea Data Net, 2006) (environmental and earth sciences), and IN-SKA (SKA Organisation, 2011) (physical sciences).

In the field of information and communications technology (ICT), a common goal is to achieve responsible research and innovation (RRI) aimed at providing a platform or ecosystem for ethics-sensitive scientific discoveries and advanced applications of social data mining on the various dimensions of social life, or, in other words, social big data science. More and more often, these data regard private aspects of our lives, such as our movements (Inkpen et al, 2018), healthcare (Rodríguez-González et al, 2019), our social interactions and our emotions (Hasan et al, 2019). In this context, it therefore becomes fundamental to take into consideration the legal and ethical aspects of the processing of personal data, especially given the entry into force of the General Data Protection Regulation (GDPR) in May 2018, but also to move forward considering already existing recommendations and exploring the frontier of novel solutions, in accordance with shared societal and moral values. For this reason, it is important that legal requirements and constraints are complemented by a solid understanding of ethical and legal views and values, such as privacy and data protection.

The rest of this chapter is structured as follows. First of all, we analyse the solution related to data protection, describing the general idea and providing an overview of some technical solutions. Then, we focus on the *right to explanation*, listing the most important properties an explanation should have. Again, we move towards a novel model that aims to cover more ethical aspects in the generation of AI systems. Finally, we conclude the chapter with some general remarks of relevance to making policy in this field.

Privacy and data protection

During the twenty-first century, individual privacy has been one of the most discussed jurisdictional issues in many countries. Indeed, the very fine level of detail of data collected by a variety of organisations comes along with potential issues, such as containing and controlling personal information. Consequently, the opportunities to release the knowledge hidden in data bring an increased risk of privacy violation of the people who are represented in it. The threat includes identification of personal aspects of people's lives, such as their home address, mobility habits and religious or political beliefs. Managing this kind of data is not a trivial task. It is not sufficient to rely only on de-identification (that is, removing the direct identifiers contained in the data) in order to preserve the privacy of the people involved. In fact, many examples of reidentification from supposedly anonymous data have been reported both in the scientific literature and in the media, from GPS trajectories (de Montjoye et al, 2013; Hern, 2014) to movie ratings of on-demand services (Narayanan and Shmatikov, 2008) and, even, from health records (Sweeney, 2002).

Several techniques and technological frameworks have been proposed to counter privacy violations, without losing the benefits of big data analytics technology (Fung et al, 2010a). Unfortunately, no general method exists that is capable of handling both generic personal data and preserving generic analytical results. Nevertheless, big data and privacy are not necessarily opposites: indeed, many practical and impactful services can be designed in such a way that the quality of results can coexist with a high protection of personal data if the Privacy-by-Design (PbD) paradigm is applied (Monreale et al, 2014). The PbD paradigm (Cavoukian, 2009, 2012; Cavoukian and Jonas, 2012), introduced by Cavoukian in the 1990s, aims to protect privacy by inscribing it into the design specifications of information technologies, accountable business practices, and networked infrastructures, from the very start. It represents a profound innovation with respect to traditional methods; the idea is to have a significant shift from a reactive model to a proactive one, that is, preventing privacy issues arising in the first place instead of remedying them. PbD has raised interest especially in the last few years because an elaboration of this paradigm is explicitly referred to in the new European GDPR (European Parliament and Council, 2016: 118). Indeed, the new regulation states that controllers shall implement appropriate technical measures for ensuring, by default, the protection of personal data.

The problem of protecting individual privacy when disclosing information is not trivial and this makes the problem scientifically attractive. It has been studied extensively also in the data mining community, under the general umbrella of privacy-preserving data mining and data publishing (Monreale, 2011; Pratesi, 2017; Pellungrini, 2020). The aim of the methods proposed in the literature is of assuring the privacy protection of individuals during both the analysis of human data and the publishing of data and extracted knowledge. Two main families of approaches treat the problem of privacy preservation: anonymity by randomisation and anonymity by indistinguishability. More recently, anonymity by encryption has also become popular.

Anonymity by randomisation

Randomisation methods are used to transform data in order to preserve the privacy of sensitive information, perturbing the data using a noise quantity. They were traditionally used for statistical disclosure control (Adam and Wortmann, 1989) and later have been extended to privacy-preserving data mining problems (Agrawal and Srikant, 2000). In the literature, there exist two types of random perturbation techniques: additive random perturbation and multiplicative random perturbation. In the additive random perturbation methodology, the perturbed dataset is obtained drawing independently from the probability distribution (Uniform or Gaussian) some noise quantities and adding them to each record in the original data set. Thus, individual records are not available, while it is possible to obtain distribution describing the behaviour of the original data set. Moreover, from the perturbed data, it is still possible to extract patterns and models, even if there was the need to develop new data mining approaches to work with aggregate distributions of the data in order to obtain mining results (Agrawal and Srikant, 2000; Agrawal and Aggarwal, 2001; Evfimievski et al, 2002; Rizvi and Haritsa, 2002; Zhan et al, 2005; Zhang et al, 2005). For privacy-preserving data mining, multiplicative random perturbation techniques can also be used. The main techniques of multiplicative perturbation are based on the work presented in Johnson and Lindenstrauss (1984).

Unfortunately, the main problem of randomisation methods is that they are not safe in case of attacks with prior knowledge (Kargupta et al, 2003). To overcome this drawback, a relatively new randomisation paradigm was developed: a recent model of randomisation, though based on different assumptions, is differential privacy. This is a privacy notion introduced by Dwork (Dwork et al, 2006). The key idea is that the privacy risks should not increase for a respondent as a result of occurring in a statistical database. Differential privacy ensures, in fact, that the ability of an adversary to inflict harm should be essentially the same, independently of whether any individual opts in to, or opts out of, the data set. This model is called ε-differential privacy, due to the level of privacy guaranteed ε. It assures a record owner that any privacy breach will not be a result of participating in the database since nothing, or almost nothing, that can be discovered from the database with his record that could not have been discovered from the one without his data (Fung et al, 2010b). Moreover, Dwork (2006) formally proved that ε-differential privacy can provide a guarantee against adversaries with arbitrary background knowledge.

Anonymity by indistinguishability

As already stated, randomisation methods have weaknesses. In some cases, it is better to apply methods that reduce the probability of record identification by public information and that are not data-independent: k-anonymity, l-diversity and t-closeness. The traditional k-anonymity framework (Sweeney, 2000) focuses on relational tables. The basic assumption is that attributes are partitioned in quasi-identifiers and sensitive attributes (Sweeney, 2002). The first kind of attributes can be linked to external information to reidentify the individual to whom the information refers (so-called linking attack); they are available in public such as age, postcode and sex. The second category of attributes instead represents the information to be protected. A data set satisfies the property of k-anonymity if each released record has at least (k − 1) other records also visible in the release whose values are indistinct over the quasi-identifiers. The k-anonymity model usually relies on methods such as generalisation and suppression to reduce the granularity of representation of quasi-identifiers. It is evident that these methods guarantee privacy but also reduce the accuracy of applications on the transformed data. The main problem of k-anonymity is to find the minimum level of generalisation that allows us to guarantee high privacy and good data precision. Indeed, Meyerson and Williams showed that the problem of optimal k-anonymisation is extremely complex to solve (Meyerson and Williams, 2004). Fortunately, many efforts have been done in this field and many heuristic approaches have been designed (see Bayardo and Agrawal, 2005; and LeFevre et al, 2005).

Unfortunately, the k-anonymity framework, in some cases, can be vulnerable (Kifer, 2009). In particular, it is not safe against homogeneity attack and background knowledge attack. The homogeneity attack easily infers the value of the sensitive attributes when a k-anonymous data set contains a group of k entries with the same value for the sensitive attributes.

In a background knowledge attack, instead, an attacker knows information useful to associate some quasi-identifiers with some sensitive attributes. So, he can reduce the number of possible values of the sensitive attributes. Against these two kinds of attack, l-diversity was proposed (Machanavajjhala et al, 2006). The basic idea is to maintain the diversity of sensitive attributes. However, in some cases, the attacker can infer the value of the sensitive attribute knowing the global distribution of the attributes. The t-closeness method (Li et al, 2007) is safe against this kind of attack. It requires that the distance between the distribution of a sensitive attribute in any equivalence class and the distribution of the attribute in the overall table has to be bounded by a threshold t, ensuring the two distributions (the original and the sanitised ones) are quite similar.

Anonymity by encryption and cryptography

Many studies have addressed the problem of supporting query execution on encrypted data. One of the most relevant is homomorphic encryption (Gentry, 2009), which supports computations without decrypting the input. This kind of encryption enables the computation of some operations (such as additions, multiplications and quadratic functions) on encrypted data and generates encrypted results, which, conveniently decrypted, correspond to the results of the same operations performed on the plain text. The weak point of this technique is in the efficiency in the query processing. Other methods append indexes (a sort of metadata) to the data and are useful for executing specific queries (Hacigümüş et al, 2002; Ceselli et al, 2005; De Capitani Di Vimercati et al, 2007). In particular, Hacigümüş et al (2002) explain how it is possible to split a query (translating specific query operations) into a server query and a client query. The first query can be executed without having to decrypt the data, while decryption and compensation query are performed at the client site. Ceselli et al (2005) focus on inference exposure, providing a model to evaluate the trade-off between performance degradation and protection ensured. Finally, De Capitani and Di Vimercati et al (2007) concentrate on data outsourcing and present a solution to the enforcement of access control and the management of its evolution.

A possible use of homomorphic encryption, that can be found in Damgård et al (2012), is in Secure Multi-party Computation (SMC) (Yao, 1982; Goldwasser, 1997), which deals with computing a certain function on multiple inputs in a distributed network. The problem in this case is to compute any probabilistic function on inputs that are distributed among the participants in the network, while ensuring independence of the inputs, correctness of the computation, and that no more information is revealed to participants in the computation. The computation can be carried out by a single participant or by a coalition of participants. As noted in Goldwasser (1997), a trivial centralised solution would be to assume a trusted centre exists, and that all users send their inputs to this trusted centre for the computation of their respective outputs. A preferable option is a solution where trust is distributed. SMC is often used in distributed environments, but regrettably it allows only some kinds of computations.

One of the first techniques is shown in Chaum et al (1988), where participants can share secrets, even if one third of the participants deviate from the protocol (that is based on not leaking secret information and on sending the correct messages). A more recent solution can be found in Gilburd et al (2004), where a new privacy model, k-privacy, is proposed for real-world large-scale distributed systems. They use a relaxed privacy model implementing efficient cryptographically secure primitives that do not require all-to-all communications. Another example is the work of Sanil et al (2004), where they implement a privacy-preserving algorithm of computing regression coefficients, which permits (honest or semi-honest) agencies to obtain the global regression equation as well as to perform rudimentary goodness-of-fit diagnostics without revealing their data.

The right to explanation

The GDPR, in its Recital 71, also mentions the right to explanation, as a suitable safeguard to ensure a fair and transparent processing in respect of data subjects. While privacy and data protection are not novel concepts, and a lot of scientific literature has been explored on these topics, the study of explainability is a new challenge.

So far, the usage of black boxes in AI and machine learning processes implied the possibility of inadvertently making wrong decisions due to a systematic bias in training data collection. Several practical examples have been provided, highlighting the 'bias in, bias out' concept. One of the most famous examples of this concept regards a classification task: the algorithm goal was to distinguish between photos of wolves and Eskimo dogs (huskies) (Ribeiro et al, 2016). Here, the training phase of the process was done with 20 images, hand-selected such that all pictures of wolves had snow in the background, while pictures of huskies did not. This choice was intentional because it was part of a social experiment. In any case, on a collection of additional 60 images, the classifier predicts 'wolf' if there is snow (or light background at the bottom), and 'husky' otherwise, regardless of animal colour, position, pose and so on.

However, one of the most worrisome cases was discovered and published by ProPublica, an independent, non-profit newsroom that produces investigative journalism with moral force. In Angwin et al (2016) and Larson et al (2016), the authors showed how software can actually be racist. In a nutshell, the authors analysed a tool called COMPAS (which stands for Correctional Offender Management Profiling for Alternative Sanctions). COMPAS tries to predict, among other indexes, the recidivism of defendants, who are ranked low, medium or high risk. It was used in many US states (such as New York and Wisconsin), to suggest to judges an appropriate probation or treatment plan for individuals being sentenced. Indeed, the tool was quite accurate (around 70 per cent overall with 16,000 probationers), but ProPublica journalists found that Black defendants were far more likely than White defendants to be incorrectly judged to be at a higher risk of recidivism, while White defendants were more likely than Black defendants to be incorrectly flagged as low risk.

From these examples, it appears evident that explanation technologies can help companies for creating safer, more trustable products, and better managing any possible liability they may have.

The five dimensions of a valid explanation

So far, we analysed the motivation to provide an explanation, which can be both legal and utilitarian. However, from a practical point of view, we need to define some dimensions, useful to understand what makes for a valid explanation. The EU-funded PRO-RES project (aiming to PROmote ethics and integrity in non-medical RESearch) that produced this book hosted a workshop about ethics, social mining and explainable artificial intelligence (ESME 2019). The discussion in this section reports on the group thinking that took place during that event.

First of all, experts tried to define *what is an explanation*, analysing the main characteristics that a good explanation should have:

- *Simplicity*. This is one of the most important properties: the simplest explanation, which requires a minimum cognitive effort to be understood, should be enough. You must be able to reason on the black box model if you are going to understand and to keep all the concepts in your mind.
- *Truth*. This seems trivial, but it must be considered by design: if an explanation is not true, probably there are some biases in the data. However, if you are visualising an advertisement for a wrong reason, it is the classification process that is wrong, while the explanation of why you are visualising the advertisement is still correct.
- *Symbolic*. The explanation should be as general and abstract as possible, and it should possibly imitate human intelligence in the performed reasoning. An example is saliency maps, which are usually a good way to compare algorithms, but they are not symbolic since they only highlight areas or pixels involved in the classification process, without providing any additional information about what that area really represents.

- *High level*, in order to have understandable explanations. This is strictly related to the symbolic property since a more abstract explanation is also generally a higher-level one. Both these characteristics are important because the explanation becomes also simpler to be understood by anyone (see also *Simplicity*, the first bullet point). Consider, as an example, an explanation that is very complex because the explainability model that generated it learned to use too detailed or irrelevant information, generating explanations that are 'overfitted'; in this case, the model can be unable to work properly with new data.
- *Local vs global?* There are different levels to have an explanation: when the explanation is local it is explaining only a single case, while the goal of a global explanation is to recap the overall logic behind a black-box model. The first case is easier for very complex models like neural networks, but to better understand the big picture we probably need something in between, a sort of *sub-global explanation*.
- *Given by causality*, not by correlation, or, even better, by counterfactual analysis or domain adaptation.
- Providing *reasoning and learning* at the same time, taking advantage of multiple data sources (for example, classifying images using both pictures and captions).
- *Actionable*. Indeed, human perception has an element of intuition which is not explainable, or it is very hard to model.
- *Trust*. We must rely on an explanation. Indeed, as highlighted also by Kersting (2020), people are not disposed to forgive a wrong explanation.
- *Stability*. Similar instances should have similar explanation for a given model. A non-deterministic explanation could be easier to provide, but it implies that understanding the model is more difficult; in addition, it violates the property of simplicity and, probably, the trust. Indeed, consistency is a fundamental property also in real life: if a person asks three different doctors for a medical opinion, the opinions must be similar in order for that person to trust them.

Second, the discussion moved on to *how to measure the understanding of an explanation*.

Generally speaking, *we can measure the level of comprehensibility of an explanation as the degree in which humans can replicate the reasoning of the machine*. The measure must be: consistent, trusted, accountable, stable but also monotonic. Indeed, as humans, we accept better explanations which follow a logic: if the measure first grows and then decreases, you will not accept the explanation. The more an explanation is following a certain *monotonicity*, the better it is. The generality of the measure, instead, is not crucial because it depends on the final user of the explanation and on the situation that we are analysing. Regarding the stability, it is preferable that the explanation does not vary too much if parameters vary. Indeed, with adaptable decision algorithms, a small variation in the inputs may change the decision too.

Third, the discussion focused on what are *interpretable data and interpretable models*. It is easy to agree that there is a need to find a trade-off between accuracy and simplicity: models are often so complex that we need to approximate the flows; however, we need to simplify models being careful not to make them too generic. The gender factor could offer a good explanation: if the majority of women in their 30s visualise an advertisement, explaining to them that they are receiving that advertisement because they are women between 30 and 40 years old, and usually women of that age appreciate the offered product or service, seems to be a good explanation. Of course, data (and how it is integrated into the model) also has its importance; for example, we cannot transform a non-interpretable variable into an interpretable one. Clearly, using such explanations is a generalisation, and it could not capture all the characteristics involved (for example, it could be that age is correlated with other variables).

Then, another crucial point regards the *business perspective* of explanation, that is, the implications for companies that have to guarantee explainable and interpretable systems and models, and whether and how these systems can be actually realised in real applications. A big problem is that companies are forced to provide explanations, but they do not want to reveal how the system is reasoning, in order to preserve business strategies and secrecy. Auditors can solve the problem of checking fairness without compromising trade secrets, *but* controllers may feel that providing too detailed an explanation is against their trade interest: providing a lot of detailed explanations to different individuals may disclose the model. A possibility is that an explanation is personal: if a user requests an explanation, it could be based only on their data (even if this is partially in contrast with the principles of stability, and to being sub-global) and it must be revealed just to the user and cannot be shared. One (not very feasible) alternative is to drop not explainable models and only use intrinsically explainable algorithms, but the possibility to use something that can explain an algorithm is substantial.

However, it seems reasonable that users are interested in knowing an explanation of their own situations, while they are not interested in a super-detailed explanation, so the intellectual property of companies seems not to be at risk. In addition, there is the problem that an explanation could require information that is not directly available on site (for example, the economic system is very complex and if you want to explain the price of some products you need to analyse the whole market); we cannot explain every single decentralised node in the big network, but only treat them as a unique giant black box. It is also very important to clarify for which categories of systems we need to provide an explanation. Social network advertising? Or just for loans, mortgages and health-related systems?

Finally, it is important to consider that the final decision belongs to users, and it must be taken by real persons: automatic decision systems should only support decisions. Thus, another crucial question is who are the *final users*, and how the explanation must adapt to its target. An explanation should surely be human-understandable when decision impacts on legal status; however, humans do not need to understand the full model, just why a certain decision was made on them. *Unfortunately*, very often, different types of people require different explanations (for example, diagnosis explained to a patient or to a doctor).

Indeed, the explanation should change along with the background. If we consider developers as the final user, an explanation allowing a prediction of a system's answer is useful enough. An explanation can be useful for debugging (for example, to find bias in the data). A more difficult objective of explanations is their potential social function, that is, a way to suggest to users how to change their behaviour in order to change a system answer and achieve their goals. Tools are needed for different categories of people and different levels of understanding. Sometimes only one factor among many can be given as an explanation. Maybe a solution could be to provide a system that offers the possibility to go gradually in deep: surely the explanation 'you are receiving this advertisement because you are a woman between 30 and 40 years old' is enough for the majority, but if a woman wants to know more, the system could add some information about her web history, 'and you visited the websites X and Y'. Of course, users still have the option (and the right) to contact the data protection authorities if a received explanation does not satisfy them. However, this path is not followed as often: as an example, it is important to point out that in the first year from the entry in force of the GDPR, the Italian Data Protection Authority received zero requests for an explanation. We need to investigate whether explainability is a right that does not interest people or if the general public is simply not aware of this right.

To conclude, a comment at the ESME 2019 workshop mentioned earlier made by Dino Pedreschi summarises this discussion well: 'Explainability is not a value, it is a tool, and we need to understand how to use it.'

Towards ethics by design and Trustworthy AI

With legal frameworks evolution, ethical concerns and guidelines are changing too. As highlighted in the World Economic Forum (2016), this is reflected by social networks continuing to update privacy policies and settings, by newsrooms making frequent updates to publishing guidelines on how they use material sourced from social media platforms, and by the continuous shifts in what is or is not considered appropriate when individuals post on social media platforms. Moreover, both active and passive data collections also raise questions. In this context, the World Economic Forum warns both people and organisations, pointing out that people need to be informed about the potential impact of their content being shared widely. On the other hand, organisations must be honest with the user about when and how the content will be used, and whether it will be syndicated to other publishers or organisations.

The World Economic Forum is not the only entity that invokes transparency. Indeed, transparency is one of the pillars in ethics and it is related to several parts of the big data process, such as seeking permission of users, explanation of terms of use, and data usage after the collection. The Organisation for Economic Co-operation and Development (OECD, 2013), UK Cabinet Office (Government Digital Service Cabinet Office, 2016) and Council of Europe (Directorate General of Human Rights and Rule of Law, 2017) state that notice and consent are fundamental tasks in big data management. They also offer other important considerations about ethics. In particular, De Mooy (2017) gives a good excursus on the history of individual control, on cultural differences between Europeans and Americans and a list of key concepts useful for addressing the challenges of privacy management.

These guidelines are: individual empowerment (through education that teaches individual basic technology and data portability), corporate accountability (through a voluntary, self-regulatory risk assessment) and collective accountability (through government-mandated entities that can assess the impact of any big data process). The OECD framework (2013) is presented along with fundamental principles that should be respected in the data usage process: collection limitation (data collected are the minimum necessary and they must be obtained by lawful and fair means), data quality (personal data should be relevant to the purposes for which they are to be used and they must be complete and up to date), purpose specification (purposes should be specified before any data collection), use limitation (data must be used and disclosed only for the specified purpose), security safeguard (data must be protected by reasonable security safeguards), openness (about development, practices and policies), individual participation (individuals should have the right to control, rectify or have their data erased) and accountability (data controllers should be accountable for complying with measures regarding the other principles). In the UK (Government Digital Service Cabinet Office, 2016), we can find a short summarisation, along with some practical examples of good and bad practices, of the six key principles they consider essential to data management: (1) to highlight the users' need and public benefit from the start of the definition of the methods; (2) to use data and tools with the minimum intrusion necessary; (3) to create robust data science models, analysing the representativeness of the data and the presence of potential discrimination features; (4) to be aware of public perception, understanding how people expect their data to be used; (5) to be clear and open about data, tools and algorithms, providing explanation in plain English; and (6) to keep data secure, following the guidelines provided by the Information Commissioner's Office (ICO, 2017). Finally, the Council of Europe (Directorate General of Human Rights and Rule of Law, 2017) drafted guidelines too. The majority of ethical principles are highly shared among different institutions, and many of them are included in the new EU Regulation. However, in the case of relatively loose regulatory environments, ethical rules are particularly important. Zook et al (2017) listed ten rules for performing ethical research on big data. Some of them are inspired by the concepts already described (for example, inserting ethics directly in the workflow of research or documenting clearly when decisions are made), while others are specifically oriented to research. For example, the importance of debating issues within a group of peers or of sharing data is listed as a fundamental task in some projects, like studies of rare genetic diseases. Last but not least, the 'Ethics guidelines for trustworthy AI' (European Commission, 2019) are a valuable help to researchers. According to the guidelines, Trustworthy AI should be:

1. lawful, respecting all applicable laws and regulations;
2. ethical, respecting ethical principles and values;
3. robust, both from a technical perspective, while taking into account its social environment, since, even with good intentions, AI systems can cause unintentional harm.

One of the most innovative parts of the document is the acknowledgement of potential tensions and the promotion of trade-offs between some ethical imperatives, such as: respect for human autonomy (ensuring respect for the freedom and autonomy of human beings); prevention of harm (and guaranteeing protection of human dignity as well as mental and physical integrity); fairness (both regarding a substantive and a procedural dimension, that is, ensuring equal and just distribution of both benefits and costs and that individuals and groups are free from unfair bias, discrimination and stigmatisation, while seeking effective redress against decisions made by AI systems and by the humans operating them); and explicability (again, processes need to be transparent, the capabilities and purpose of AI systems openly communicated, and decisions – to the extent possible – explainable to those directly and indirectly affected). Moreover, some of the authors (Quintarelli, 2020) stressed some other interesting points: (1) the process automation affects the extent and the speed in the data domain, thus, traditional methods are not working anymore; as a consequence, we need to rethink our rules and the way to assign and verify responsibilities; (2) automated systems at some point could perform wrong predictions or actions, so we need to provide for remedies for such errors, enabling a redress-by-design paradigm; (3) it is not fair to evaluate each single instance of a problem, that is, the personal cases, but we need to evaluate also the benefits of a solution for the whole society.

How to reconcile ethical and industrial objectives

In this subsection, we want to provide an overview of some practical solutions that could help in promoting ethics, even in a business domain. Unless otherwise specified, the discussion reported is again a result of the event ESME 2019, the PRO-RES workshop about ethics, social mining and explainable artificial intelligence.

We start with an example of the application of ethical concerns in everyday life, with particular connection to private companies that affect society, and then we try to figure out some possible general solutions. Applying ethical concerns in everyday life also means that we confront each other also with Web 2.0 and online social networks (OSNs). These relatively new concepts implied clear advantages (reducing distances and democratising the information) but also novel issues. Indeed, several other contemporary problems (fake news is the first among them) are due to the possibility of remaining anonymous. In the real world, actions have different consequences with respect to digital life, and real life often has some form of self-regulatory system (as an example, if I live in a small village and I lie often, I soon become seen as unreliable for other persons). Thus, to lever *individual responsibility* for each action and opinion in the digital world setting too seems to be the right way to proceed. Of course, in a digital work this responsibility level is harder to achieve with regard to the real world, nevertheless, from a technical point of view, mechanisms that regulate this aspect are nearly possible: for example, a certified digital identity can be provided to every online user.

This is particularly important because an OSN is not necessarily a 'bad guy' who tries to break the rules, and it seems unfair that it should be the only entity in charge of supervising the users' contents. In addition, it also seems unfair that the online platform has the responsibility to establish if a certain content is illegal (and several discussions came after, for example, the 2021 US Capitol riot[1]), but an independent authority could help, participating in these disputes; thus, outsourcing of legal decisions might be a solution. Nevertheless, the fact that owners of OSNs are private companies that can and should autonomously decide whether to publish content or not does not remove the responsibility of the role of the company itself in society. Recently, some steps have been made by private companies to enforce control and integrity of published content (Halevy et al, 2020). Of course, a clear drawback in removing anonymity is a possible limitation of freedom of speech and to be too over-blocking, so another aspect that must be also considered is the trade-off between accountability and the freedoms of expression and information (in some countries, anonymity is fundamental to protect users). However, this dilemma is not new since law usually must balance between opposing rights.

The first problem that we need to face is that laws and ethics do have a certain cultural dependency. For example, in the US, nudity is considered a very serious problem, while in India, hate against castes is a sensitive issue. Again, in Italy, one of the major problems is cyberbullying, while Germany has a law against hate speech and fake news. Moreover, the same problem could have different severity: nudity is a concept that can be different in different countries and even in different locations in any one country! Thus, the *need for global ethical values* contrasts with the fact that each company uses data in different ways and operates in different countries, so common ground could not easily be fairly or equitably established.

Nevertheless, given the international reach of much big data *we need an ethical framework of fairly common standards and values*, where legislation is only the basis. Indeed, the GDPR does not cover all the aspects related to data protection; thus, being GDPR-compliant is only the first step. Other aspects that should be considered are:

- *user-centric model*: we need to work for the individuals;
- *substitute privacy with ethics*: indeed, privacy is only one aspect; transparency is one another pillar of ethics, as we have already discussed;
- provide *examples of business models that are ethics-aware*.

A possibility is represented by the positive-sum model (opposite to the zero-sum one), given by the ethics-by-design paradigm. Ann Cavoukian presents this model (Cavoukian, 2018), an extension of the famous Privacy-by-Design one (thanks to which we became finally aware that we can obtain both privacy and utility in machine learning) that also includes transparency, accountability, algorithmic responsibility and security. Dr Cavoukian, who created the Global Privacy and Security by Design organisation,[2] remarks that investing in prevention is more cost-effective in the long term, and she pointed out the importance of evaluating both algorithm and data in the explainability questions. Companies should understand that ethics is an added value, and, in the long run, this is convenient for companies too. An ethics-aware business leads to more trust from customers; this implies that more users will use the company's product/service (the reputation of a company plays a significant role in the acquisition and retention of clients); more users mean more data and, thus, more money for the company. Indeed, all the participants agree that access data is one of the primary goals of each company, and ethics-by-design can help to gain access to data and to manage it at best.

An *alternative model towards a less profit-centred concept of values* is possible, and the participants identify some necessary ingredients:

- *awareness* of people, both from users and people working in companies;
- encourage *interdisciplinarity*;
- *public incentives*, to overcome the general lack of interest from companies;
- *sustainability*: ethical environment is, in all respects, an environment that we need to protect;
- *ethics-by-design*: ethics is an added value – it is a resource and not merely a cost.

Clearly, some costs are still necessary: to create an interdisciplinary team (for example, legal experts for compliance with the law, social media expert for improving the communication of values, ethical philosophers for analysing the whole aspects, computer scientists for implementing solutions), to implement technological tools that help in explaining the behaviour of a black box and tools for ensuring privacy.

The interest from governments in enforcing ethics is crucial: we cannot wait for private companies creating alternatives, for example, to some overused tools. It is obviously inconceivable and utopian that a company develops, for example, a new ethics search engine able to actually replace Google. As academic institutions, instead, we can contribute to this, if there is a clear direction and effort from the EU government. Indeed, a common European ethical framework might also affect other countries and provide a model that can be adopted worldwide (as the GDPR already did).

As academics, we have the *moral obligation to push toward the creation of new models, and we can contribute by providing practical ideas and solutions* (bringing evidence that they can work), so companies could invest in them.

Conclusion

In this chapter, we described progress and open challenges related to ethics in AI systems and machine learning processes. In particular, we gave an overview of what is mainly done: anonymisation (encryption or removal of personally identifiable information), access control (selective restriction of access to places or resources) and policy enforcement (of rules for the use and handling of resources). We also outline the problems of accountability (the evaluation of compliance with policies and provision of evidence) and data provenance (attesting to the origin and authenticity of information). Then, we talked about transparency (explanation of information collection and processing) and explainability (of algorithms, that is, its reasoning or, at least, a justification of a given decision). In particular, we analysed the main characteristics that a good explanation should have, how to measure an explanation, and the business perspective of explanation. We argued that we surely need to provide alternative technologies, but, more importantly, we need to find *alternative business models*, that can be applied by private companies. In such a way, we can finally build an economy that 'works for people' and, as advocated by Ursula von der Leyen, this can permit the move from 'need to know' to 'need to share' (von der Leyen, 2019). Finally, we move the discussion to a more complete ethical approach to AI, considering privacy and data governance in the equation, but also human agency and oversight, robustness and safety, transparency and accountability, non-discrimination and fairness, and societal and environmental well-being.

Indeed, AI is a collection of technologies that combine data, algorithms and computing power. On that basis, as stated also by the European Commission (2020), an AI ecosystem can bring the benefits of the technology to the whole of European society and economy:

- *for citizens* to reap new benefits, for example improved healthcare, fewer breakdowns of household machinery, safer and cleaner transport systems, better public services;
- *for business development*, for example a new generation of products and services in areas where Europe is particularly strong; and
- *for services of public interest*, by improving the sustainability of products and by equipping law enforcement authorities with appropriate tools to ensure the security of citizens, with proper safeguards to respect their rights and freedoms.

Given the major impact that AI can have on our society and the need to build trust, it is vital that European AI is grounded in our values and fundamental rights such as human dignity and privacy protection. Furthermore, the impact of AI systems should be considered not only from an individual perspective, but also from the perspective of society as a whole. As policymakers grapple with these new technologies and applications, careful attention needs to be given to their ethical implications.

Acknowledgements

This work was supported by the European Commission through the Horizon 2020 research and innovation project 'PROmoting integrity in the use of RESearch results (PRO-RES)' under grant agreement No 788352, by the H2020-INFRAIA-2018–2020 / H2020-INFRAIA-2019–1 European project 'SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics' under grant agreement No 871042, and by the H2020 ICT-48 European project 'TAILOR: Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization' under grant agreement No 952215. The funders had no role in developing the research and writing the manuscript.

The authors need to thank all the people participating in ESME 2019: The PRO-RES workshop about Ethics, Social Mining and Explainable Artificial Intelligence. The complete list of people, who provided fundamental inputs to the discussions reported in this manuscript, can be found at https://kdd.isti.cnr.it/esme2019/

Notes

[1] See: https://www.washingtonpost.com/technology/2021/01/09/trump-twitter-banned-apps/ and https://eu.usatoday.com/story/tech/2021/01/08/twitter-permanently-bans-president-trump/6603578002/ and https://www.ctpost.com/news/slideshow/Q-A-How-can-Twitter-ban-Trump-215406.php

[2] See: https://gpsbydesign.org/

References

Adam, N.R. and Wortmann, J.C. (1989) ‘Security-control methods for statistical databases: a comparative study’, *ACM Computing Surveys*, 21(4): 515–56.

Agrawal, D. and Aggarwal, C.C. (2001) ‘On the design and quantification of privacy preserving data mining algorithms’, in P. Buneman (ed), *Proceedings of the Twentieth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 21–23 May, 2001, Santa Barbara, CA*, New York: Association for Computing Machinery, pp 247–55.

Agrawal, R. and Srikant, R. (2000) ‘Privacy-preserving data mining’, in W. Chen, J.F. Naughton and P.A. Bernstein (eds) *SIGMOD ’00: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, May 15–18, 2000, Dallas TX*, New York: Association for Computing Machinery, pp 439–50.

Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016) ‘Machine bias: there’s software used across the country to predict future criminals. And it’s biased against blacks’, ProPublica, 23 May, available from: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [accessed 23 September 2020].

Bayardo, R.J. and Agrawal, R. (2005) ‘Data privacy through optimal k-anonymization’, in K. Aberer, M.J. Franklin and S. Nishio *(eds) Proceedings of the 21st International Conference on Data Engineering, ICDE 2005, 5–8 April 2005, Tokyo, Japan*, Washington, DC: IEEE Computer Society, pp 217–28.

Cavoukian, A. (2009) ‘Privacy by design: the 7 foundational principles’, revised 2011, Toronto, ON: Information and Privacy Commissioner of Ontario, available from: https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf [accessed 2 February 2022].

Cavoukian, A. (2012) ‘Privacy by design [leading edge]’, *IEEE Technology and Society Magazine*, 31(4): 18–19.

Cavoukian, A. (2018) 'AI ethics by design', Toronto, ON: Ryerson University, available from: https://www.ryerson.ca/content/dam/pbdce/papers/AI_Ethics_by_Design.docx [accessed 23 September 2020].

Cavoukian, A. and Jonas, J. (2012) 'Privacy by design in the age of big data', 8 June, Toronto, ON: Information and Privacy Commissioner of Ontario, available from: https://jeffjonas.typepad.com/Privacy-by-Design-in-the-Era-of-Big-Data.pdf [accessed 2 February 2022].

Ceselli, A., Damiani, E., De Capitani Di Vimercati, S., Jajodia, S., Paraboschi, S. and Samarati, P. (2005) 'Modeling and assessing inference exposure in encrypted databases', *ACM Transactions on Information and System Security (TISSEC)*, 8(1): 119–52.

Chaum, D., Crépeau, C. and Damgård, I. (1988) 'Multiparty unconditionally secure protocols', in *STOC '88: Proceedings of the twentieth annual ACM Symposium on Theory of Computing, Chicago, May 2–4, 1988*, New York: Association for Computing Machinery, pp 11–19.

Damgård, I., Pastro, V., Smart, N.P. and Zakarias, S. (2012) 'Multiparty computation from somewhat homomorphic encryption', in R. Safavi-Naini and R. Canetti (eds) *Advances in Cryptology: CRYPTO 2012*, Heidelberg: Springer pp 643–62.

De Capitani Di Vimercati, S., Foresti, S., Jajodia, S., Paraboschi, S. and Samarati, P. (2007) 'Over-encryption: management of access control evolution on outsourced data', in *Proceedings of the 33rd International Conference on Very Large Data Bases, University of Vienna, Austria, September 23–27, 2007*, np [US]: VLDB Endowment, pp 123–34.

de Montjoye, Y.A., Hidalgo, C.A., Verleysen, M. and Blondel, V.D. (2013) 'Unique in the crowd: the privacy bounds of human mobility', *Scientific Reports*, 3: art 1376, available from: https://doi.org/10.1038/srep01376 [accessed 2 February 2022].

De Mooy, M. (2017) 'Rethinking privacy self-management and data sovereignty in the age of big data', Gütersloh: Bertelsmann Stiftung, available from: https://cdt.org/wp-content/uploads/2017/04/Rethinking-Privacy_2017_final.pdf [accessed 2 February 2022].

Directorate General of Human Rights and Rule of Law (2017) 'Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data', 23 January, Strasbourg: Council of Europe, available from: https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806ebe7a [accessed 2 February 2022].

Dwork, C. (2006) 'Differential privacy', in M. Bugliesi, B. Preneel, V. Sassone and I. Wegener (eds) *Automata, Languages and Programming: 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10–14, 2006, Proceedings, Part II*, Berlin: Springer, pp 1–12.

Dwork, C., McSherry, F., Nissim, K. and Smith, A. (2006) 'Calibrating noise to sensitivity in private data analysis', in S. Halevi and T. Rabin (eds) *Theory of Cryptography: Proceedings of the Third Theory of Cryptography Conference, TCC 2006, New York, March 4–7, 2006*, Berlin: Springer, pp 265–84.

EOSC Pilot Consortium (2018) The European Open Science Cloud for Research Pilot Project, available from: https://eoscpilot.eu/ [accessed 2 February 2022].

European Commission (2019) 'Ethics guidelines for trustworthy AI', 8 April, available from: https://wayback.archive-it.org/12090/20201227221227/https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai [accessed 2 February 2022].

European Commission (2020) 'White paper on artificial intelligence: a European approach to excellence and trust', 19 February, Brussels: European Commission, available from: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en [accessed 2 February 2022].

European Parliament and Council (2016) General Data Protection Regulation, *Official Journal of the European Union*, 4 May, L119, available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL [accessed 2 February 2022].

Evfimievski, A.V., Srikant, R., Agrawal, R. and Gehrke, J. (2002) 'Privacy preserving mining of association rules', in O.R. Zaïane (ed) KDD '02*: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23–26, 2002, Edmonton, AL*, New York: Association for Computing Machinery, pp 217–28.

Forgó, N., Hänold, S., van den Hoven, J., Krügel, T., Lishchuk, I., Mahieu, R., Monreale, A., Pedreschi, D., Pratesi, F. and van Putten, D. (2020) 'An ethico-legal framework for social data science', *International Journal of Data Science and Analytics*, 11(4): 377–90.

Fung, B.C.M., Wang, K., Chen, R. and Yu, P.S. (2010a) 'Privacy-preserving data publishing: a survey of recent developments', *ACM Computing Surveys*, 42(4): art 14, available from: https://dl.acm.org/doi/10.1145/1749603.1749605 [accessed 2 February 2022].

Fung, B.C.M., Wang, K., Fu, A.W.C. and Yu, P.S. (2010b) *Introduction to Privacy-Preserving Data Publishing: Concepts and Techniques*, Boca Raton, FL: CRC Press.

Gentry, C. (2009) 'Fully homomorphic encryption using ideal lattices', in M. Mitzenmacher (ed), *STOC'09: Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing, May 31–June 2, 2009, Bethesda, MD*, New York: Association for Computing Machinery, pp 169–78.

Gilburd, B., Schuster, A. and Wolff, R. (2004) 'k-TTP: a new privacy model for large-scale distributed environments', in W. Kim and R. Kohavi (eds) *KDD '04: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle WA, August 22–25, 2004*, New York: Association for Computing Machinery, pp 563–8.

Goldwasser, S. (1997) 'Multi party computations: past and present', in J.E. Burns and H. Attiya (eds) *PODC '97: Proceedings of the Sixteenth Annual ACM Symposium on Principles of Distributed Computing, Santa Barbara CA, August 21–24, 1997*, New York: Association for Computing Machinery, pp 1–6.

Government Digital Service Cabinet Office (2016) 'Data science ethical framework', 19 May, updated 13 June 2018, available from: https://www.gov.uk/government/publications/data-science-ethical-framework [accessed 2 February 2022].

Hacigümüş, H., Iyer, B., Li, C. and Mehrotra, S. (2002) 'Executing SQL over encrypted data in the database-service-provider model', in B. Moon, D. DeWitt and M.J. Franklin (eds) *SIGMOD '02: Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data, Madison WI, June 3–6, 2002*, New York: Association for Computing Machinery, pp 216–27.

Halevy, A. Ferrer, C.C. Ma, H., Ozertem, U., Pantel, P., M Saeidi, M., Silvestri, F. and Stoyanov, V. (2020) 'Preserving integrity in online social networks', arXiv, 25 September, art 10311v3, available from: https://arxiv.org/abs/2009.10311 [accessed 3 February 2022].

Hasan, M., Rundensteiner, E. and Agu, E. (2019) 'Automatic emotion detection in text streams by analyzing Twitter data', *International Journal of Data Science and Analytics*, 7(1): 35–51.

Hern, A. (2014) 'New York taxi details can be extracted from anonymised data, researchers say', *The Guardian*, 27 June, available from: https://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymised-data-researchers-warn [accessed 2 February 2022].

ICO (Information Commissioner's Office) (2017) 'Guide to data protection', updated 29 April 2019, available from: https://ico.org.uk/media/for-organisations/guide-to-data-protection-1-1.pdf [accessed 3 February 2022].

Inkpen, D., Roche, M. and Teisseire, M. (2018) 'Guest editorial: special issue on environmental and geospatial data analytics', *International Journal of Data Science and Analytics*, 5(2/3): 81–2.

Johnson, W.B. and Lindenstrauss, J. (1984) 'Extensions of Lipchitz mapping into Hilbert space', *Contemporary Mathematics*, 26: 189–206.

Kargupta, H., Datta, S., Wang, Q. and Sivakumar, K. (2003) 'On the privacy preserving properties of random data perturbation techniques', in *ICDM '03: Proceedings of the Third IEEE International Conference on Data Mining*, New York: IEEE, pp 99–106.

Kersting, K. (2020) 'Making deep neural networks right for the right scientific reasons', workshop on bias and fairness in AI. BIAS 2000, available from: https://sites.google.com/view/bias-2020/recordings [accessed 3 February 2022].

Kifer, D. (2009) 'Attacks on privacy and deFinetti's theorem', in C. Binnig (ed) *SIGMOD '09: Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data, Providence RI, 29 June 2009–2 July 2009*, New York: Association for Computing Machinery, pp 127–38.

Larson, J., Mattu, S., Kirchner, L. and Angwin, J. (2016) 'How we analyzed the COMPAS recidivism algorithm', ProPublica, 23 May, available from: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm [accessed 23 September 2020].

LeFevre, K., DeWitt, D.J. and Ramakrishnan, R. (2005) 'Incognito: efficient full-domain k-anonymity', in F. Ozcan (ed) *Proceedings of the ACM SIGMOD International Conference on Management of Data, Baltimore, June 2005*, New York: Association for Computing Machinery, pp 49–60.

Li, N., Li, T. and Venkatasubramanian, S. (2007) 't-closeness: privacy beyond k-anonymity and l-diversity', in R. Chirkova, A. Dogac, M.T. Özsu, and T.K. Sellis (eds) *Proceedings of the 23rd International Conference on Data Engineering, ICDE 2007, 15–20 April, 2007, Istanbul*, New York: IEEE, pp 106–15.

Machanavajjhala, A., Gehrke, J., Kifer, D. and Venkitasubramaniam, M. (2006) 'l-diversity: privacy beyond k-anonymity', in L. Liu, A. Reuter, K.-Y, Whang and J. Zhang (eds) *Proceedings of the 22nd International Conference on Data Engineering, ICDE 2006, 3–8 April 2006, Atlanta, GA*, New York: IEEE, p 24.

Meyerson, A. and Williams, R. (2004) 'On the complexity of optimal k-anonymity', in C. Beeri (ed) *PODS '04: Proceedings of the Twenty-Third ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, New York: Association for Computing Machinery, pp 223–8.

Monreale, A. (2011) 'Privacy by design in data mining', PhD thesis, University of Pisa.

Monreale, A., Rinzivillo, S., Pratesi, F., Giannotti, F. and Pedreschi, D. (2014) 'Privacy-by-design in big data analytics and social mining', *EPJ Data Science* 3: art 10, available from: https://doi.org/10.1140/epjds/s13688-014-0010-4 [accessed 3 February 2022].

Narayanan, A. and Shmatikov, V. (2008) 'Robust de-anonymization of large sparse datasets', in *2008 IEEE Symposium on Security and Privacy (SP 2008)*, Washington, DC: IEEE, pp 111–25.

OECD (Organisation for Economic Co-operation and Development) (2013) *The OECD Privacy Framework*, available from: https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf [accessed 3 February 2022].

Pellungrini, R. (2020) ‘Modeling and Predicting Privacy Risk in Personal Data’, PhD thesis, University of Pisa.

Pratesi, F. (2017) ‘Privacy Risk Assessment in Big Data Analytics and User-Centric Data Ecosystems’, PhD thesis, University of Pisa.

Quintarelli, S. (2020) ‘Algorithmic accountability’ [Affidabilità e responsabilità degli algoritmi] Fondazione Ugo Bordoni webinar [in Italian], available from: https://youtu.be/jLbSAXdwLrs [accessed 23 September 2020].

Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) ‘“Why should I trust you?”: explaining the predictions of any classifier’, in B. Krishnapuram (ed) *KDD ’16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining San Francisco August 13–17, 2016*, New York: Association for Computing Machinery, pp 1135–44.

Rizvi, S.J. and Haritsa, J.R. (2002) ‘Maintaining data privacy in association rule mining’, in *Proceedings of the 28th International Conference on Very Large Data Bases Hong Kong 20–23 August 2002*, np [US]: VLDB Endowment, pp 682–93.

Rodríguez-González, A., Vakali, A., Mayer, M.A., Okumura, T., Menasalvas-Ruiz, E. and Spiliopoulou, M. (2019) ‘Introduction to the special issue on social data analytics in medicine and healthcare’, *International Journal of Data Science and Analytics*, 8(4): 325–6.

Sanil, A.P., Karr, A.F., Lin, X. and Reiter, J.P. (2004) ‘Privacy preserving regression modelling via distributed computation’, in W. Kim (ed) *KDD ’04: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York: Association for Computing Machinery, pp 677–82.

Sea Data Net (2006) Pan-European Infrastructure for Ocean and Marine Data Management, available from: https://www.seadatanet.org/ [accessed 3 February 2022].

SKA Organisation (2011) Square Kilometre Array: Exploring the Universe with the World’s Largest Radio Telescope, available from: https://www.skatelescope.org/science/ [accessed 3 February 2022].

SoBigData (2015) Social Mining and Big Data Ecosystem: A Research Infrastructure (RI), available from: http://project.sobigdata.eu/ [accessed 3 February 2022].

Sweeney, L. (2000) *Uniqueness of Simple Demographics in the U.S. Population*, Technical Report LIDAP-WP4, School of Computer Science, Data Privacy Laboratory, Carnegie Mellon University, Pittsburgh, PA.

Sweeney, L. (2002) ‘k-anonymity: a model for protecting privacy’, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5): 557–70.

Uria-Recio, P. (2018) '5 principles for big data ethics', Towards Data Science, 14 September, available from: https://web.archive.org/web/20190324111737/https://towardsdatascience.com/5-principles-for-big-data-ethics-b5df1d105cd3 [accessed 3 February 2022].

von der Leyen, U. (2019) 'A union that strives for more: my agenda for Europe: political guidelines for the next European Commission 2019–2024', available from: https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf [accessed 3 February 2022].

World Economic Forum (2016) 'The impact of digital content: opportunities and risks of creating and sharing information online', Geneva: WEF, available from: www3.weforum.org/docs/GAC16/Social_Media_Impact_Digital.pdf [accessed 3 February 2022].

Yao, A.C. (1982) 'Protocols for secure computations', in *23rd Annual Symposium on Foundations of Computer Science (SFCS 1982)*, New York: IEEE, pp 160–4.

Zhan, J.Z., Matwin, S. and Chang, L.W. (2005) 'Privacy-preserving collaborative association rule mining', in S. Jajodia and D. Wijesekera (eds) *Data and Applications Security XIX: Proceedings of the 19th Annual IFIP WG 11.3 Working Conference on Data and Applications Security, Storrs, CT, USA, August 7–10, 2005*, Berlin: Springer, pp 153–65.

Zhang, P., Tong, Y., Tang, S. and Yang, D. (2005) 'Privacy preserving naive Bayes classification', in X. Li, S. Wang and Z.Y. Dong (eds) *Advanced Data Mining and Applications: Proceedings of the First International Conference, ADMA 2005, Wuhan, China, July 22–24, 2005*, Berlin: Springer, pp 744–52.

Zook, M., Barocas, S., boyd, d., Crawford, K., Keller, E., Gangadharan, S.P., Goodman, A., Hollander, R., Koenig, B.A., Metcalf, J., Narayanan, A., Nelson, A. and Pasquale, F. (2017) 'Ten simple rules for responsible big data research', *PLoS Computational Biology*, 13(3): art e1005399, available from: https://doi.org/10.1371/journal.pcbi.1005399 [accessed 3 February 2022].

10

The AI industry and regulation: time for implementation?

Gabi Lombardo

Introduction

The AI industry is booming. Despite a deep economic crisis caused by the COVID-19 pandemic, digital technologies have seen remarkable success. When lockdowns/confinements began, more of us had to become digitally 'enabled' citizens – and even more so for those businesses forced to send employees to work from home. This even involved people historically less comfortable with sharing data and engaging with digital content being pushed forward into the world of new technologies, sharing their personal information, registering on all types of social media and websites, using video conferencing tools and other connected platforms.

More importantly, society has a growing interest in the many opportunities that AI and online data can offer, which we have only begun to tap into – for example: improving content moderation on social media; supporting clinical diagnosis in healthcare; and detecting fraud in financial services. Even those sectors that are relatively mature in their adoption of digital technology, such as financial services or retail, have yet to maximise the benefits of AI and data analytics. Particularly relevant also is the role played by the EU recovery fund which foresees a 20 per cent investment in digitalisation and other instruments.[1]

This accelerated digital adoption is not without its problems. Three barriers to progress require particular attention: the poor quality and quantity of data; a lack of coordinated policy and practice across public and private sectors; and transparency around AI and digital data use (CDEI, 2020: 4). The AI and data industry developed so fast and so globally that there is no accepted regulatory environment for good practice, beyond data protection laws such as the EU's General Data Protection Regulation (GDPR), which are only a partial solution. The big data giants – Facebook, Google, Amazon and so on – have been left to set the tone for the whole industry and strongly advocate for self-regulation. Given the pressure and the capital mobilised by these giants, it is not a surprise that so far there has been little progress towards national guidelines and regulations for the industry.

The growth of AI and data is characterised by an interesting dilemma. Governments want better self-regulation from industry, but industry says it needs governments to explain the details of what to regulate. In many cases, governments are constrained by two sometimes opposing ideological positions. First, to promote a free market where the giants of the data industry promise self-regulation because they are in the best position to recommend fair regulatory frameworks that respect privacy and human rights. Second, that governments should work together to develop the expertise to design and encourage implementation of national or supernational rules. Despite the rhetoric of political leaders, no government wants to forge its own path, knowing that the big tech companies can simply shift their business elsewhere – as demonstrated by Facebook's and Google's threats to withdraw services from Australia in response to attempts to regulate the social media space there (Scroxton, 2021).

Indeed, lawmakers and regulators have still not even arrived at a broad consensus on what 'AI' itself is, a clear prerequisite for developing a common standard to enable its governance. Some definitions, for example, are tailored so narrowly that they only apply to sophisticated uses of machine learning, which are relatively new to the commercial world; other definitions (such as the one as in the recent EU proposal) appear to cover nearly all software systems involved in decision-making, which would apply to systems that have been in place for decades. Diverging definitions of AI are simply one among many signs that we are still in the early stages of global efforts to regulate AI.

From an ethical perspective, regulation is desperately needed to protect individuals, groups and communities. Smaller companies are slowly starting to change their ways when it comes to developing algorithms with social impact, with many beginning to view consumer trust as a competitive advantage. Meanwhile, the biggest players see their algorithms as the essence of their competitive edge, and aggressively protect their intellectual property, guaranteed to reduce transparency.

Recent research suggests the public is open to greater use of data: 72 per cent of survey respondents supported the use of data-driven technology during the pandemic, but they expressed ongoing concerns over its governance (CDEI, 2021).

This chapter considers the debate around regulation and governance of the AI and data industry. The aim is to offer a snapshot of the current situation, presenting the different positions that dominate the discussion at national and international levels. Possible pathways for future developments will be suggested.

The chapter arises out of the work of a Coordination and Support action funded by the European Commission (Grant No 788352), the PRO-RES project[2] which aims to promote ethics and integrity in non-medical research. PRO-RES has designed a guidance framework regarding the delivery of responsible research and innovation. The partners of the project have strongly advised that promoting ethics in research does not mean producing rigid, prescriptive sets of rules, but rather the project aims to provide a clear backdrop on value and principles as well as a toolbox and a library of reference, to inform anyone engaging with scientific literature and evidence-based studies. The outcome of the project then provides a framework that adapts to changes over time, similar to what the European Commission (EC) proposal is addressing in terms of the regulations for AI.

The case of the EU proposal: transparency, ethics and responsibility

On 21 April 2021, the European Commission introduced a proposal for legislation to govern the use of AI, acting on its aim to draw up rules for the technology sector over the next five years and on its legacy as the world's leading regulator of digital privacy. At the heart of the issue is the will to balance the need for rules with the desire to boost innovation, allowing the old continent to assert its digital sovereignty. On where the needle should be, opinions are divided – and the publication of the Commission's draft proposal will not be the end of the discussion. But how will such rules fit in with broader plans to build European tech platforms that will compete globally with other regions? How will new requirements on algorithmic transparency come across to the general public? And what kind of implementation effort will this require from start-ups, mid-size companies and big tech? Another set of questions concerns the role of a single European platform for Member States' national data. From health registries to education material, the COVID-19 crisis has accelerated the demand for storing large amounts of data for national epidemiological purposes. Countries in Europe, even the most advanced, are ill-prepared to respond to such challenges and distrust the ability of European institutions to host and accommodate their needs, particularly because a key European disadvantage lies in the lack of significant European digital corporations with global influence (Shapiro, 2020).

Prior to the proposal's release, on 19 February 2020, the EC published a White Paper on AI – 'A European Approach to Excellence and Trust' (European Commission, 2020). The White Paper sets out policy options on how to achieve the twin objectives of promoting the uptake of AI and of addressing the risks associated with certain uses of such technology. This proposal published in April 2021 aims to implement the second objective for the development of an ecosystem of trust by proposing a legal framework that helps to ensure AI will be trustworthy. Furthermore, the EC proposal delivers on the political commitment by President von der Leyen, who announced in her political guidelines for the 2019–24 Commission, 'A

Union that Strives for More' (von der Leyen, 2019), that the EC would put forward legislation for a coordinated European approach to the human and ethical implications of AI. The core of the EU AI recommendations can be split into three parts: AI systems should be lawful, ethical and robust. Lawful AI applications are those that respect common standards. Ethical AI applications should respect agreed rules which are based on guiding principles including: a human-centric and human-made AI; safety, transparency and accountability; safeguards against bias and discrimination; right to redress; social and environmental responsibility; and respect for privacy and data protection.

High-risk AI technologies, such as those with self-learning capacities, should be designed to allow for human oversight at any time. If a functionality is used that would result in a serious breach of ethical principles and could be dangerous, the self-learning capacities should be disabled and full human control should be restored.[3]

Robust AI applications take both a technical and a social environment perspective into consideration with regards to system behaviour. More importantly, a trustworthy environment for European companies means a stronger position for the EU market and a position in which EU institutions can overcome Member State scepticism and Europe could become a beacon for trusted technology.

To implement these three core parts, the EU Trustworthy AI recommendations list seven requirements for an AI system. In other words, the proposal calls for an AI industry which is based on the approach 'ethical by design'.[4] The proposal's requirements apply to all those who are involved in planning, developing and managing AI systems. This long list includes developers, data scientists, project managers, line-of-business owners and even the users of the applications. The core requirements are:

- *Focus on human agency and oversight*: AI systems need to support human objectives, enable humans to flourish, support human agency and fundamental rights and support overall goals of a healthy human society.
- *Technical robustness and safety*: AI systems should 'do no harm' and even predict and prevent harm from occurring. They must be developed to perform reliably, have safe failover mechanisms – in other words a backup operational mode that automatically switches to a standby database, server or network if the primary system fails – that minimise intentional as well as unintentional harm and prevent damage to people or systems.
- *Privacy and data governance*: AI systems should maintain people's data privacy as well as the privacy of the models and supporting systems.
- *Transparency*: AI systems' developers and owners should be able to explain their decision-making as well as provide visibility into all elements of the system.

- *Diversity, non-discrimination and fairness*: as part of the focus on human agency and rights, AI systems must support society's goals of inclusion and diversity, minimise aspects of bias and treat humans with equity.
- *Societal and environmental well-being*: in general, AI applications should not cause societal or environmental unrest, make people feel like they are losing control of their lives or jobs, or work to destabilise the world in one manner or another.
- *Accountability*: at the end of the day, some human needs to be in charge. The systems might be working in an autonomous fashion, but humans should be the supervisors of the machine. There needs to be an established path for responsibility and accountability for the behaviour and operation of the AI system through the system's lifecycle.

Therefore, the April 2021 proposal sets harmonised principles for the development, placement on the market and use of AI systems in the EU following a proportionate risk-based approach. Such an approach proposes a common definition of AI, that will endure into the future. Certain particularly harmful AI practices are prohibited as contravening EU values, while specific restrictions and safeguards are proposed in relation to certain uses of remote biometric identification systems for the purpose of law enforcement. The proposal lays down a solid risk methodology to define 'high-risk' AI systems that pose significant risks to the health and safety or fundamental rights of persons.

A closer analysis of the EC proposal

The EU has identified two types of system that require regulation: those deemed to pose an unacceptable risk, and those it believes present a critical risk. AI systems considered a clear threat to the safety, livelihoods and rights of people will be banned. These include AI systems or applications that manipulate human behaviour to circumvent users' free will and systems that enable 'social scoring' by governments.

The EU proposal lists eight applications of AI deemed to be of high risk. Broadly speaking, these cover critical infrastructure, systems for managing crime and the judicial process, and any system whose decision-making may have a negative impact on an EU citizen's life, health or livelihood. The remit of these systems covers areas such as AI used to deny access to education or training, worker management, credit scoring and where AI is used to prioritise access to private and public services and border control.

More importantly, at the core of the proposal remain the key principles of transparency, ethics and responsible AI. Starting with transparency, the proposal states that humans need to have visibility into how the AI comes to its decisions as well as what data it uses. Without visibility, it is impossible to understand and dissect the reasons behind AI decisions if something goes wrong. Transparency gives people the opportunity to improve their systems by having visibility into how they fail or where they make mistakes. Transparency is more than just an additional feature; it is necessary for overall system accountability.

Transparency is not sufficient to address the issue of ethics. Even if we know how the system is working, it is important to know that the actions consequent on the application of AI are ethical. Companies are making use of algorithmic decision-making that has been shown to be prone to bias. For example, bias has been identified in the use of AI for recruitment purposes (Köchling and Wehner, 2020) or in the justice system (Noriega, 2020; Zajko, 2021). These biases can then become entrenched and magnified in the systems – as they often are in humans, if oversight is insufficient. Applications such as facial recognition have run into challenges with accuracy and the tendency for organisations to put too much emphasis on what is a probabilistic match. The question here is not only about the system's functionality or transparency, but rather the context in which the AI is being used.

Related to the issue of ethics is the concept of accountability in AI. Even if the systems are transparent and they are operated ethically, it is important for organisations to ensure that any outcomes are responsibly handled. If these systems hold important decisions in the balance, then monitoring by employees is key. While these systems might be ethical on face value, they need aspects of responsibility to make them trustworthy.

This latest proposal complements existing European Union law on non-discrimination with specific requirements that aim to minimise the risk of algorithmic discrimination, in particular in relation to the design and the quality of data sets used for the development of AI systems combined with obligations for testing, risk management, documentation and human oversight throughout the AI systems' lifecycle.

From the AI industry perspective, the recommendations translate into some key features:

- Maintain data privacy and security. Look across the AI system lifecycle and make sure that elements that interact with data, metadata and models are secured and maintain data privacy as required.
- Reduce the bias of data sets to train AI models. Examine training data sets for sources of potential bias and make sure that communities are represented in a fair and equitable way.
- Provide transparency into AI and data usage. Organisations should let AI system users know how their data is being used to train or power AI systems and provide visibility into aspects of data selection, usage and even the business model that the AI system supports. To the extent that the AI

system might be invisible to the user, responsible AI usage suggests you should let your users know they are interacting with an AI-based system.
- Keep the human in the loop. Even when AI systems are operating in an autonomous fashion, there should always be a human monitoring the system performance. There should be an appointed human system owner or group of humans who are responsible. Users should also know who to reach out to when the AI systems are exhibiting problematic behaviours.
- Limit the impact of AI systems on critical decision-making. If the AI system is being used for critical life-or-death or high-impact decisions, there should always be an identified failover process or human oversight to make sure that no harm is done.

Applying these key points gives users more confidence in the AI system and allows the AI to deliver the expected value without any fear of irresponsible behaviour or outcomes.

"Companies should remember that regulation also helps them because it creates a level playing field, where you know that your competitor is bound by the same rules as you", comments Catelijne Muller, a Dutch lawyer and a member of the EU High Level Expert Group on AI (Sapra, 2021).[5]

The counterfactual argument

The approach in the UK since its post-Brexit deal is still a little unclear, but it seems to be fast moving towards a very different position from its neighbours in the EU. The main aim of the current UK government is to demonstrate the ability to set a regulatory environment free from the influence of EU in many economic areas. As a result, their position has started showing a clear divergence from the EU proposal and to diverge from many AI regulations.

Set up by Prime Minister Boris Johnson, the Taskforce on Innovation, Growth and Regulatory Reform (TIGRR), consisting of three pro-Brexit Conservative MPs and former ministers, and has called for key protections to be cut from the UK's implementation of GDPR as it relates to automated decision-making (Duncan Smith et al, 2021). TIGRR recommends scrapping Article 22 of GDPR, which concerns 'the right not to be subject to a decision based solely on automated processing, including profiling'. Article 22 had been seen as establishing a 'right to explanation' to data subjects who have had decisions made about them in an automated fashion (Iphofen and Kritikos, 2021). Acknowledging the potential for controversy, the taskforce report says: 'If removing Article 22 altogether is deemed too radical, GDPR should at a minimum be reformed to permit automated decision-making and remove human review of algorithmic decisions' (Duncan Smith et al, 2021: 53). Clearly this is a point which brings the UK into opposition to the EU proposal which, as mentioned before, seeks the opportunity to have human supervision on critical life-or-death or high-impact automated decisions. The task force makes some clear claims. The report authors believe that loosening the burden of regulation is necessary to promote innovation to the benefit of the UK AI sector.

Welcoming the recommendations of the taskforce, Johnson wrote that it is 'obvious that the UK's innovators and entrepreneurs can lead the world in the economy of the future … this can only happen if we clear a path through the thicket of burdensome and restrictive regulation' (Skelton, 2021). Such a move would be controversial within the UK too. Some MPs have already raised concerns that moving away from AI regulations like those in the EU proposal could in fact impact negatively on the capacity of UK AI products to be welcome in the European market and as such retain its competitive edge across the continent of Europe. Furthermore, as *Computer Weekly* reported, trade unions have already objected to the proposal to ditch Article 22. "Scrapping Article 22 could be the green-light to the expansion of automated processing, profiling and transfer of personal data into private hands. We need data laws fit for the challenges of the digital economy, not a race to the bottom on standards", said Andrew Pakes, director of communications and research at Prospect Union (Skelton, 2021).

Changing UK AI regulations could also bring into question the EU's recent agreement to offer data adequacy to the UK, especially considering the EU's inclusion of a review process over wider fears that the UK may dilute the protections inherent in GDPR. The question is still open, as are many other aspects of this difficult post-Brexit deal.

Some experts have examined the position of the US in this area. In a report published in January 2021, US expert Alex Engler claimed that: 'This year is poised to be a highly impactful period for the governance of artificial intelligence (AI)' (Engler, 2021). In his analysis, President Joe Biden is already capitalising on the increased investment that the Trump administration approved for hundreds of millions of dollars in AI research funding, and additionally his federal agencies are already working to comply with executive guidance on how to use and regulate AI. A new National AI Initiative Office has been set up and will coordinate all AI initiatives in synergy with Congress.

Two recent publications of the US-based Center for Strategic and International Studies, one on the EU Digital Services Act and Digital Markets Act (Broadbent, 2020) and another on artificial intelligence (Broadbent, 2021), suggest that the administration centres transatlantic discussions on the high-tech regulatory matters crucial to the competitiveness of US tech companies in Europe and also relevant to strategic competition with China. It is in fact in the interest of both the US and Europe to hold the line against China, as was often repeated in the 2021 AI Summit organised by Politico. China seeks to export its intrusive model of data governance and AI regulation – a model anchored in state control of all information and communication, draconian surveillance, data localisation, and other protectionist and autocratic practices. To succeed, Europe and the US should agree on a basic framework of top-line, democratic, regulatory principles for AI that can be promoted with trading partners in Asia-Pacific, where China is proselytising its model as an element of the Belt and Road Initiative.[6]

In truth, the US is balancing multiple priorities. It is committed to ensuring the technology can be built on democratic ideals. Congress has already proposed legislation like the Algorithmic Accountability Act, which is close to the EU regulations. Also, there are initiatives like the Joint Artificial Intelligence Center and National Security Commission on Artificial Intelligence that have stated that AI is needed mainly in the interest of security, an arguable view from the perspective of European human rights and lawyers. At the same time, US institutions see AI as an economic catalyst and, together with other governments, believe that a balance between ethical, fair and unbiased AI must not stifle innovation. For some companies in the US, the European path suggests a framework that is still too prescriptive and some of the requirements may hinder small businesses and start-ups. In fact, while Europe is moving quickly to craft concrete proposals for the EU-wide regulation of data, digital services and AI, the US has followed a slower and more fragmented approach where the only laws putting guardrails on AI are at the state level.[7]

At the federal level, the Federal Trade Commission (FTC) last year issued guidance emphasising the transparent, explainable, and fair use of AI tools (Smith, 2020). The FTC issued further guidance in April 2021 warning companies against biased, discriminatory, deceptive or unfair practices in AI algorithms. The National Security Commission for Artificial Intelligence's March 2021 Final Report urged the adoption of a cohesive and comprehensive federal AI strategy (NSCAI, 2021).

The US position remains quite distant from the EU proposal and addresses directly some of the key concerns that the EU proposal is meant to resolve. The US is not the only country where the debate about regulating AI is viewed in competition with business development and where ethical concerns seems to be second in order of importance compared to empowering technological companies.

Citizens and AI

The concept of trustworthiness in AI is all about humans putting their confidence in machine-based systems. Trust is hard won and it is vitally important for those looking to put AI into real-world use that they pay close attention to these issues of trustworthiness and responsibility. As AI becomes an ever-increasing part of our daily lives, trustworthiness will make the difference between AI systems that are relied upon and those that are avoided due to legitimate concerns or individual fears.

When the political debate touches upon 'ethical barriers', it is widely acknowledged that we need to develop practical guidelines because AI ethics presents major issues for society. More importantly, people need to understand where they must take control of their data, and where data is needed. At the same time, regulators are developing simple frameworks and audit arrangements that can be easily applied and explained to people. New careers are likely to emerge like actuaries, accountants and lawyers who help companies audit algorithms for bias, fairness, accountability and ethics. The bigger risk perceived today is that systems now use deep learning, meaning the capacity of a machine to extract progressively higher-level features from raw data, whereas in the past, AI systems tended to be modelled on human decision-making. Because of deep learning, today's systems are far more opaque and less controlled. The availability of data to improve AI algorithms impacts on the use of data for the public good. In fact, public data is often held by a few very large private US companies, not the public sector. Apple, Amazon, Facebook and Google are in a much better position than public sector organisations to advance because they have monopolistic access to public data. To create a more even playing field, such monopolistic control of platforms would have to be dismantled.

At the same time, concerns emerge about the centralised control of data by governments, as discussed earlier with the Chinese model, and the debate focuses on who should be entrusted to hold, manage and profit from the data. Societies are facing big challenges to protect their privacy, interests and individuals' identities, profiles and independence. The European proposal on regulations is aiming mainly to make a responsible industry which will act ethically and guarantee customers' interests.

AI applies to regulated industries: do more rules mean more costs for the industry?

Regulations will foster some major changes in the high-tech industry. Some would say that such a regulatory framework could entail a breakup of America's largest tech firms, by prohibiting them from operating and competing on digital platforms at the same time, and that this shift may imply that tremendous costs would be imposed upon consumers and companies alike.

It is fair to say that for several years now, there has been a growing pushback against the perceived 'unfairness' of the tech industry. The main arguments address the unfairness of large tech platforms favouring their own products at the expense of entrepreneurs who use their platforms; incumbents acquiring start-ups to squash competition; and tech companies that spy on their users and use their data to sell them things they don't need (House Committee on the Judiciary, 2019). On the other hand, critics say there is a chance that the reforms proposed by the House Judiciary Committee Antitrust Report, for example, would merely exacerbate the status quo (Miller and Mitchell, 2021). More importantly, it must be recognised that over the last decade, the tech sector has been the crown jewel of the US's economy and has been a factor pushing technology development across the world. While firms like Amazon, Google, Facebook and Apple have grown at an amazing pace, countless others companies have flourished in their wake.

Google's and Apple's app stores have given rise to a booming mobile software industry. Platforms like YouTube and Instagram have created new venues for advertisers and ushered in a new generation of entrepreneurs including influencers, podcasters and marketing experts. Social media platforms like Facebook and Twitter have disintermediated the production of news media, allowing ever more people to share their ideas with the rest of the world (mostly for better, and sometimes for worse). Amazon has opened up new markets for thousands of retailers, some of which are now going public.

The recurrent question is whether it is possible to regulate this thriving industry without stifling its unparalleled dynamism. Acquisition by a 'big tech' firm is one way for start-ups to rapidly scale and reach a wider audience, while allowing early investors to make a quick exit. Self-preferencing can enable platforms to tailor-make their services to the needs and desires of users. In the online retail space, copying rival products via house brands provides consumers with competitively priced goods and helps new distributors enter the market.

Sceptics may think that all these practices would be heavily scrutinised or banned outright by new regulations. Beyond its direct impact on the quality of online goods and services, this huge shift would threaten the climate of permissionless innovation that has arguably been key to Silicon Valley's success. Nothing in the EU proposal seems to really point at the quality of online goods and services but rather focuses on the risks of the AI systems. The distinction between high- and low-risk AI shows that there are no premises for a tight market framework. Some critics say: 'It leaves Big Tech virtually unscathed. It lacks a focus on those affected by AI systems, apparently missing any general requirement to inform people who are subjected to algorithmic assessments. Little attention is paid to algorithmic fairness in the text of the regulation as opposed to its accompanying recitals' (MacCarthy and Propp, 2021).

AI has applications in many products, some of which already fall under existing regulations (for example, antitrust or transport regulations). Products like cars and aircraft are already subject to regulation designed to protect the public from harm and ensure fairness in economic competition. In general, the approach to regulation of AI-enabled products to protect public safety will be informed by assessment of the aspects of risk that the addition of AI may increase or reduce. The EC proposal goes in this direction and does not propose new regulations but rather points at the existing one as models of efficiency and best practices.

At the same time, as companies begin refining their practices to abide by new regulations, some EC panel experts[8] suggested that they must also consider tools besides technology itself to eliminate bias. 'Diverse teams help to represent a wider variation of experiences to minimize bias. Embrace team members of different ages, ethnicities, genders, educational disciplines, and cultural perspectives', says Francesca Rossi, IBM's AI ethics global leader (IBM, 2019). Indeed, such training and technology that will be 'ethical by design', including all the checks and additional requirements for implementing trustworthy technology, may have some heavy costs. Experts claim that larger companies will of course be advantaged even in adopting the new regulatory system, whereas smaller businesses may lack the infrastructure, tools and resources to comply. According to Coadec (the UK-based Coalition for a Digital Economy) which campaigns on behalf of the start-up sector: 'Regulation and taxes designed to target the biggest companies will have unintended consequences, damage start-ups and lead to more unequal outcomes' (Allied for Startups, 2018: 2).

The ethical perspective: the EU proposals as a beacon to set standards for other countries.

Strengthened by the GDPR experience, the EC and the European Parliament are encouraging Member States to support their proposal. It would be the first step towards a new way to engage with AI companies and products which is in line with European values and principles. As has been mentioned, in this case regulations are not designed for the sake of limiting high-tech companies' capacities and potential, but rather to ensure that, especially in deep-learning AI, we can maintain a very ethical and transparent approach. At its core, the EC proposal remains faithful to protecting human rights and European citizens. As such, the policy strives to maintain ethical values. It has also set the tone for an ecosystem for innovation which is meant to be sustainable and trustworthy.

Economically and politically, though, it raises the issue of how it fits into the global landscape and what impact it may have on other markets and non-European companies. More importantly, with AI (and digital technology in general) being the battleground of the future economy, the question is how policies and regulations may impact on European progress in this area. Too often China is mentioned as the competitive model which will not enforce ethical standards or allow them to determine policy directions. With a poor record in human rights, China's policymakers will be less concerned about citizens' protection.

The US and Europe are forging closer alliances and they represent the most substantial share of AI global market and technology development. Political and economic tensions will affect this sector as any other, but regulations could create a homogenous playing field that would also support technology adoption and make trustworthy technology more successful for citizens' adoption. At the same time, risk assessments and costs analysis have been undertaken to assess the impact of such a proposal on EU companies and to ensure that what some entrepreneurs see as a risky strategy could eventually have a comparative advantage on the global market. Too many questions are still open and it is hard to make a real assessment as yet.

Conclusion

The proposal is published and will go through a long process of scrutiny by policymakers in the Member States, and in the European Parliament who will need to adopt it as the European approach for AI in the ordinary procedure. Only once these are adopted, will the regulations be directly applicable across the EU. At the same time, Member States are also working on a coordination plan for the implementation of the actions stated in the proposal. No doubt the road to implementation will be long and will require further clarification of the details as the industry continues to grow and prosper.

Notes

[1] In particular, they include:

- The Recovery and Resilience Facility: 20 per cent of its funds must be spent on the digital transition of Member States, including on digital skills.
- The Digital Europe Programme: promoting digital skills is a core element of this new funding, which has a budget of around €200 million for 2021 and 2022.
- The European Social Fund Plus: a fund to support Member States in reforming national education and training systems to support key skills.
- The European Global Adjustment Fund: supports training in digital skills to help laid-off workers find another job or set up their own business.
- Horizon Europe: finances grants for masters, PhD and postgraduate research activities in all fields including digital through Marie Skłodowska-Curie actions as well as the European Institute of Innovation & Technology.

[2] https://prores-project.eu, a European Commission–funded project aiming to PROmote ethics and integrity in non-medical RESearch.

[3] Guidelines were already set up by the legislative initiative of the European Parliament issued in February 2020. See: https://www.europarl.europa.eu/news/en/press-room/20201016IPR89544/parliament-leads-the-way-on-first-set-of-eu-rules-for-artificial-intelligence and https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/646174/EPRS_BRI(2020)646174_EN.pdf

[4] See SIENNA project: https://www.sienna-project.eu/

[5] Catelijne Muller is a member of the EU High level Expert Group on AI. See: https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

[6] The Belt and Road Initiative is a global infrastructure development strategy adopted by the Chinese government in 2013 to invest in nearly 70 countries and international organisations. See Belt and Road Initiative research reports from the World Bank, https://www.worldbank.org/en/topic/regional-integration/brief/belt-and-road-initiative

[7] Legislation related to AI 16 April 2021.

[8] High Level Expert Group on AI. See: https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

References

Allied for Startups (2018) 'The Impact of Regulation on the Tech Sector: Informing a regulatory environment which leads to a stronger tech ecosystem in Europe – Full Report', December, available from: http://coadec.com/wp-content/uploads/2018/12/The-Impact-of-Regulation-on-the-Tech-Sector.pdf [accessed 4 February 2022].

Broadbent, M. (2020) 'The Digital Services Act, the Digital Markets Act, and the New Competition Tool: European initiatives to hobble U.S. tech companies', Center for Strategic and International Studies, 10 November, available from: https://www.csis.org/analysis/digital-services-act-digital-markets-act-and-new-competition-tool [accessed 4 February 2022].

Broadbent, M. (2021) 'What's ahead for a cooperative regulatory agenda on artificial intelligence?', Center for Strategic and International Studies, 17 March, available from: https://www.csis.org/analysis/whats-ahead-cooperative-regulatory-agenda-artificial-intelligence [accessed 4 February 2022].

CDEI (Centre for Data Ethics and Innovation) (2020) *AI Barometer Report*, June, available from: https://www.gov.uk/government/publications/cdei-ai-barometer [accessed 4 February 2022].

CDEI (Centre for Data Ethics and Innovation) (2021) *AI Barometer Report*, December, available from: https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation [accessed 4 February 2022].

Duncan Smith, I., Villiers, T. and Freeman, G. (2021) *Final Report: Taskforce on Innovation, Growth and Regulatory Reform*, July, available from: https://www.gov.uk/government/publications/taskforce-on-innovation-growth-and-regulatory-reform-independent-report [accessed 4 February 2022].

Engler, A. (2021) 'Six developments that will define AI governance in 2021', Brookings, 21 January, available from: https://www.brookings.edu/research/6-developments-that-will-define-ai-governance-in-2021/ [accessed 4 February 2022].

European Commission (2020) On Artificial Intelligence: A European *A*pproach to *E*xcellence and *T*rust, COM(2020) 65 final, Brussels: European Commission, available from: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf [accessed 4 February 2022].

House Committee on the Judiciary (2019) Antitrust Investigation of the Rise and Use of Market Power Online and the Adequacy of Existing Antitrust Laws and Current Enforcement Levels, available from: https://judiciary.house.gov/issues/issue/?IssueID=14921 [accessed 4 February 2022].

IBM (2019) *Everyday Ethics for Artificial Intelligence*, available from: https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf [accessed 4 February 2022].

Iphofen, E. and Kritikos, M. (2021) 'Regulating artificial intelligence and robotics: ethics by design in a digital society', *Contemporary Social Science*, 16(2): 170–84.

Köchling, A. and Wehner, M.C. (2020) 'Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development', *Business Research*, 13(3): 795–848.

MacCarthy, M. and Propp, K. (2021) 'Machines learn that Brussels writes the rules: the EU's new AI regulation', Brookings, 4 May, available from: https://www.brookings.edu/blog/techtank/2021/05/04/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/ [accessed 4 February 2022].

Miller, T.C. and Mitchell, T. (2021) 'Dynamic competition in digital markets: a critical analysis of the House Judiciary Committee's antitrust report', Mercatus Center, George Mason University, 27 January, available from: https://www.mercatus.org/publications/antitrust-and-competition-policy/dynamic-competition-digital-markets-critical-analysis [accessed 4 February 2022].

NSCAI (National Security Commission on Artificial Intelligence) (2021) *Final Report*, March, available from: https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf [accessed 4 February 2022].

Noriega, M. (2020) 'The application of artificial intelligence in police interrogations: An analysis addressing the proposed effect AI has on racial and gender bias, cooperation, and false confessions', *Futures*, 117: art 102510, available from: https://doi.org/10.1016/j.futures.2019.102510 [accessed 4 February 2022].

Sapra, B. (2021) 'Why AI regulation need not hamper innovation', Wired, 5 April, https://wired.me/technology/artificial-intelligence/why-ai-regulation-need-not-hamper-innovation/ [accessed 4 February 2022].

Scroxton, A. (2021) 'Google threatens to cut off Australia', Computer Weekly, 22 January, available from: https://www.computerweekly.com/news/252495215/Google-threatens-to-cut-off-Australia [accessed 4 February 2022].

Shapiro, J. (2020) 'Introduction: Europe's digital sovereignty', in C. Hobbs (ed) Europe's Digital Sovereignty: From Rulemaker to Superpower in the Age of US–China Rivalry, Berlin: European Council on Foreign Relations, available from: https://ecfr.eu/publication/europe_digital_sovereignty_rulemaker_superpower_age_us_china_rivalry/#Introduction:_Europe's_digital_sovereignty [accessed 4 February 2022].

Skelton, S.K. (2021) 'UK taskforce calls for cutting GDPR protections', Computer Weekly, 18 June, available from: https://www.computerweekly.com/news/252502634/UK-taskforce-calls-for-cutting-GDPR-protections [accessed 4 February 2022].

Smith, A. (2020) 'Using artificial intelligence and algorithms', Bureau of Consumer Protection, 8 April, Washington, DC: Federal Trade Commission available from: https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms [accessed 4 February 2022].

von der Leyen, U. (2019) 'A union that strives for more: my agenda for Europe: political guidelines for the next European Commission 2019–2024', available from: https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf [accessed 3 February 2022].

Zajko, M. (2021) Conservative AI and social inequality: conceptualizing alternatives to bias through social theory. *AI & Society*, 36(3): 1047–56.

11

Cardiovascular disease prevention and health promotion in times of a pandemic: a global health case study

Linda Gibson, Nchafack Almighty, Sibyl Anthierens, Tebogo Maria Mothiba and Hilde Bastiaens

Introduction

This chapter provides a case study of the experiences from six sites in five countries within the Scaling-up Packages of Interventions for Cardiovascular disease prevention in selected sites in Europe and sub-Saharan Africa (SPICES), and we explore the effects of the pandemic on delivering cardiovascular disease (CVD) health promotion and prevention activities across various health systems in low-, middle- and high-income settings (Belgium: Antwerp; France: Brest; South Africa: Limpopo; Uganda: Makerere; and the UK: Nottingham and Sussex). The SPICES project is a five-year European-funded (Horizon 2020) project that started in 2018. Implementing and delivering CVD health prevention-and-promotion activity across a diverse range of health systems and infrastructures in low-, middle- and high-income settings has been complex. One of our main objectives was to identify and compare the contextual factors across study sites that influence the scale-up of a comprehensive CVD prevention intervention. The unexpected impact of COVID-19 across the implementation sites has meant that our activities had to be fundamentally altered.

With the COVID-19 pandemic, the course of the SPICES project has changed to suit the different contextual realities of the evolution of the pandemic and the restrictions set by the respective governments. Ethics is key to this discussion as it raises issues of what, if any, research can continue during emergency response efforts in the pandemic and whether research methodologies can be adapted in pre-existing projects such as ours in public health emergencies while also adhering to the ethical requirements of the funding body, in this case the EU Horizon 2020 programme.

CVD health education and health promotion cuts across the SPICES project through a series of comprehensive integrated interdisciplinary and inter-sectorial interventions. The 1986 Ottawa Charter defines health promotion as empowering people to improve their health by having greater control over it (WHO, 2019). The Charter remains as relevant today as when it was written and its three strategies (advocate; enable; and mediate) and five action areas (healthy public policy; creating supportive environments; strengthening community actions; developing personal skills; and reorienting health services) that recognise the need to promote community participation, empowerment and social justice are key drivers of approaches to health prevention-and-promotion (Potvin and Jones, 2011) as used within the SPICES project.

In Uganda and South Africa interventions are being carried out in rural and semi-urban areas while interventions in the three high-income countries focus on vulnerable groups in urban settings (Belgium, UK) and rural settings (France). The focus of the SPICES project is 'real-world' implementation, looking at how to scale up and support uptake of proven interventions to a larger group of people in different contexts. A participatory approach is taken in this project, involving relevant stakeholders from the communities (local and regional/national) throughout the whole process. The interventions focus on risk profiling for cardiovascular diseases (CVDs) in communities and GP practices (in some settings) and coaching people with increased risk on lifestyle change. Profilers and coaches (both lay people and professionals) are trained within the SPICES project on motivational conversations.

CVDs account for 31 per cent of all global deaths annually with over three quarters of these deaths occurring in low- and middle-income countries. CVDs are largely caused by tobacco use, lack of physical activity, unhealthy diets and alcohol abuse and usually take the form of heart attacks and strokes from increased blood pressure and blood glucose, and obesity (WHO, 2020b). Those living with non-communicable diseases (NCDs) can be said to be more vulnerable and at a higher risk of severe COVID-19 related illness and death (WHO, 2020a). Pre-existing CVDs are associated with severe COVID-19 outcomes and higher risks of death while COVID-19 could cause cardiovascular conditions such as myocardial injury (Nishiga et al, 2020).

The COVID-19 pandemic has also caused huge disruption in the delivery of global health care for all diseases including the late diagnosis and treatment of CVDs (WHO, 2021). As health services, including screening services, experienced severe disruption and reassignment of staff to support the strain on services, the longer-term impact on NCDs and, in this instance, CVD prevention services, will have to be assessed globally in the future.

In addition, the crisis impacted on the prevention of chronic conditions such as CVD since it seriously influenced lifestyle behaviour. The public health enforcement of COVID-19 preventative measures such as lockdowns and physical distancing have significantly reduced levels of physical activity

(WHO, 2021) and Hall et al (2021) argue that COVID-19 may further worsen the already existing *pandemic* of sedentary behaviour.

Because of the nature of the SPICES project, focusing on healthy lifestyle, CVD prevention and its implementation, and collaborative approach, the COVID-19 pandemic had a huge impact on the course of the project. It impacted on the scientific work itself but also induced ethical discussions within the consortium of doing research during a pandemic and on the role a scientific project can and needs to play in these circumstances.

In this chapter we present reflections from the different project sites. We draw from guided conversations with the SPICES site lead researchers about the impact of the pandemic on the project and the ethical issues this raised and how they were dealt with. We also explore some of the lessons learned and consider how health systems might reprioritise health promotion and prevention for CVD and other NCDs in a way that will minimise risk and vulnerability to future epidemics/pandemics.

Methodology

Drawing on qualitative research interview methodology we undertook a set of guided conversations to elicit the reflections of the SPICES teams to understand more fully the disruption that had been caused by COVID-19, the impact on site activities, how teams dealt with the disruption and what mitigation activities they undertook. We present reflections from the different project sites (Belgium, France, UK, South Africa and Uganda) on how they have dealt with and coped with the impact of pandemic on their work and what ethical issues this has raised for our community-based research in the SPICES project.

The data for these case studies were collected through guided conversations held with team members from the different sites over a period of two weeks at the beginning of 2021 when there was still a great deal of variation between the settings in terms of type of public health restrictions and their timings. A set of core questions were agreed by the authors in relation to this study. These conversations were led by senior investigators from the SPICES Nottingham, South Africa and Belgium sites through recorded Microsoft Teams video meetings and were conducted in English (the working language of the consortium). The conversations lasted between one and one and a half hours approximately.

An a priori framework for analysis was used to capture the various themes emerging from the conversations. The framework held four categories based on the topic guide (health policy context; impact of COVID-19; ethical issues raised; lessons learned/next steps) and was agreed upon by the core research team. Using a pragmatic approach to data analysis, the data were then manually analysed for themes and subthemes. All interviews were analysed by two of the chapter authors, and at each stage of this process, data were discussed within the study team.

Findings

Two key themes relating to ethics emerged from the conversations that we focus on here. The first, the impact of COVID-19 on the delivery of the SPICES project as a whole and the adaptation to the implementation in order to be able to continue to deliver the project in the different sites. The second, how the teams managed to overcome the ethical dilemmas of continuing to conduct research within the site communities during pandemic.

Theme 1: Ethical challenges for the SPICES research project due to the COVID-19 pandemic

The COVID-19 pandemic significantly affected the implementation of the SPICES project in the field and raised real ethical issues for the academic teams in terms of needing to respond quickly to a changing and fluid situation. Research that was planned or ongoing needed to be reassessed and reviewed to comply with local and country-level public health control measures. To add complexity, the timing and the level of the public health lockdown measures were not uniform across sites so different sites had to respond differently according to the local situation. At the peak of the pandemic, however, restrictions on face-to-face data collection were applied by all the universities involved. All changes to data collection required a modification request and approval from the respective sites' ethics committees.

Remodelling the interventions

Interventions were no longer possible as originally designed. Ethics permissions had to be remodelled. This raised difficulties as a large EU-funded project over five years with a clearly defined set of protocols and objectives, giving rise to tensions in being unable to deliver to our expected objectives and compromising our study design. It also triggered ethical reflections on what constitutes good use of public money in changed circumstances:

> SPICES Antwerp: 'We did get with the whole project, six million euro … I think it's our obligation to put it to good use. We can debate about what is good use. Is good use only doing what you put in the proposal and sticking really very strictly to it? I think we are not that kind of group. I think we always try to say, "OK if that is not possible, is there something else that is relevant to the CVD prevention that we can do with that money?"

> Personally, I think within the Antwerp SPICES project, the collaboration that we have built up with the local prevention organisations … to set up health collaboration kiosks … is something we didn't say from the beginning, but it fits within the philosophy of what we wanted to do and I think this is making good use of public money. Then, of course, there is the tension of the measurement and the scientific output because you have to make sure you do evaluate. And for me, that is the biggest struggle.'

Given the very different contextual settings in which each of the partner sites operate and the unclear trajectory of the pandemic it was decided that each site would adapt as it felt the need, but with a clear eye on our research objectives The UK and European partners delivered CVD prevention interventions remotely/online where possible and given the availability of the internet in high-income settings, this provided access to try to interview and carry out coaching interventions:

> SPICES Antwerp: 'I think there was a difference between the first and the second wave of COVID. Maybe not very much but there is some difference. If I remember correctly … in the first wave, at a certain point, we weren't allowed to do anything anymore because you couldn't do research and you couldn't be outside without a specific reason … In GP practices, almost everything moved online in the first wave. All the welfare organisations stopped working face to face … Then, the second wave came and the practices and the organisations didn't really close. It was really restricted with masks … also, research was considered as essential [at this point].'

> SPICES Uganda: 'With COVID-19, you needed to be very innovative to keep things moving forward for sure. And the only strategy to make sure that we are back on the ground was to see how we could integrate COVID-19 in our programme and indeed that integration, we actually started it within the School of Public Health ourselves. We were actually the first to do that. … We even had to develop a COVID-19 mitigation plan which we submitted to the review committee and they approved it …'

In the Nottingham site it was decided to use SPICES instruments (INTERHEART survey; telephone interviews) to track the impact of extended lockdowns on heart health risk:

> SPICES Nottingham: 'From all the epidemiology that was coming out and being reported, it looked like these measures would be in place

> for several months and it looked like perhaps what we ought to do was to look at the impact of those measures on heart health because we had a great baseline. So, we knew what people's heart risk score was before the pandemic started and we felt that it would be useful in public health terms to understand what the longer term impacts of sustained social isolation, social distancing and the whole lockdown package may have on their heart risk.'

At the beginning of the first major European lockdown people seemed to become more aware of their health and more open to discussing their health and knowing their CVD risk profile. However, initial gains were not sustained as the different waves took hold. In France in the setting of Brest cases were relatively low but policy came directly from the centre in Paris and was rolled out across the country:

> SPICES Brest: 'engaging with the community champions was increasingly difficult due to curfews, holiday time and also a preferred want for access to experts rather than trained lay people.'

As the restrictions in the UK began to lengthen there had been continued disengagement by SPICES communities, participant numbers falling away month by month, online and telephone engagement became extremely challenging as people got tired of the pandemic and its impacts on their lives. Activation and deactivation of participants led to research fatigue and it became difficult to sustain activities. Initial gains were hard to maintain as lockdown dragged on.

In Africa public health restrictions heavily impacted on the work of the community health workers (CHWs) who are grassroots health workers located within communities and villages. In Limpopo research activity was forced to stop and then they were unable to return to their original field site and work with the CHWs:

> SPICES Limpopo: 'When the lockdown came and we had to halt everything and last year when things were sort of getting better, we thought we will then go back but you know, government was also coming on board saying it is not a good idea for CHWs to go into the field to interact with these people because remember, these are vulnerable people in terms of COVID-19 infections. … So, a decision had been that the CHWs would be operating from the clinics and for us as the research team, we would no longer have the privilege of joining the CHWs sometimes to visit households as was embedded in our proposal.'

Impact on research design and data collection

Some of the randomised control trial (RCT) designs in the sites of Uganda, Brest and Sussex posed problems during the pandemic as nothing was able to be controlled anymore. RCTs are prone to rigid protocols and are seen as the gold standard in health and medicine. Yet during the pandemic, communities increasingly became a site of attention by researchers in terms of the impact of COVID-19; they were already overburdened with coping with economic and mental health impacts and needed to be able to define their own priorities. There was a drop in recruitment across the sites, as may be expected, and some outcome measures had to be adapted (for example waist-hip measuring was unable to be carried due to social distancing rules) and so some of the statistical data were unable to be collected, potentially threatening the ability to collect enough 'hard evidence' to prove efficacy. Therefore, it becomes more difficult to publish findings in high-impact journals.

> SPICES Uganda: 'In the midst of the survey, we couldn't continue, we stopped. … And indeed we did not continue with the survey, even after lifting the lockdown. So, that was one of the severe impacts and given our stepped wedge [RCT] nature of the project … indeed that affected our whole methodology.'

However, the different sites responded in diverse, thoughtful, and creative ways, from adapting research designs and data collection methods to fostering researcher resilience and rethinking about the researcher–researched relationship:

> SPICES Antwerp: 'The thing that comes to my mind is what we have tried with the online [interventions]. We have also discussed with our stakeholder groups and what we find is that with this vulnerable group of people, it is really difficult. So, maybe that is not the way forward for everybody and maybe we need to try and balance and also try to see how we can do things face to face outside.'

Some primary methods of data collection had to be conducted remotely where possible. The different sites had to look and shift towards alternative methods of data collection moving from face-to-face data collection to distant, virtual data collection:

> SPICES Nottingham: 'We decided to resurvey on a timetable, every two months we said initially. From June to August, then October. We were going to continue every two months to resurvey using

> INTERHEART to be able to track changes in CVD risk due the fact that maybe people couldn't exercise so much, maybe they have been more impacted by stress and depression.'

Resilience within the teams

The different site teams had to try to remain flexible and there were open discussions between the consortium about how to move forward and adapt the project where necessary, once it became clear that the COVID-19 restrictions would last for a long time. It was necessary to accept limitations and adapt the expectations of the implementation of the SPICES project in conversation with the EU funders. It was also key to look out for the different members of the team as working from home became the standard and there was the risk of isolation. Issues arose around working virtually with our participants, especially for the less-experienced members of the teams and those undertaking doctoral research, which was heavily impacted:

> SPICES Nottingham: 'We have gone to telephone interviews, you know, telephone interviews are a pretty standard part of things. But, what is clear I think, it's pretty much harder to do work over the phone than it is face to face. Which really reinforced for me the advantages of doing face-to-face work whether it is one to one or group based. And I do miss it'.

Regular virtual team meetings and support for each other through our consortium networks were important to have in place, although it was challenging at times as the trajectory of the restrictions carried on in various iterations. Meeting virtually over long periods of time has certainly presented challenges:

> SPICES Uganda: 'When we meet physically … there are some things we can demonstrate better offline. All our meetings are now virtual. Of course, we try but there are a lot of things you cannot achieve when you do things virtually. Of course learning experiences get missed. Sometimes, the fact that we can't get sufficient bandwidth … like all our Zoom meetings will never be video meetings. So, there are a lot of expressions in communication that we actually miss.'

This international global health project had been working closely together over the three years previously and so partners had built up considerable group capital and rapport from our work together and this has been key to sustaining the integrity of the project. The consortium had visited all the sites and understood the different cultural contexts of where the project was being implemented. This meant we were able to jointly prioritise the challenges faced, respond quickly and flexibly to fluid situations, and still communicate, work and plan together.

Opportunities

The challenges imposed by the pandemic pushed the teams out of their comfort zone towards the use of different methods and this in turn has been beneficial for the researchers' development and academic growth. For example, Sussex moved more towards taking a participatory action research approach:

> SPICES Sussex: 'The slight strength for us is that we are working with community organisations and whilst the whole model has changed, it feels like they are more in touch with the communities than we are. … We included some people we otherwise wouldn't have because of travel restrictions or childcare. … There have been some positives … we lost about one or two organisations but gained as many as we lost from new relationships with organisations on the basis of this virtual model.'

What became clear to the teams was the value of the qualitative data, which was rich and able to better capture the contextual and experiential nature of what was happening within the different settings. The researchers also had to learn some new skills in order to communicate and engage with people in a human and compassionate way. More emphasis was put on personalised communication instead of a generic response to have better engagement. All the sites also moved to spend more time on desk-based aspects of research such as literature reviewing and having dedicated time to write articles and papers.

In summary, the ethical issues the sites had to grapple with were how far to adhere to the original proposal (in which case, the project could come to a halt) and how far to adapt the project to the changing context(s). And this raised many questions for us such as how you isolate the impact of interventions (even if you can find a way to deliver them) from the impacts of COVID control measures.

Theme 2: Ethical dilemmas of continuing to conduct research within the SPICES site communities during the pandemic

Ethical implications were a key element in the discussion and reflections in the site teams when making decisions about how to continue with the project at the different sites during the pandemic and to seek how the implementation of the SPICES project could be used to enhance the well-being of our research participants and population.

Impact of the pandemic on our SPICES project communities

Across the sites one common concern about the impact of lockdown was on the restrictions to any research activity which meant that many of the project's intervention activities had to come to a halt. In all settings the concern was with dealing with COVID-19 and trying to ensure that the health systems did not become overwhelmed. In the SPICES project face-to-face fieldwork and data collection halted and, where possible, moved online in the European SPICES study sites, although at different times of the various waves:

> SPICES Sussex: '… [name] protocol that she published said we were going to have a stepped wedged randomised trial design and that was something Tom and I quickly abandoned as soon as the pandemic hit because we just couldn't see how we were going to force our community organisations into a trial structure. We needed a much more collaborative approach with them.'

Virtual activity was not feasible for African sites since the relevant infrastructure was lacking. For the research investigators this raised issues of how much or how little pressure to put on stakeholders for collaboration when they were dealing with COVID-19. In Europe much household, social and economic activity transferred online but for SPICES researchers in a number of the settings transferring and keeping the impetus of the interventions was difficult. Communities and stakeholders were focused mainly on coping with COVID-19:

> SPICES Brest: 'I think it would be very hard to focus again on CVD prevention because all stakeholders and all the national and local policies are against COVID. … This focus is all the strategies, all the efforts, all the financial resources and the other sicknesses that are more prevalent … nobody cares.'
>
> SPICES Sussex: 'We did lose a couple of relationships with some community groups quite early on in that first lockdown … they were furloughed or I guess their focus on what they wanted to work on changed. … I guess that led to us to making virtual relationships with a couple of our sites we have got now. They started completely online from that first lockdown. … In some ways, there have been some positives because everyone has been going through the same thing and there have been some common ground that everyone has been able to meet on.'

Asking people to participate in research during a pandemic can cause additional unnecessary stress and recruitment became difficult for certain sites. The sites needed time to pause and reflect on whether data collection could be postponed. The teams had to reflect on the ethical questions concerning the relevance of the SPICES research and the possibility of over-demanding from certain communities:

> SPICES Limpopo: 'If we were going to continue seeing participants in their homes, that is a natural environment in terms of research and the dynamics of seeing the family … would be different as having to interact with the participants in the clinics as participants would come either alone or with one family member. … That has changed the relationship with the participants.'

Therefore, in line with the participatory action research approach taken in certain sites (such as Antwerp and Sussex), it was important the teams went into discussions with the community and the organisations in order to move forward with the project and find out how they could adapt in order to ensure a mutual interest for the community, the partner organisations and the research project.

> SPICES Sussex: 'We talked a lot about it like Kathleen said but when we actually got the first hints … when two of the organisations started saying maybe we could do this online, and then it almost happened in a week for us. We all just went, OK, this is what we are doing and since then, we have kind of been on that path.'

Impact on the research communities/stakeholders

Community engagement was a key strategy within our research project to working with our vulnerable/at-risk populations. These are the very communities most affected by the pandemic and so the COVID restrictions placed a dual burden on these groups. It also meant that communities we were working with had other priorities to focus on rather than longer-term conditions such as CVD:

> SPICES Sussex: 'There was a three- or four-month hiatus and within this time we also had some staff members on furlough, so the team went back to kind of a low level of resource and we kind of communicated with our organisations … to say we are kind of taking a break at the moment but we want to stay in touch … we don't want you to feel abandoned because there had been a previous break in the project and that was problematic for some of the organisations.'

In Uganda and Limpopo, where there was limited access to target communities and partners due to restrictions, the project was able to support national COVID prevention strategies through the Uganda SPICES networks and personnel where this was allowed. They were able to move quickly to repurpose field teams and retrain volunteers to promote COVID prevention, distribute equipment, and retain CVD prevention where possible through a highly centralised primary healthcare system and good access and communication with the Ministry of Health. This meant that CHWs were able to restart activities by combining COVID-19 and CVD interventions after further training. CHWs had a great role in providing education about COVID-19 transmission and safety within communities and were very appreciated after initial hesitation from the public to engage in any prevention activities due to fear of COVID-19:

> SPICES Uganda: 'We also had to train them [CHWs] on COVID-19. They had basic training on COVID-19 so that they could provide a combined response of COVID-19 and CVD prevention. … We then went on to procure for them temperature guns and the healthcare workers appreciated this. With temperature guns, buying into the project was really safe and quick even during the pandemic. That strategy helped to get back on our feet.'

> SPICES Limpopo: 'We will no longer have the participants being met by the CHWs in their homes … that in itself has certainly changed that dynamics of our relationship with the participants and the CHWs.'

Lessons learned

Online interventions are difficult with vulnerable people where often there is a digital divide across the sites and culturally the power of face-to-face interaction is important. As the focus on COVID-19 took hold across health systems, communities and the media, it was difficult to maintain a focus on CVD. In Uganda and Limpopo where community engagement is at the core of the primary health system they were able to train the CHWs in COVID prevention measures using the existing CVD intervention programmes (although these did stop and start) and they were able to refocus programmes on both NCDs and infectious diseases. In Limpopo, the model of holistic training of CHWs supports the integrated approach and is, we would argue, a key lesson to learn for high-income settings.

> SPICES Limpopo: 'There is a bit of a positive as people have been brought a bit closer to technology. Technologies are now being embraced because that is how a whole lot of things are being

> delivered. … The pandemic is going to affect communities differently given their level of advancement and relationship with technology …
>
> Vulnerable people and those who have been struggling before maybe because they didn't have the privilege to go to school and so they can't read and write for instance. So, this wave is moving fast and now, we have to take them along … chances are high that they may be left behind. …
>
> As much as we may want to reprioritise and refocus, it means government and communities have to find each other so that the reprioritisation may be could end up being such that NCDs should receive more attention than infectious diseases … or striking a balance at a different level than the balance that has been existing.'

In Europe the sites saw an initial turn away from any NCD focus. As the pandemic continues, however, there is a recognition beginning to emerge of the importance of prevention in primary care and in community organisations:

> SPICES Antwerp: 'Prevention is the most important thing right now … especially with COVID-19, we can really call for more focus on prevention given all we now know. … Prevention is an important issue that should be taken up even if it is at a very low level. Now, in practice, we are trying to adapt the project workload, at least for the scientific part of it, the data collection they do for us, the interviews … are an extra burden we put on them but we try to adapt the project in a way that the burden is as low as possible.
>
> What I also find in communication with our local stakeholders is that it is important not to pressure them to include large numbers of patients in the project. That is the luck we have with our research design … that we do get very rich qualitative data.'

Nottingham, however, had experienced early in the project the impact of austerity policies which meant that there was very little health promotion or third-sector involvement so there was little to work with in this regard:

> SPICES Nottingham: 'We found that a lot of [health prevention-and-promotion] work was decommissioned. This also kind of reflects the national picture where health promotion has really been rolled back because of austerity. We have had ten years of austerity really where community groups were just not able to function anymore or functioning under really difficult circumstances …
>
> It is important to understand that for the UK context and the Nottingham context, a lot of health promotion … the responsibility

> of health promotion has moved away from health services as primary care … the responsibility of health promotion and public health has moved to the local authority/local government who have also been starved of cash.'

Discussion

This case study has explored the ethical issues that have been raised due to the impact of the global pandemic of COVID-19 on an international health project focused on implementing CVD prevention and health promotion activities across five countries and six sites. Two main themes emerged from the guided conversations that we had with the different consortium sites: first, the impact of COVID-19 on the delivery of a global health prevention-and-promotion project and the ethical challenges faced by the teams in adapting the design and implementation activities yet still trying to maintain integrity with the overall research design. The second theme raised ethical issues of conducting research that focuses on longer-term behavioural impacts on CVD and heart health with and in communities when during the pandemic they are coping with immediate severe public-health restrictions imposed by governments to try to keep their populations safe. These restrictions were used across all of the SPICES sites to contain the transmission of COVID-19 since March 2020.

While there have been some commonalities in the implementation of these restrictions, as one might expect there have also been some notable variations across Europe and sub-Saharan Africa and we do not endeavour to make generalisations, rather to offer our reflections. The restrictions happened in waves across the sites and so the measures, reporting and implementation of restrictions is variable. This has meant that for our global health project that research activities have been very stop and start in all the settings. Early impact of the pandemic was in the European sites while South Africa and Uganda appeared to be less affected. However, at the time of writing, our African colleagues are grappling with a 'third wave' and vaccine inequity, another key ethical issue being explored elsewhere. The trajectory of the pandemic globally has still to play out.

What lessons can be learned from some of the ethical dilemmas the SPICES project has experienced? First, being able to adapt the project to a changing context is necessary but this raises ethical issues about the disruption of data collection, measurement and scientific output that funders are expecting. One of the key ethical issues in undertaking research with vulnerable communities was about the need to be sensitive to the pressures the communities themselves are under and how the project adapts activities to still deliver interventions that are useful and still prioritise health-promotion activities. These are critical but in times of physical restrictions, social and economic impacts on mental health and job security, for example, demonstrate the need to be agile and responsive researchers. From our experience within the SPICES project one of the things we learned was that where participatory approaches are used in research, it helps to adapt the research with communities and stakeholders in a more supportive way that speaks to their needs. Flexible research designs in implementation research are valuable and important.

Second, in undertaking large-scale international global projects, the value and importance of local contextual understanding of policy and the communities in which research takes place is key. It was clear, although perhaps unsurprising, that CVD health prevention-and-promotion programmes such as the SPICES project would be unable to carry on with their planned implementation activities given the scale of the crisis, the disruption at national levels to health services and the mobilising of all human and physical resources to deal with the crisis. Some of the SPICES investigators are themselves clinicians and so needed to be deployed on the frontline of the emergency response. Much of the concern in all countries has been a fear that services may become overwhelmed and so nationally and internationally the realisation that those who were most affected were those suffering from chronic and NCDs took a while to kick in. Ill-health prevention, screening and other primary care services were severely disrupted as national public health restrictions were imposed. As we have seen in the SPICES European sites much of primary care and community organisations shut down due to COVID-19 as the public health response was largely led by a bioscientific/security perspective. Despite warnings from other epidemic/pandemic global episodes, as in SARS, H1N1 and, more recently, Ebola, there were critical failures in global and national health policy responses and preparedness, not least the issue of focusing on a technical response alone (Bowsher et al, 2020). One of the key lessons that WHO reflected on after the Ebola crisis in 2014/15 in West Africa was the importance of community engagement and the need to incorporate social science responses using a social model of health that understood the multilayered nature of those communities in responding to health emergency (WHO, 2020a). Yet it was also clear from our experience in the SPICES project that, especially in high-income settings, the response was highly centralised and, in a number of the settings, community organisations, self-help and voluntary organisations were shut down at the peak of the waves.

In sub-Saharan Africa, in Uganda specifically, the response was different. Health promotion and the primary care system is a core part of the National Health Plan in Uganda and colleagues from Makerere University, in tandem with the Ministry of Health, were, in the first wave, able to respond quickly to put a risk mitigation plan in place. CHWs and frontline staff were able to be trained in pandemic responses quickly through the existing screening and health promotion and education activities programmes that were already in place. Strong local intelligence that prioritised a coordinated primary care response meant that expertise could be shared more quickly and while certainly there was some loss, many of the prevention programmes were able to continue. Uganda's history of focusing on health promotion and ill-health prevention, and its experience of controlling previous epidemics, have also come into play.

In Limpopo we saw the impact of an integrated primary care approach to prevention at policy level, enabling CHWs to stay operational during the restrictions although the research was unable to carry on. However, at points, the system was under threat as levels of lockdown increased and CHWs were unable to go into the communities they worked in. But the system still coped to some degree as people were asked to attend the primary care services to see the CHWs for their preventative care, although again this was not completely the case.

These examples lead us to reflect on the more global ethical tensions and shifts in the locus of power and control from the individual and community, where most upstream health-promotion activity is focused, to the need of the state to implement large-scale public health restrictions necessary in the control of infectious diseases in a pandemic scenario. It also reveals where the priorities of health systems lie. While NCDs is an emerging issue in south and sub-Saharan Africa, they also have a history of dealing with outbreaks of infectious disease and epidemics and have an emphasis on building strong preventative health systems. Consequently, they have experienced the pandemic differently from Europe in the way they have navigated its impact. There is bidirectional learning to be considered in the delivery of heart health interventions across communities. The structuring of health systems in high-income settings such as Europe is led by a high-cost, high medical interventionist approach. The rolling out of this approach globally to many low- and middle-income settings means that national governments' prioritisation of donor-led heavily medicalised programmes have overshadowed other initiatives more focused on the promotion of health and well-being for citizens and communities through an integrated community-based primary care system approach.

Conclusion

This overview of a global health implementation research project on CVD prevention and health promotion during a pandemic has focused on some of the ethical tensions that arose for us as researchers. Using the methodology of guided conversations, we have been able to reflect on the impact of COVID-19 on our research activities as a group. Reflection and learning from our colleagues in the different sites were an important feature in our

SPICES project from the beginning and this has been key to enabling us to navigate through uncertain and difficult ethical dilemmas as they have arisen. We would argue that this is an important skill set to build into any research activities which policymakers should consider.

While we cannot make any generalisations from our work due to the high variation of both the nature and impact of COVID-19 in the different sites it seems to us that our experience talks to both specific and wider recommendations. CVD prevention as a focus is one approach in minimising risk and vulnerability to COVID-19 and hence we argue for a reprioritising of health promotion activity within our health systems to mitigate the (re)emergence of highly infectious disease within vulnerable populations.

At a wider policy level, investing in strong, preventative primary health care systems is essential to protect the most vulnerable of our populations, who are more likely to suffer the dual burden of disease, both communicable and non-communicable. This is not a novel observation. Many in the global health community have consistently argued for this but we would suggest the pandemic has shone a light on the ethical imperative for us as global health researchers to enable and advocate to policymakers for those people in vulnerable communities with whom we work to hasten efforts towards this goal. We urge policymakers to grasp this unique moment in history to shift health systems towards a more preventative and integrated approach.

Acknowledgement

The SPICES project receives funding from the European Commission through the Horizon 2020 Research and Innovation Action Grant Agreement No 733356. We would like to thank the SPICES country teams for their participation in this case study.

References

Bowsher, G., Bernard, R. and Sullivan, R. (2020) ‘A health intelligence framework for pandemic response: lessons from the UK Experience of COVID-19’, *Health Security*, 18(6): 435–43.

Hall, G., Laddu, D.R., Phillips, S.A., Lavie, C.J. and Arena, R. (2021) ‘A tale of two pandemics: how will COVID-19 and global trends in physical inactivity and sedentary behavior affect one another?’, *Progress in Cardiovascular Diseases*, 64: 108–10.

Nishiga, M., Wang, D.W., Han, Y., Lewis, D.B. and Wu, J.C. (2020) ‘COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives’, *Nature Reviews: Cardiology*, 17(9): 1–16.

Potvin, L. and Jones, C.M. (2011) ‘Twenty-five years after the Ottawa Charter: the critical role of health promotion for public health’, *Canadian Journal of Public Health*, 102(4): 244–8.

WHO (World Health Organization) (2019) 'Health promotion: the 1st International Conference on Health Promotion, Ottawa, 1986', World Health Organization, available from: https://www.who.int/teams/health-promotion/enhanced-wellbeing/first-global-conference [accessed 11 August 2021].

WHO (World Health Organization) (2020a) 'Ebola then and now: eight lessons from West Africa that were applied in the Democratic Republic of the Congo', World Health Organization, 10 April, available from: https://www.who.int/news-room/feature-stories/detail/ebola-then-and-now [accessed 3 August 2021].

WHO (World Health Organization) (2020b) 'Cardiovascular diseases (CVDs)', World Health Organization, 11 June, available from: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 3 April 2021].

WHO (World Health Organization) (2021) 'COVID-19 significantly impacts health services for noncommunicable diseases', *World Health Organization*, 1 June, available from: https://www.who.int/news/item/01-06-2020-covid-19-significantly-impacts-health-services-for-noncommunicable-diseases [accessed 2 April 2021].

12

Conducting ethical research in crisis situations: COVID-19

Dónal O'Mathúna

Introduction

Disasters, humanitarian emergencies and pandemics are characterised by crisis conditions that require urgent attention. Disaster risk reduction (DRR) aims at preventing new risks, reducing existing risk and managing residual risk. On a global level, DRR policies and strategies are developed by the United Nations Office of Disaster Risk Reduction (UNDRR), formerly known as UNISDR. Disasters lead to risk and losses in terms of 'lives, livelihoods and health' and impact 'economic, physical, social, cultural and environmental assets of persons, businesses, communities and countries' (UNDRR, 2020). The *UNDRR Strategic Framework 2022–2025* identifies four strategic objectives, the first of which is directly relevant to this chapter's topic: that countries use DRR and sustainable development policies based on robust evidence, quality information and good practice (UNDRR, 2021). Research plays a key role in all these areas. Some of these policies can be developed ahead of a disaster based on findings from similar events, predictions and general principles. However, those general policies will then need to be applied to the specific conditions of the current disaster. In other situations, scientific evidence will be lacking on the interventions, responses and policies that are best for a specific crisis. This leads to much uncertainty about how to respond to a crisis and makes the tasks of responders, practitioners and policymakers even more challenging.

Such was clearly the case with the COVID-19 pandemic. While the response to the pandemic could have been improved by better pre-pandemic planning and preparation, and more extensive dissemination of, training in, and research on existing plans, this chapter will instead focus on the role of research in the early days of the pandemic and the ethical challenges that such research raised. In March 2020, the World Health Organization (WHO) published a research roadmap for the pandemic that highlighted the importance of research and innovation 'during, after, and in anticipation of public health emergencies'. It linked research to 'a moral obligation to learn as much as possible, as quickly as possible'. When 'policy and practice' is based on such research, it 'can save lives and needs to be integrated into the response from the start' (WHO, 2020: 4).

This declaration of support for crisis research points to one of the key ethical challenges with research during crises which was seen clearly with COVID-19 and continues to have ramifications for policy and practice. What, if any, adjustments to normal research practice and policy should be made in order that research findings would be made available 'as quickly as possible', as WHO put it? Should research methods be adapted or replaced so that results become available sooner than otherwise? Should ethics review, peer review, safety monitoring, regulatory review or other governance structures be changed so that research can be conducted as quickly as possible? A number of such items will be discussed here to exemplify how research might be conducted rapidly while maintaining standards. Examples from COVID-19 research will be selected to point out the potential benefits with some approaches, and some of the challenges they raise. The goal is to learn from these examples to see how research might be conducted better during any future crisis.

Choosing a methodology

As the pandemic developed, it quickly became clear that existing treatments and interventions were not effective with COVID-19 patients. Neither were there interventions to prevent COVID-19 infections, such as prophylactic medications or vaccines. Research to develop *de novo* treatments would likely take too long, so the search was undertaken to determine if existing, sometimes already approved, treatments could be repurposed for use against the SARS-CoV-2 coronavirus or if experimental medications in development for other conditions might be effective here. A variety of candidates were identified for further investigation, primarily those with a track-record of effectiveness against viruses, especially other coronaviruses, or in relieving the symptoms associated with COVID-19. Early candidates included antimicrobial agents like hydroxychloroquine (HCQ) and chloroquine, ivermectin and lopinavir, approved drugs like interferon and dexamethasone, and experimental agents like remdesivir.

The next question centred on what type of methodology should be used to demonstrate whether or not these agents should be used or recommended for COVID-19. This debate revolves around the various types of studies that can be performed. Policymakers, media reporters and individual patients may not see the need to understand the technical details of these different research methods, but a basic understanding of their aims and purposes is crucial in making decisions about whether proposed medications should be recommended or used. This became particularly important during the pandemic as research findings were often first heard about through press releases and media reports, and particularly through social media. To evaluate the claims and counterclaims that the general public was being exposed to, some knowledge of research methods and trial design became an important part of health literacy. A brief overview of those methods will be provided here as the terminology is important to understand the ethical controversies described later. Further details and definitions of these research methods are widely available (for example, O'Mathúna et al, 2022).

Anecdotal evidence is that which reports people's accounts of what they did, saw or heard. People reported that certain symptoms occurred with COVID-19, that they were treated in various ways, and then recovered after so long, or unfortunately didn't. We tell such stories regularly to share our experiences. In a clinical context, such accounts may be written up and published as case studies. Details are shared so that other clinicians, patients and policymakers can learn from others and start to see possible patterns. *Case studies* provide important insight into what is happening, and can be used to generate hypotheses about where research is needed. So many case studies of COVID-19 were published that several systematic reviews were conducted to show the overall prevalence of various symptoms and how they were being treated (Borges do Nascimento et al, 2020). Most importantly, while case studies describe what happened, for a complex area like people's health, they are very poor at explaining why changes took place. For example, let's take the case of a person with COVID-19 who takes aspirin, and recovers quickly. The reason for their recovery might have been aspirin, or some other aspect of how they took care of themselves, or some combination of those factors, or none of them; they might have been fortunate to get a mild version of the infection and recover quickly. We may be encouraged by their experience, but we should not assume that therefore aspirin is a good way to treat COVID-19.

Further research is needed to address different questions and could involve a range of research methods. A *case-control study* retrospectively examines a group of individuals with a goal of identifying important aspects of the people or their care. For example, researchers might look at the medical records of patients who died from COVID-19 as well as the records of patients who recovered. They could search for differences between the two groups that might suggest why some recovered and others didn't. Interesting proposals might arise, such as more patients who survived received a specific treatment compared to those who died. These provide important hypotheses that need to be explored, but again should not be used to explain why some survived because this could have arisen for multiple reasons.

A *cohort study* could be undertaken where one group (or cohort) of COVID-19 patients is observed while receiving one treatment, and another group also observed while receiving another course of care. After some time, the two groups would be tested to see which had better outcomes. This type of study has many advantages, and may be the only method feasible, especially in a crisis. However, people can be placed in each cohort for many reasons, some of which can introduce bias to the results. Patients in one cohort might have been sicker or younger than in another cohort. The two cohorts might have lived or been treated differently. In general, multiple reasons can exist for any differences identified between the cohorts. This complicates understanding why any differences were found.

The *randomised controlled trial* (RCT) method can address these limitations. RCTs have two or more groups and people (called subjects or participants) are randomly assigned to each group. The goal is to have the overall composition of each group be as similar as possible in all relevant characteristics. Then each group is treated as similarly as possible during the study, except that one group receives the experimental treatment, and the others do not. This other control group may receive a placebo (like a sugar pill), standard care or some other treatment, while the experimental group receives the treatment being tested. Other steps are taken to control the study, such as 'blinding' the participants, researchers and other clinicians from knowing who is receiving the experiment medication. Other studies have shown that when people or those interacting with them know whether they are in the experimental or control group, this affects the results (Karanicolas, 2010). The overall goal is to have two or more groups that differ only in whether or not they receive the test treatment so that the researchers can be confident that any differences in outcomes came from the treatment. The outcomes will be measured carefully, and subjected to statistical tests to determine if any differences are due to random variation or can be attributed to the experimental treatment. If the differences are large enough, they are said to be 'statistically significant'.

RCTs are widely regarded as the gold standard research method for identifying whether or not an intervention is effective (Hariton and Locascio, 2018). If the experimental group shows significantly better results compared to the control group, and the only difference between the groups was that one received the intervention and the other didn't, we can be confident that the intervention *caused* the benefits. However, this assumes that the study was designed and conducted properly, that everything went according to plan, and that outcomes were measured, reported and analysed accurately. As with any human activity, it is important to carefully examine how the study was conducted and reported, a process called critical appraisal (Will, 2016; O'Mathúna et al, 2022). Those using research results, especially clinicians, policymakers and media reporters, should critically appraise the studies they use to inform their decisions, practice and policy recommendations. The examples here will demonstrate why this has been particularly important during the COVID-19 pandemic, and the consequences of times when it has been lacking.

Randomised controlled trials and the pandemic

Under normal circumstances, RCTs take time to design, organise, ethically approve, conduct, analyse and report the findings. They are also expensive to conduct. For these and other reasons, questions were raised about whether RCTs were appropriate in the early stages of the COVID-19 pandemic. Some reported that, 'In critical situations, large randomized controlled trials are not always feasible or ethical', and they concluded that small sample sizes, non-random allocation procedures, unblinded studies, and unvalidated end points 'may be acceptable' (Kim et al, 2020: 820, 819). Similar viewpoints were expressed during the 2014–16 Ebola epidemic in West Africa (Upshur and Fuller, 2016). The argument is that during a crisis, when people are dying from a serious and novel disease, to allocate patients to different treatment groups randomly and in ways that clinicians and patients were 'blinded' (did not know what they were receiving) is unethical. This was viewed as especially problematic if some patients received a placebo. The argument is made that this abandons some patients to receiving no care, that patients are denied informed choice, and that meanwhile others are receiving a potentially effective treatment. The claim is made that some treatment is better than none, a view that London and Kimmelman (2020) have shown to be based on faulty assumptions.

One aspect of what justifies the use of an RCT during a pandemic is the role of *clinical equipoise*. This term refers to situations where the true effectiveness of an intervention is unknown and is an important ethical justification for research (Rabinstein et al, 2016). In such situations, the experimental intervention might benefit patients, but it might also harm them. Hence, clinicians and policymakers need evidence from research to guide their practice, recommendations and policies. Part of the purpose of research is to figure out whether an experimental intervention provides more benefit than harm. Being experimental does not guarantee benefit, as is sometimes suggested. At the same time, ethical guidelines now require that those in control groups receive all the standard (that is, normal) care they would receive if they were not in the study, and thus should not be 'abandoned' (CIOMS, 2016). The control group could thereby benefit more than the experimental group if those patients avoided the experiment's adverse effects. An example of this occurred in an RCT in Brazil involving critically ill COVID-19 patients (Borba et al, 2020). Eighty-one patients were assigned randomly to two groups receiving different doses of chloroquine. In the group receiving the higher dose, 16 of the 41 patients died (39.0 per cent), while 6 of the 40 in the low-dose group died (15.0 per cent). As a result, the study was terminated early to avoid further harm through the experimental intervention. If this study had been observational, and not an RCT, the greater harm from the higher dose might not have been as clear. Other research methods are important to address other research questions, but they do not show as clearly whether one intervention is as effective or as safe compared to another (NASEM, 2020). Situations exist when RCTs are not feasible or ethical, but they can be conducted during a pandemic (and some were), and therefore should be used to address questions of efficacy and safety. When other study designs are relied upon to make practice and policy decisions, problems arise, as demonstrated in the following example.

Hydroxychloroquine (HCQ) received much media exposure in March 2020 when a study in France reported benefits (Gautret et al, 2020). HCQ and chloroquine are both US Food and Drug Administration (FDA)-approved to prevent and treat malaria, while HCQ is also approved to treat certain autoimmune disorders, which generated additional interest compared to chloroquine. The design details of Gautret and colleagues' study need to be examined as it was not blinded and therefore everyone involved knew who received HCQ and who did not. This introduces a high risk of bias compared to RCTs where everyone is blinded. In addition, some of the patients who received HCQ also received azithromycin, another drug proposed to treat COVID-19, while others did not receive azithromycin. This complicates the analysis, making it difficult to arrive at any conclusions regarding the effectiveness of HCQ alone. Nevertheless, media interest in HCQ was triggered, and the presidents of France and the US came out in support of access to HCQ (Sciama, 2020). The problem, though, is that the design of this study precluded any conclusions concerning whether HCQ was effective or safe. The results of this observational trial were encouraging, but rather than leading to general recommendations for its use, further research was warranted, including RCTs to investigate effectiveness.

Research rigour during pandemics

When effectiveness research of the most appropriate design is conducted, policymakers and clinicians still need to examine whether it was conducted and reported properly by critically appraising its publications. Even in the midst of a pandemic, when evidence is needed urgently, methodological and ethical rigour must be maintained. The urgent need for evidence should not deteriorate into desperation that accepts any evidence at any price. 'But with speed borne of desperation comes risk and confusion – of trials too small to yield answers, of treatments overhyped, and of uncertainty about how to design the best studies possible' (Couzin-Frankel, 2020). Without rigorous studies, designed and conducted appropriately, results can be meaningless or incorrect. This has direct impact on people's health. When ineffective interventions are promoted or effective interventions overlooked, more people may become infected and even die, people's suffering is prolonged and resources are wasted. Policies based on the research that is not appropriate or well-conducted will then need to be changed as more rigorous and reliable results become available. Uncertainty and confusion will reign and may even undermine public trust in research and the scientific process. The importance of using the right sort of research to address various questions places an ethical obligation on users of studies to take the time to carefully examine their details before basing reports, policies or recommendations on them.

After research is conducted, publication normally involves submitting manuscripts to journals which then conduct rigorous peer review. This gives other independent researchers an opportunity to evaluate and critically appraise a study and advise the journal editors on whether they should publish the article or what changes should be made before publishing it. While this process is far from perfect, it allows careful evaluation of an article by other experts in the area before the results are circulated further (Smith, 2006). However, that process takes time as it usually involves other researchers volunteering their time to carefully review the details of the manuscript, provide feedback, give authors time to respond to questions or concerns, and possibly repeat the cycle a few times. A journal may decline to publish the article, and the authors may then start the process with another journal. During the COVID-19 pandemic, changes were introduced to speed up the peer-review process, with some concerns expressed about whether or not this compromised the integrity of the process (Horbach, 2021). Another development was the more common use of 'preprints'. During peer review, articles are typically kept in confidence by everyone, whereas a preprint exists when authors make their manuscripts openly available online while the peer-review process occurs. Some journals run their own preprint websites, while other websites are independent of any particular journal. This makes the results available to clinicians, policymakers and others more quickly, but also means that users must remember that preprints are not peer reviewed. This places a greater responsibility on readers to evaluate the results carefully and carry out their own critical appraisal (CASP, 2019).

Another way that readers can check the rigour of a study is by comparing the published results with how the study was originally planned. Clinical trials are increasingly required to publish a detailed description of their study methods before the study is started in what is called a study protocol. Having these publicly available helps to discourage researchers from arbitrarily changing their studies, possibly in ways that increase the risk of bias. Changes may be required during the research, but then they should be described and justified in published reports. Protocols also help to encourage researchers to publish their results, whether they support the intervention's effectiveness or not. This helps to counteract publication bias, where it has been found that studies with positive findings (that an intervention was effective) tend to be published more easily and quickly than those which find that an intervention is no better than the control (Dwan et al, 2013).

Problems developed with COVID-19 interventions when preprints were relied on too much. Given the publicity focused on HCQ, much interest arose in the results of an RCT from China that became available as a preprint in April 2020 (Chen et al, 2020). News reports noted that COVID-19 patients given HCQ recovered more quickly, one of which reported an infectious disease expert stating that this study is 'going to send a ripple of excitement out through the treating community' (Grady, 2020).

The details of the study revealed a different picture and raised serious questions about how rigorously it had been conducted. Some of this can be seen when the preprint is compared with the original protocol, revealing several major changes (Ferner and Aronson, 2020). The protocol stated that the study would involve three groups of 100 people each receiving 100 mg HCQ twice daily, or 200 mg HCQ twice daily, or a starch pill (a placebo). The preprint described only two groups: the higher dose group and another group receiving standard care (without a placebo). In addition, 62 participants were enrolled (31 in each group), not 300. This raises questions about whether enough participants enrolled to satisfy statistical requirements. Other significant changes were made and not discussed or justified in the preprint. These details raise concerns about why changes were made, as they weakened the rigour of the methods and introduced a high risk of bias. For reasons like this, news media, practitioners and policymakers should be tentative in their recommendations based on preprints and wait for peer review (or conduct their own thorough critical appraisal). To date, this study has not been published in a peer-reviewed journal.

Research on HCQ has continued to test its effectiveness in treating and preventing COVID-19. When several studies are published on a similar topic, they can be summarised in a *systematic review*. If the studies are similar enough, their quantitative data can be combined in a *meta-analysis*. Cochrane systematic reviews are recognised internationally as being of high quality and conducted by independent investigators. A Cochrane systematic review of HCQ published in February 2021 included 12 RCTs conducted in several countries (Singh et al, 2021: 2). The trials varied considerably in many ways and included hospitalised and outpatient participants. Overall, the authors concluded that HCQ 'has little or no effect' on risk of death or serious complications from COVID-19. The results were so clear-cut, they concluded, 'No further trials of chloroquine or hydroxychloroquine should be conducted for the treatment of COVID-19'. For preventing COVID-19, few trials were found and those suggest that it is unlikely that HCQ is effective at protecting people from infection. An accompanying editorial decried the 'false hope' communicated by some about HCQ that has contributed to 'a climate of mistrust' in science and public policy (Gould and Norris, 2021).

Another preprint has been found to be the source of even greater problems. Ivermectin is an FDA-approved drug used orally to treat parasitic worms and topically for head lice and rosacea. Two researchers were awarded the 2015 Nobel Prize in Physiology and Medicine for discoveries that led to ivermectin (Formiga et al, 2020). Billions of doses have been distributed for use in animals and humans, with an excellent safety record (Chaccour et al, 2020). In April 2020, a peer-reviewed article was published showing that ivermectin prevented the replication of the SARS-CoV-2 virus (Caly et al, 2020). Crucially, the authors highlighted that the study was conducted in cells in laboratory experiments, not in humans or even animals. This is called an *in vitro* study (literally meaning, in glass, as opposed to *in vivo*). Even more importantly, the doses used in the laboratory were 50 to 100 times higher than the equivalent doses approved for use in humans. In spite of this, public interest was fuelled by media reports that often did not describe both of these crucial details (Slisco, 2020). This is another example of how understanding the methodological details of a study is vital before it informs practice or policy.

Interest in ivermectin for treating COVID-19 was particularly high in Latin America where it is used widely against parasites. This resulted in shortages for its approved uses, which points to another ethical issue with the inappropriate use of a medication without evidence of benefit (Chaccour et al, 2020). The *potential*, but uncertain, benefit of a drug for a new disease must be balanced against the *actual* benefits demonstrated for approved uses. The same problem occurred with HCQ when its use for COVID-19 led to shortages for those who needed it as an approved treatment for autoimmune diseases and malaria (DeJong and Wachter, 2020).

Using a limited resource for unapproved purposes can have far-reaching consequences for others. The resulting scarcity of ivermectin led to people in Latin America (and later in the US) using veterinary formulations of ivermectin. This became so widespread that the FDA issued a warning about the different types of products (Solomon, 2020). Policymakers in Latin America faced ethical dilemmas over how to address the situation. Víctor Zamora, Peru's health minister, stated that while they were developing guidelines and policies for clinicians, they did not have time 'to wait for scientific evidence' (Offord, 2020). Dr Eduardo Gotuzzo, a tropical medicine expert, asked, 'What do you do?' for seriously ill patients: 'Give them water?' (cited in Offord, 2020). Without further support from research, Peru added ivermectin to its COVID-19 clinical guidelines in May 2020, followed by Bolivia, and later Brazil and Chile (Davey, 2020). Doctors reported being pressured to prescribe ivermectin, while researchers reported difficulties conducting RCTs because patients did not want to take any chance of being assigned to a group that didn't receive ivermectin. Gradually, however, use of ivermectin revealed problems, leading one Peruvian global health researcher to state, "I think people have lost faith in science … and it has been very, very bad for us in Latin America" (cited in Offord, 2020). Policymakers and clinicians will experience pressure to give guidance in the absence of evidence, but doing so is fraught with danger and may have serious consequences, short term and long term.

The situation with ivermectin has been worsened by problematic preprints. In April 2020, a preprint reported findings from a large medical database owned by a US company called Surgisphere (Patel et al, 2020a). This database was reported to have collected details on COVID-19 patients in hospitals around the world. Surgisphere's owner, Dr Sapan Desai, co-authored this preprint stating that COVID-19 patients had improved survival rates after receiving ivermectin. The death rate was 18.6 per cent for COVID-19 patients not receiving ivermectin but was 7.7 per cent for those receiving ivermectin. Patel and colleagues posted a new version of their study (2020b) after they compared similar patients with one another (rather than reporting overall averages). They now reported a death rate of 8.5 per cent in patients not receiving ivermectin compared to 1.4 per cent for similar patients taking ivermectin.

While ivermectin was becoming more widely used in Latin American, Carlos Chaccour, a Venezuelan physician and researcher, familiar with ivermectin, identified serious discrepancies with the Surgisphere data (Davey, 2020; Offord, 2020). He noted that the data included 52 COVID-19 patients receiving ivermectin before it was being recommended for COVID-19. Three patients on ventilators were included from African hospitals, but only two COVID-19 patients were known in Africa at this time, and neither received ventilator support. A third patient was later identified, but did not require a ventilator. Chaccour had worked in Africa for many years and questioned whether many African hospitals had the electronic patient record system that Surgisphere claimed they used to collect data. Other articles using Surgisphere data were published in the *New England Journal of Medicine* (Mehra et al, 2020) and *The Lancet* (Mehra et al, 2020), two of the leading medical journals in the world. Although they temporarily had a major impact on HCQ research and policy, they were later retracted. In this case, investigative news reporters identified inconsistencies and contradictions in the articles that have led to all the Surgisphere publications being viewed as potentially fraudulent (Davey and Kirchgaessner, 2020).

The idea of people publishing fraudulent research results may seem incomprehensible during a pandemic which has claimed the lives of millions and caused untold suffering. The problem has grown to the point where the former editor of the *BMJ*, another leading medical journal, asked whether it should be assumed that 'health research is fraudulent until proven otherwise?' (Smith, 2021). The implications for policymakers and practitioners are significant. Whether fraud or sloppiness or incompetence, all research reports must be thoroughly evaluated for their methodological rigour and ethical quality before being used to influence policy or practice.

One more example will be provided, taking up the ongoing saga with ivermectin. Following on from the initial reports of ivermectin's use for COVID-19, RCTs and other types of studies were conducted. A number of meta-analyses of ivermectin for COVID-19 have been published. One included ten RCTs and found that 'ivermectin did not reduce all-cause mortality, length of stay or viral clearance vs. controls' and concluded that '[ivermectin] is not a viable option to treat COVID-19 patients' (Roman et al, 2021). Another review with meta-analyses included 18 RCTs and 'found large, statistically significant reductions in mortality, time to clinical recovery, and time to viral clearance' and concluded that 'an oral agent effective in all phases of COVID-19 has been identified' (Kory et al, 2021). How is a busy clinician or policymaker to decide which systematic review to rely on? Critical appraisal of reports is crucial as this reveals the strengths and weaknesses of each review (O'Mathúna, 2022).

One RCT plays a large role in influencing the results of many ivermectin meta-analyses and systematic reviews. This article was first made available on a preprint website in November 2020 and subsequently updated twice (Elgazzar et al, 2020). With 600 subjects, it was at the time the single largest RCT available on ivermectin for COVID-19 and thus has the largest impact on systematic review findings. As of 14 July 2021, this study should never be included in another review or allowed to influence policy or practice. The study has been withdrawn by the website hosting the preprint, originally stating this was 'due to ethical concerns', but this was later changed to 'an expression of concern communicated' with the website staff that is 'now under formal investigation'.

An in-depth critical appraisal of the Elgazzar et al RCT was undertaken initially by a medical student in London, but then verified by other researchers (Davey, 2021). The preprint first raised questions by having large sections of its introduction allegedly plagiarised from other sources (Lawrence, 2021). The preprint included a link to the study's raw data, which allowed other researchers to identify many discrepancies and concerns (Brown, 2021). For example, the preprint stated that subjects were all over 18 years old, but the raw data included patients aged under 18; the raw data included some collected before the study was reported to have begun; and most of the descriptive data (like mean ages, gender and so on) differ between the raw data and the preprint. Such discrepancies could be due to poor research technique or fraud, and might not affect the study's conclusions about the effectiveness of ivermectin. However, they point to very poor research practice, and should lead to even greater scrutiny of the effectiveness data. Policymakers should be very hesitant to rely on research that raises so many concerns. Thorough appraisal of this study and its data revealed that dozens of data cells appear to have been copied from one participant to another, with small changes introduced in some places. This raises serious questions about the authenticity of the data. Whether fraudulent or not, the article has been withdrawn, and with it one of the largest studies supporting the effectiveness of ivermectin for COVID-19. What remains are several smaller studies of ivermectin, many of which did not find ivermectin effective for a number of outcomes. At the time of writing, the evidence from rigorous and ethical research does not support the use of ivermectin in the treatment of COVID-19. The WHO reached the same conclusion, but did recommend that ivermectin be studied further in RCTs (WHO, 2021). Many such trials are under way, and will hopefully provide more definitive answers on ivermectin and other treatments for COVID-19 (Hart, 2021).

Good research during the pandemic

This chapter should not end on a note of despondency over the state of pandemic research. In contrast, the literature includes encouraging examples of research at its best even with the challenges of conducting research in the midst of a pandemic. Remdesivir was the first COVID-19 treatment to receive FDA approval in October 2020 (FDA, 2020). Remdesivir had shown encouraging antiviral results in animal and laboratory studies and some promise against two other coronavirus diseases, SARS and MERS, prior to COVID-19 (Beigel et al, 2020). Beigel and colleagues enrolled more seriously ill COVID-19 patients in a large RCT called the ACTT-1 Trial to test the efficacy of remdesivir. After enrolling over 1,000 patients, the safety monitoring committee stopped the trial because it was clear that patients receiving remdesivir were benefiting more than those receiving placebo. Those taking a placebo took an average of 15 days to recover, while those taking remdesivir took 11 days. The mortality rate showed a positive trend, going from 11.9 per cent to 7.1 per cent with remdesivir, but this was not statistically significant. They also found that the most seriously ill patients, those on ventilators, did not benefit from the drug. The trial was stopped for ethical reasons: so that all patients would have the opportunity to receive remdesivir. Because the trial was an RCT, it demonstrated significant differences between the groups of patients and provided clear evidence to guide practice and policy. It also demonstrated that rigorous research can be designed and conducted during a pandemic, with other researchers stating that, ‘Conducting such a clinical trial only a few months after SARS-CoV-2 was discovered is an extraordinary achievement’ (Dolin and Hirsch, 2020: 1886).

While ACTT-1 used standard RCT methods, others have introduced appropriate changes into trial methods to meet the challenges of conducting research during a pandemic. RCTs are designed to be conducted as planned and without looking at the data until the trial is completed (or only under very specific criteria, usually to do with safety issues). This helps to minimise bias, but has been criticised for making RCTs overly inflexible. During a crisis, especially when little evidence exists to guide decisions, greater flexibility has been called for. *Adaptive design* methodology is one way to introduce appropriate flexibility into trials. The basic strategy is to design flexibility into a trial by building into protocols how trials could change as preliminary data becomes available. Such pre-planned adaptations are intended to avoid researchers making unplanned (or ad hoc) changes, which can introduce bias.

Many adaptations can be planned, including when and why to stop groups with interventions showing themselves to be ineffective or harmful, criteria to change doses as evidence shows this is needed, adding new interventions to be evaluated, adapting the sample size, refocusing the trial on specific patient groups who are responding better than others, and many other possibilities (Pallmann et al, 2018). However, the additional flexibility introduces greater complexity and requires much more statistical expertise to analyse the results appropriately. The WHO-sponsored Solidarity Trial used an adaptive design to evaluate the effectiveness of several interventions used to treat COVID-19. It also built in ways to address global health disparities and research capacity building by providing training, support and funding to allow researchers in low-income settings to get involved in the study and recruit participants from their countries. This had the additional benefit of being able to identify whether certain populations or settings around the world were impacted differently by the medications being tested. The Solidarity Trial has enrolled more than 11,000 participants at 405 hospitals in 30 countries and continues to provide high-quality evidence on the effectiveness of COVID-19 interventions (WHO Solidarity Trial Consortium, 2021). Other innovative approaches to conducting public health research have been developed during COVID-19 (Beckman and Smith, 2021).

The path to developing any new treatment, especially for a new disease, should be expected to have twists and turns. The Solidarity Trial included remdesivir and its interim analysis concluded that remdesivir provided no benefit to the COVID-19 patients it enrolled (Carey, 2020). As a result of this and other trials, WHO has recommended against its use, in contrast to the FDA's approval of its use. While such conflicting conclusions are frustrating for patients, clinicians and policymakers, such is the way with a new disease and the search for new treatments. Results often vary and fluctuate, in contrast to the dramatic improvements alleged of ivermectin. Unfortunately, some uncertainty continues, but this is normal in research, and has to be understood by policymakers. Sometimes inconsistent evidence develops at first, and over time a clearer picture emerges about something's effectiveness and whether particular patients benefit more or less. This is part of the normal research and scientific process, requiring that we remain humble and duly cautious when presenting research results. For practitioners and policymakers, this means that practice and policy will need to be regularly reviewed and updated, with guidelines changing if and when the evidence from research changes.

A chapter on research during COVID-19 would not be complete without some discussion of COVID-19 vaccines. In spite of the ongoing controversy over their implementation, they are another remarkable success story for research. During the early days of the pandemic, researchers cautioned that 'vaccine development is a lengthy, expensive process. Attrition is high, and it typically takes multiple candidates and many years to produce a licensed vaccine' (Lurie et al, 2020). Dozens of candidates were developed, and before the end of 2020, vaccines had started to receive emergency use authorisation.

While the speed of their development and approval has raised safety questions, several issues contributed to their rapid development. Many governments, foundations and corporations put large amounts of funding towards this research. Phases during vaccine development that are normally sequential were overlapped, which introduced financial risk, but not necessarily safety risks (Thompson, 2020). For example, the Gates Foundation spent billions of US dollars on constructing seven factories, even though only one or two might eventually be suitable for producing COVID-19 vaccines (Hamilton, 2020). At the same time, the shorter time for development means that longer-term data on safety and effectiveness was not available at the time of emergency use authorisation and is still being collected while the vaccines are being distributed. Such balancing of the benefits of halting the pandemic against the potential harms of distributing the COVID-19 vaccines is part of the challenging work of public health policymakers. It points to the importance of having rigorous research evidence available to inform policy, of having careful critical appraisal of such research, and careful ethical decision-making that is transparent and accountable to the public. Much more could be said about COVID-19 vaccines, including ethical issues about their international distribution and the policies that have allowed wealthier nations to stockpile much of the supply, leaving low-income countries with little or none (Bollyky and Brown, 2020). Vaccine challenge trials, which deliberately expose healthy volunteers to the SARS-CoV-2 virus after receiving an experimental vaccine, raise many ethical challenges (Chappell and Singer, 2020). While they have been proposed for COVID-19 vaccines, they have not been needed because the virus continues to circulate widely in communities. Research has contributed immensely to the pandemic response, and much more research remains to be done.

Conclusion

Research plays an important role in generating evidence to inform and guide policy in many areas, including pandemics and other crises. A wide variety of research methods is available, each of which addresses different types of research questions. No one research method is the most rigorous or most ethical, but each method must be examined against the standards for that particular method. Users of research should understand which methods best address which type of research question (such as whether the question is about something's effectiveness or someone's experience of something).

Those who read and make recommendations based on research need to understand the basic features of different types of research methods, such as *in vitro* studies, case studies, cohort studies, randomised controlled trials (RCTs), systematic reviews and meta-analyses. This is key to identifying whether or not the right method has been used to answer the corresponding question.

All research reports must be critically appraised before being used to influence policy or practice. Those advising patients, practitioners or policymakers must be skilled and trained in critical appraisal, or be able to access sources where original research has been critically appraised by others. For example, high-quality systematic reviews will include a detailed critical appraisal of the studies they reviewed.

During a pandemic or crisis, steps should be taken to make research available more rapidly, but not at the expense of rigour and quality. Preprints are an example of how results can be made available more quickly, but they increase the importance of readers carefully evaluating the results. Preprints must be treated as tentative findings until after they have been peer reviewed and critically appraised. Press releases, blogs and other sources of research results must be treated with a greater degree of scepticism than other forms of pre-appraised and peer-reviewed evidence.

Pandemics involving novel diseases, like COVID-19, are particularly challenging because of the scale of uncertainty, at least at first, and how evidence will change as time goes on. This uncertainty and variability is part of the complex situation and how research develops in a new area, and needs to be understood by decision-makers and communicated transparently and honestly. Policymakers should be prepared to revisit and possibly revise policies as new evidence becomes available. This planned review should be communicated with new policy announcements.

The evidence from research evolves and often changes as time goes on and more trials are conducted. Policymakers should expect new data to become available that may provide additional support for existing policies, or suggest changes and adaptations. The changing public health mandates during COVID-19 are to be expected given how the virus, society and research develops and adapts. Policymakers must therefore monitor the data coming from many different sources as a crisis continues and evolves. Secondary data analysis, such as systematic reviews and meta-analyses are crucial to understanding the 'direction of travel' of many factors. As with primary research studies, these analyses must also be critically appraised before informing policies. In addition, the effectiveness of the policies themselves must be monitored and adapted accordingly as the evidence becomes available.

Pandemics and disasters have happened throughout history and will continue to happen. Policymakers should ensure that pandemic preparedness plans and disaster risk reduction plans are prepared at every level and for various situations, that they are piloted and tested, and that they are reviewed regularly. Research has an important role here also, to evaluate the effectiveness of these plans.

References

Beckman, J. and Smith, A. (2021) 'PRINCIPLE: accelerated recruitment for clinical trials through rapid access to large, linked datasets and integrated patient engagement', Health Research Data UK, 1 April, available from: https://www.hdruk.ac.uk/case-studies/accelerated-recruitment-for-clinical-trials-through-rapid-access-to-large-linked-datasets-and-integrated-patient-engagement [accessed 2 August 2021].

Beigel, J.H., Tomashek, K.M. et al (2020) 'Remdesivir for the treatment of Covid-19: final report', *New England Journal of Medicine*, 383(19): 1813–26.

Bollyky, T.J. and Brown, C.P. (2020) 'Vaccine nationalism will prolong the pandemic', Foreign Affairs, 29 December, available from: https://www.foreignaffairs.com/articles/world/2020-12-29/vaccine-nationalism-will-prolong-pandemic [accessed 2 August 2021].

Borba, M.G.S., Val, F.F.A. et al (2020) 'Effect of high vs low doses of chloroquine diphosphate as adjunctive therapy for patients hospitalized with Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infection: a randomized clinical trial', *Journal of the American Medical Association Network Open*, 3(4): art e208857, available from: https://doi.org/10.1001/jamanetworkopen.2020.8857 [accessed 6 February 2022].

Borges do Nascimento, I.J., von Groote, T.C. et al (2020) 'Clinical, laboratory and radiological characteristics and outcomes of novel coronavirus (SARS-CoV-2) infection in humans: a systematic review and series of meta-analyses', *PLoS ONE*, 15(9): art e0239235, available from: https://doi.org/10.1371/journal.pone.0239235 [accessed 6 February 2022].

Brown, N. (2021) 'Some problems in the dataset of a large study of Ivermectin for the treatment of Covid-19', *Nick Brown's Blog*, 15 July, available from: https://steamtraen.blogspot.com/2021/07/Some-problems-with-the-data-from-a-Covid-study.html [accessed 2 August 2021].

Caly, L., Druce, J.D. et al (2020) 'The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 *in vitro*', *Antiviral Research*, 178: art 104787, available from: https://doi.org/10.1016/j.antiviral.2020.104787 [accessed 6 February 2022].

Carey, B. (2020) 'W.H.O. rejects antiviral drug remdesivir as a Covid treatment', *The New York Times*, 19 November, available from: https://www.nytimes.com/2020/11/19/health/remdesivir-covid-19.html [accessed 2 August 2021].

CASP (Critical Appraisal Skills Program) (2019) 'CASP checklists', available from: https://casp-uk.net/casp-tools-checklists/ [accessed 2 August 2021].

Chaccour, C., Hammann, F., Ramón-Garcín, S. and Rabinovich, N.R. (2020) 'Ivermectin and COVID-19: keeping rigor in times of urgency', *American Journal of Tropical Medicine and Hygiene*, 102(6): 1156–7.

Chappell, R.Y. and Singer, P. (2020) 'Pandemic ethics: the case for risky research', *Research Ethics*, 16(3/4), available from: https://doi.org/10.1177/1747016120931920 [accessed 6 February 2022].

Chen, Z., Hu, J., Zhang, Z., Jiang, S., Han, S., Yan, D., Zhuang, R., Hu, B. and Zhang, Z. (2020) 'Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial', MedRxIV, 10 April, available from: https://doi.org/10.1101/2020.03.22.20040758 [accessed 2 August 2021].

CIOMS (Council for International Organizations of Medical Sciences) (2016) *International Ethical Guidelines for Health-Related Research Involving Humans*, Geneva: CIOMS, available from: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf [accessed 2 August 2021].

Couzin-Frankel, J. (2020) '"We've got to be able to move more quickly": the pandemic reality of COVID-19 clinical trials', *Science*, 16 June, available from: https://doi.org/10.1126/science.abd3588 [accessed 2 August 2021].

Davey, M. (2020) 'Unreliable data: how doubt snowballed over Covid-19 drug research that swept the world', *The Guardian*, 4 June, available from: https://www.theguardian.com/world/2020/jun/04/unreliable-data-doubt-snowballed-covid-19-drug-research-surgisphere-coronavirus-hydroxychloroquine [accessed 2 August 2021].

Davey, M. (2021) 'Huge study supporting ivermectin as Covid treatment withdrawn over ethical concerns', *The Guardian*, 15 July, available from: https://www.theguardian.com/science/2021/jul/16/huge-study-supporting-ivermectin-as-covid-treatment-withdrawn-over-ethical-concerns [accessed 2 August 2021].

Davey, M. and Kirchgaessner, S. (2020) 'Surgisphere: mass audit of papers linked to firm behind hydroxychloroquine *Lancet* study scandal', *The Guardian*, 10 June, available from: https://www.theguardian.com/world/2020/jun/10/surgisphere-sapan-desai-lancet-study-hydroxychloroquine-mass-audit-scientific-papers [accessed 2 August 2021].

DeJong, C. and Wachter, R.M. (2020) 'The risks of prescribing hydroxychloroquine for treatment of COVID-19: first, do no harm', *JAMA Internal Medicine*, 180(8): 1118–19.

Dolin, R. and Hirsch, M.S. (2020) 'Remdesivir: an important first step', *New England Journal of Medicine*, 383(19): 1886–7.

Dwan, K., Gamble, C., Williamson, P.R. and Kirkham, J.J. (2013) 'Systematic review of the empirical evidence of study publication bias and outcome reporting bias: an updated review', *PLoS ONE*, 8(7): art e66844, available from: https://doi.org/10.1371/journal.pone.0066844 [accessed 6 February 2022].

Elgazzar, A., Eltaweel, A., Youssef, S.A., Hany, B., Hafez, M. and Moussa, H. (2020) 'Efficacy and safety of ivermectin for treatment and prophylaxis of COVID-19 pandemic', version 3, Research Square, 28 December [withdrawn], available from: https://www.researchsquare.com/article/rs-100956/v3 [accessed 2 August 2021].

FDA (US Food and Drug Administration) (2020) 'FDA approves first treatment for COVID-19', FDA, 22 October, available from: https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19 [accessed 2 August 2021].

Ferner, R.E. and Aronson, J.K. (2020) 'Hydroxychloroquine for COVID-19: what do the clinical trials tell us?', *Centre for Evidence-Based Medicine*, 14 April, available from: https://www.cebm.net/covid-19/hydroxychloroquine-for-covid-19-what-do-the-clinical-trials-tell-us/ [accessed 2 August 2021].

Formiga, F.R., Leblanc, R., de Sousas Rebouças, J., Fariase, L.P., de Oliveira, R.N. and Pena, L. (2020) 'Ivermectin: an award-winning drug with expected antiviral activity against COVID-19', *Journal of Controlled Release*, 329: 758–61.

Gautret, P., Lagier, J.-C. et al (2020) 'Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial', *International Journal of Antimicrobial Agents*, 56(1): art 105949, available from: https://doi.org/10.1016/j.ijantimicag.2020.105949 [accessed 6 February 2022].

Gould, S. and Norris, S.L. (2021) 'Contested effects and chaotic policies: the 2020 story of (hydroxy) chloroquine for treating COVID-19', *Cochrane Database of Systematic Reviews*, paper ED000151, 25 March, available from: https://doi.org/10.1002/14651858.ED000151 [accessed 2 August 2021].

Grady, D. (2020) 'Malaria drug helps virus patients improve, in small study', *New York Times*, 1 April, available from: https://www.nytimes.com/2020/04/01/health/hydroxychloroquine-coronavirus-malaria.html [accessed 2 August 2021].

Hamilton, I.A. (2020) 'Bill Gates is funding new factories for potential coronavirus vaccines', World Economic Forum, 6 April, available from: https://www.weforum.org/agenda/2020/04/bill-gates-7-potential-coronavirus-vaccines [accessed 2 August 2021].

Hariton, E. and Locascio, J.J. (2018) 'Randomised controlled trials: the gold standard for effectiveness research', *BJOG*, 125(13): 1716.

Hart, R. (2021) 'Ivermectin – widely used to treat Covid-19 despite being unproven – is being studied in the U.K. as a potential treatment', Forbes, 23 June, available from: https://www.forbes.com/sites/roberthart/2021/06/23/ivermectinwidely-used-to-treat-covid-19-despite-being-unprovenis-being-studied-in-the-uk-as-a-potential-treatment [accessed 2 August 2021].

Horbach, S.P.J.M. (2021) 'No time for that now! Qualitative changes in manuscript peer review during the Covid-19 pandemic', *Research Evaluation*, 30(3): 231–9.

Karanicolas, P.J., Farrokhyar, F. and Bhandari, M. (2010) 'Blinding: who, what, when, why, how?', *Canadian Journal of Surgery*, 53(5): 345–8.

Kim, A.H.J., Sparks, J.A. et al (2020) 'A rush to judgment? Rapid reporting and dissemination of results and its consequences regarding the use of hydroxychloroquine for COVID-19', *Annals of Internal Medicine*, 172(12): 819–21.

Kory, P., Meduri, G.U., Varon, J., Iglesias, J. and Marik, P.E. (2021) 'Review of the emerging evidence demonstrating the efficacy of ivermectin in the prophylaxis and treatment of COVID-19', *American Journal of Therapeutics*, 28(3): e299–e318, available from: https://doi.org/10.1097/MJT.0000000000001377 [accessed 6 February 2022].

Lawrence, J. (2021) 'Why was a major study on ivermectin for COVID-19 just retracted?' Grftr, 15 July, available from: https://grftr.news/why-was-a-major-study-on-ivermectin-for-covid-19-just-retracted/ [accessed 2 August 2021].

London, A.J. and Kimmelman, J. (2020) 'Against pandemic exceptionalism', *Science*, 368(6490): 476–7.

Lurie, N., Saville, M., Hatchett, R. and Halton, J. (2020) 'Developing Covid-19 vaccines at pandemic speed', *New England Journal of Medicine*, 382(21): 1969–73, available from: https://www.nejm.org/doi/full/10.1056/NEJMp2005630 [accessed 7 February 2022].

Mehra, M.R., Desai, S.S., Kuy, S.R., Henry, T.D., and Patel, A.N. (2020a) 'Cardiovascular disease, drug therapy, and mortality in Covid-19', *New England Journal of Medicine*, 382(25): art e102 [retracted], available from: https://www.nejm.org/doi/full/10.1056/nejmoa2007621 [accessed 7 February 2022].

Mehra, M.R., Desai, S.S., Ruschitzka, F. and Patel, A.N. (2020b) 'Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis', *The Lancet*, 22 May, [retracted], available from: https://doi.org/10.1016/S0140-6736(20)31180-6 [accessed 2 August 2021].

NASEM (National Academies of Sciences, Engineering, and Medicine) (2020) *Evidence-Based Practice for Public Health Emergency Preparedness and Response*, Washington, DC: National Academies Press.

Offord, C. (2020) 'Surgisphere sows confusion about another unproven COVID-19 drug', The Scientist, 16 June, available from: https://www.the-scientist.com/news-opinion/surgisphere-sows-confusion-about-another-unproven-covid19-drug-67635 [accessed 2 August 2021].

O'Mathúna, D.P. (2022) 'Ivermectin and the integrity of healthcare evidence during COVID-19', *Frontiers in Public Health*, 10: 788972.

O'Mathúna, D.P., Fineout-Overholt, E., Graham, D. and Canada, A. (2022) 'Critically appraising quantitative evidence for clinical decision-making', in B.M. Melnyk and E. Fineout-Overholt (eds), *Evidence-Based Practice in Nursing & Healthcare: A Guide to Best Practice* (5th edn), Philadelphia, PA: Wolters Kluwer, chapter 6.

Pallmann, P., Bedding, A.W. et al (2018) 'Adaptive designs in clinical trials: why use them, and how to run and report them', *BMC Medicine*, 16(29), available from: https://doi.org/10.1186/s12916-018-1017-7 [accessed 7 February 2022].

Patel, A.N., Desai, S.S., Grainger, D.W. and Mehra, M.R. (2020a) 'Ivermectin in COVID-19 related critical illness', ISGlobal.org, 6 April, [retracted], available from: https://www.isglobal.org/documents/10179/6022921/Patel+et+al.+2020+version+1.pdf/fab19388-dc3e-4593-a075-db96f4536e9d [accessed 2 August 2021].

Patel, A.N., Desai, S.S., Grainger, D.W. and Mehra, M.R. (2020b) 'Usefulness of ivermectin in COVID-19 illness', ISGlobal.org, 20 April, [retracted], available from: https://www.isglobal.org/documents/10179/6022921/Patel+et+al.+2020+version+2.pdf/adf390e0-7099-4c70-91d0-e0f7a0b69e14 [accessed 2 August 2021].

Rabinstein, A.A., Brinjikji, W. and Kallmes, D.F. (2016) 'Equipoise in clinical trials: angst and progress', *Circulation Research*, 119(7): 798–800.

Roman, Y.M., Burela, P.A., Pasupuleti, V., Piscoya, A., Vidal, J.E. and Hernandez, A.V. (2021) 'Ivermectin for the treatment of COVID-19: a systematic review and meta-analysis of randomized controlled trials', *Clinical Infectious Diseases*, ahead of print, available from: https://doi.org/10.1093/cid/ciab591 [accessed 7 February 2022].

Sciama, Y. (2020) 'Is France's president fueling the hype over an unproven coronavirus treatment?' *Science*, 9 April, available from: https://www.sciencemag.org/news/2020/04/france-s-president-fueling-hype-over-unproven-coronavirus-treatment [accessed 2 August 2021].

Singh, B., Ryan, H., Kredo, T., Chaplin, M. and Fletcher, T. (2021) 'Chloroquine or hydroxychloroquine for prevention and treatment of COVID-19', *Cochrane Database of Systematic Reviews*, 2021(2): art CD013587, available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013587.pub2 [accessed 2 August 2021].

Slisco, A. (2020) 'Anti-parasite drug used since 1980s may help stop coronavirus, new study says', Newsweek, 3 April, available from: https://www.newsweek.com/anti-parasite-drug-used-since-1980s-may-help-stop-coronavirus-new-study-says-1496083 [accessed 2 August 2021].

Smith, R. (2006) 'Peer review: a flawed process at the heart of science and journals', *Journal of the Royal Society of Medicine*, 99(4): 178–82.

Smith, R. (2021) 'Time to assume that health research is fraudulent until proven otherwise?' The BMJ Opinion [blog], 5 July, available from: https://blogs.bmj.com/bmj/2021/07/05/time-to-assume-that-health-research-is-fraudulent-until-proved-otherwise/ [accessed 2 August 2021].

Solomon, S. (2020) 'FDA letter to stakeholders: do not use ivermectin intended for animals as treatment for COVID-19 in humans', US Food and Drug Administration, 10 April, available from: https://www.fda.gov/animal-veterinary/product-safety-information/fda-letter-stakeholders-do-not-use-ivermectin-intended-animals-treatment-covid-19-humans [accessed 2 August 2021].

Thompson, S.A. (2020) 'How long will a vaccine really take?' The New York Times, 30 April, available from: https://www.nytimes.com/interactive/2020/04/30/opinion/coronavirus-covid-vaccine.html [accessed 2 August 2021].

UNDRR (United Nations Office for Disaster Risk Reduction) (2020) 'Terminology', available from: https://www.undrr.org/terminology [accessed 2 August 2021].

UNDRR (United Nations Office for Disaster Risk Reduction) (2021) *UNDRR Strategic Framework 2022–2025*, available from: https://www.undrr.org/publication/undrr-strategic-framework-2022-2025 [accessed 2 August 2021].

Upshur, R. and Fuller, J. (2016) 'Randomized controlled trials in the West African Ebola virus outbreak', *Clinical Trials*, 13(1): 10–12.

WHO (World Health Organization) (2020) A Coordinated Global Research Roadmap: 2019 Novel Coronavirus, available from: https://www.who.int/blueprint/priority-diseases/key-action/Coronavirus_Roadmap_V9.pdf [accessed 2 August 2021].

WHO (World Health Organization) (2021) 'WHO advises that ivermectin only be used to treat COVID-19 within clinical trials', 31 March, available from: https://www.who.int/news-room/feature-stories/detail/who-advises-that-ivermectin-only-be-used-to-treat-covid-19-within-clinical-trials [accessed 2 August 2021].

WHO (World Health Organization) Solidarity Trial Consortium (2021) 'Repurposed antiviral drugs for Covid-19: interim WHO Solidarity Trial results', *New England Journal of Medicine*, 384(6): 497–511.

Will, R. (2016) 'Critical appraisal: a checklist', Students 4 Best Evidence, available from: https://s4be.cochrane.org/blog/2016/09/06/critical-appraisal-checklist/ [accessed 2 August 2021].

13

Ethics in research for resilience and societal collapse

Ian Roderick

> In old days there were angels who came and took men by the hand and led them away from the city of destruction. We see no white-winged angels now. But yet men are led away from threatening destruction: a hand is put into theirs, which leads them forth gently towards a calm and bright land, so that they look no more backward; and the hand may be a little child's.
>
> George Eliot, *Silas Marner* (1861)

Introduction

You could be forgiven for thinking that dire warnings of 'the end of the world' are the rantings of a tiny minority. Soothsayers and prophets have long claimed that all is about to collapse in misery and shame – usually because of a deity taking revenge on our collective lack of concern, greed and venality. The doomsters either get their dates wrong or the deity informs them of a last-minute reprieve and sets a new date. The eschatological prophesies have since shifted away from gods to nature itself, an anthropomorphised container of all life and fount of well-being. Named after a mythical goddess, Gaia, the Earth, is out to get us (Lovelock, 2006) due to the careless damage we are collectively imposing on her. We now threaten our own demise by destroying the systems of life that provide for us. Nature considers humans much like a disease, running a temperature for a short while and then recovering having eliminated the problem – us.

Thomas Malthus's *An Essay on the Principle of Population* (1798) predicted population growth outpacing the production of food leading to a collapse of civilisation. World population then was around one billion people; the population is now nearing eight billion people, and the first of the United Nations' (UN's) Sustainable Development Goals (SDG) is to attain zero hunger in the world by 2030 – confounding Malthus's prediction. Human ingenuity, scientific advancement in understanding, and a massive exploitation of fossil fuel supplied the energy needed to increase food production.

The consumption of many other goods increased as did the expansion of our exploitation of nature and resources: so much so that we talk of peak production, which is the time when resources start the downhill path to depletion and nature gets weaker and thinner until it ceases to function for our benefit. This hollowing out of ecosystems, loss of biodiversity and depletion of mineral capital has been hidden from common knowledge behind the glamour of immediate wealth – at least for some.

Other problems were hinted at. In the late 19th century, Svante Arrhenius and colleagues suggested that human-created emissions of carbon dioxide would eventually lead to global warming. However, because of the relatively low rate of carbon dioxide production at that time, Arrhenius thought global warming would take thousands of years, and he expected it would be beneficial to humanity. When *The Limits to Growth* (Meadows et al, 1972) was first published, climate change was just one factor in the 'pollution' variable of the world model. In the same year, at the first UN environment conference, in Stockholm, climate change hardly registered on the agenda – and the global population reached about four billion people.

Over the next 50 years we saw the Intergovernmental Panel on Climate Change (IPCC) created to collate and assess the evidence (1988), the Earth Summit was held in Rio de Janeiro (1992) where governments agreed on the UN Framework Convention on Climate Change with the objective of the 'stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system' (UN, 1992: art 2), then in 1997 the Kyoto Protocol was agreed, when developed nations pledged to reduce emissions by an average of 5 per cent by the period 2008–12, the hockey stick graph appeared around then, and on we went through a long series of high-level meetings, pledges, targets and arguments, arriving at the Paris Accord in 2016 when 196 governments pledged to keep the world's average temperature rise to only 2°C above what it had been before the Industrial Revolution – and preferably to keep the increase below 1.5°C. As I write, we await the next meeting, Glasgow COP26, in 2021, which will assess progress and make more pledges; and the world is waking up to the possibility of massive disruptions, severe dislocations of people and perhaps societal collapse. The concept of net-zero carbon emissions is taking centre stage, but the full consequences are still actors in the wings.

Extensive research has grown our knowledge and understanding of the world in turn throwing up a plethora of interconnected complex global issues and potential tipping points. Coterminous with climate change, the list of threats heading towards us is growing almost daily – biodiversity loss, eutrophication of seas, plastic pollution, novel entities in the environment and, of course, new diseases crossing over from animals to humans. Research into these potentially existential threats has generated intense interest in thinking about the future.

Future studies have expanded, especially as many of these threats and trends have a sensational presentation – *the end of the world*. 'Sustainability' became the watchword of the first phase of social and environmental concerns but then 'resilience' joined in. Surviving indefinitely was not enough – 'meeting the needs of the present without compromising the ability of future generations to meet their own needs' (UN General Assembly, 1987: para 27) – is not sufficient to prepare for the shocks and sudden changes that we might create to our fragile systems. It became important to scan the horizon, have foresight and invest in future studies. But there are consequences, so: what are the ethical considerations around future studies? And what are the ethics of research that shape the thoughts and opinions about the future?

This enquiry was triggered by a request to draw up a policy for working with people on deep questions about the survival of humanity and the fear that the consequences of this work might lead to extreme pessimism, anxiety and suicidal thoughts. How do we ethically conduct action research generating ideas about futures that might push some people into severe depression and existential questioning? This work also exposed another concern that asking how humanity will survive the next decade might lead some people to conclude that a violent response is the only answer. The urgency of the problem might stimulate some to think that we cannot wait for normal, slow political processes but instead we must force the issue. For those organising this work, non-violent methods are a basic tenet, so this was a significant wake-up call.

Future studies

Appropriately labelling the different studies of the future is a problem. Italian artists of the Futurism movement in the early 20th century glorified modernity and war – aiming to liberate Italy from the weight of its past, endowing the term 'futurism' with a dictatorial meaning. Other terms used for these studies are futures research, futuristics, futures thinking, futuring and futurology. Foresight (the ability to predict what will happen or be needed in the future) became a popular word in policymaking areas as a description of working with the future. The UK government established a set of foresight projects (Foresight, 2019) to give evidence to create policies that are more resilient to the future. The term *future studies* is used in an academic context, and the term *foresight* in a practical context when applying future studies.

Future studies appear to be dominated by the needs of businesses to understand market trends and shifting fashions, and all governments have foresight units that are horizon scanning, predicting the future and looking out for changes, political potholes and stumbling blocks. Globally, think tanks are involved in foresight in one way or another – looking out for security issues and geopolitical stresses or predicting financial markets or seeking new technologies. Many of these efforts are large scale and involve considerable investment. The World Economic Forum has recently launched a digital 'strategic intelligence' service to play with around 200 interacting megatrends.[1]

Economics contains a large component of forecasting, comprising statistical projections and business cycle analysis. Projections of trends in finance and economic activity are a key part of the 'dismal science', being major factors in government decision-making. The expectations of borrowing requirements, inflation, tax receipts, employment and, of course, public opinion are all about looking to the future. These actions are driven largely by the underlying desire to control the direction of travel, controlling the minds of people in the present without letting them know all that is going on.

Studying the future affords power and control, or influence over others. This requires decisions about how to use that power and that leads us into the world of morality and ethics. What is the right way to use this power? Since the future is uncertain, we do not have assured knowledge: we only have plausible stories to which we might attach probability estimates, and from among the plausible stories we decide the one story that we prefer and promote that to our advantage over others. Any dilemmas that arise in choosing a course of action must be expressed with the caution that the future will unfold in its own way and in its own time; it will emerge from the complexity of our systems and what we base our decisions on may turn out to be substantially different from what we proclaim.

Societal collapse and apocalypse

> The truth is that beyond our opinions, the facts matter. And the fact is that we live in an age of a choice of ruin. Climate change. Inequality. Fascism. Extremism. Mass extinction. Countries fracturing. Societies breaking down. None of these things are up for debate – and the question isn't even whether you're deeply and profoundly concerned by them – it's why you *aren't*.
>
> Umair Haque in Eudaimonia

Umair Haque's (2022) quote offers an example of how messages like this are broadcast; strictly, it was a narrow cast as this article first appeared in 2021 on Medium, which is an online blog channel, but the readership is substantial and consists of the 'informed public'.[2] It was submitted into popular culture, into a channel that has an air of authority and believability and where research work often gets interpreted for mass consumption. The blog piece illustrates that some people are certain about where the world is going and are keen to persuade others to their point of view: that the world is facing catastrophe. The apocalyptic narrative is powerful and seductive, a useful tool for those who wish to change the world. It is the consequences of their interpretations that concerns us.

'Our current global system is on the brink of collapse', began an email I received, not from some frantic activist but from the Head of US Operations, Sustainable Business, Reuters Events. Obviously, it was a hyperbole to get my attention because if they genuinely believed it then what are they doing just organising events for businesses to have a get-together. Apocalypse and collapse are now common currency and their meanings have shifted – and debased.

Is society about to collapse or is it already collapsing, and we are running in mid-air over the cliff edge with seconds before we plunge into the depths?

The considerable literature on collapse probably begins with Edward Gibbon's *The Decline and Fall of the Roman Empire* (Gibbon, 1776–89). Written in the latter part of 18th century, it traces more than 1,500 years of history and places the reason for that collapse on barbarian invasions and the slow erosion of civic virtue and ethics among its citizens – Rome became decadent. In his *A Study of History* (1934–39), Arnold Toynbee considers the growth and decline of civilisations, and their disintegration. He points out the evidence for this happening many times – the cyclical nature of expansion and contraction of civilisations. Joseph Tainter published *The Collapse of Complex Societies* (1988), in which he applies systems and complexity sciences to explain how collapse happens – societies often failing, in some sense, under their own weight. In *Collapse: How Societies Choose to Fail or Survive* (2005) Jared Diamond uses the evidence from previous civilisations to compare and extrapolate to the issues we face today.

More recent 'apocalyptic' books take on more of this tone, towards the complete global collapse of civilisation. The main driver of this is seen as climate change. In *The Uninhabitable Earth* (2019) David Wallace-Wells describes his frightening story of the future – the fly leaf of the book starts off: 'It is worse, much worse, than you think'.

There is something alluring about societal collapse, it makes great Hollywood movies, and the adrenaline flows while we scare ourselves. The dystopian, post-apocalyptic genre is full of bad news – from the earliest H.G. Wells's *The Time Machine* and *The War of the Worlds*, through *Mad Max*, *Blade Runner*, *The Hunger Games*, *Interstellar*, *The Handmaid's Tale*, *Black Mirror* and many more. Societies collapse in these movies, and we are entertained to examine what life might be like if circumstances suddenly change and we are thrown out of our comfort zone to scratch an existence, which is 'worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short' (Hobbes, 1651: Part 1, Ch 13). We should remember Orson Welles's radio production of *The War of the Worlds*, which many people believed. There were rumours of suicide – none were confirmed, thank goodness. It is interesting to note that far-right reading material has always used collapse, as in apocalyptic visions, which tend nicely to the ideology of a rebirth of a 'pure' order – under their control of course.

The concept of society collapsing has spawned a new discipline of collapsology. This neologism is gradually gaining usage as the study of how societies and civilisations come to an end (and potentially transform). It primarily explores issues like climate change as caused by human economic and geopolitical systems. It is not in line with the idea of a cosmic, apocalyptic 'end of the world', but makes the hypothesis of the end of the current world, how systems degrade, fail and disintegrate, how they lose values and often revert to earlier forms.

In a different context, we have seen the rise of mass movements such as Extinction Rebellion and Deep Adaptation Forum following decades of doom predictions about the limits to growth and the headlong rush of humanity past the boundaries of the planet.

At one end of a collapse spectrum, people are concerned about the future, they wish to be prepared for opportunity or disaster and recovery (they see a graceful degradation within business as usual); in the middle ground we have those who are certain of collapse of some kind and wish to cope with what that means, in a sense of rebirth; at the far end we have the die-hards who will survive an apocalyptic collapse in their redoubts until it's all over ready to emerge into a barren, purged world or else they will somehow flee to another planet to find a brand new barren world.

Limits to growth and research into possible collapse

In 1972 the Club of Rome (a small international group of influential people from academia, civil society, diplomacy and industry) published the results of a study modelling the world's population, food production, industrialisation, pollution and consumption of non-renewable natural resources. The model predicted that on the present course the limits to growth on earth would become evident by 2072, leading to 'sudden and uncontrollable decline in both population and industrial capacity' (Meadows et al, 1972: 23). This meant societal or civilisation collapse was inevitable without a substantial change in consumption and pollution.

In 2004, an update was published called *Limits to Growth: The 30-Year Update*. The authors observed that:

> It is a sad fact that humanity has largely squandered the past 30 years in futile debates and well-intentioned, but halfhearted, responses to the global ecological challenge. We do not have another 30 years to dither. Much will have to change if the ongoing overshoot is not

> to be followed by collapse during the twenty-first century. (Meadows et al, 2004: xvi)

In 2009, a group of Earth system scientists produce a framework called Planetary Boundaries that defines a 'safe operating space for humanity'. The framework is based on scientific evidence that human actions since the industrial revolution have become the main driver of global environmental change. They identified nine 'planetary life support systems' essential for human survival and evaluated how far these systems have been pushed already. They estimated how much further humans can go before planetary habitability is threatened. Their conclusion was that 'transgressing one or more planetary boundaries may be deleterious or even catastrophic due to the risk of crossing thresholds that will trigger non-linear, abrupt environmental change within continental-scale to planetary-scale systems' (Rockström et al, 2009).

The planetary boundaries mark the safe zone for the planet. As of 2009, two boundaries had already been crossed, while others were in imminent danger of being breached and there are two that we do not even know how to measure.

Climate change modelling is happening on an extensive scale by many teams across the world. The International Panel on Climate Change (IPCC) was created to provide policymakers with regular scientific assessments on climate change, its implications and potential future risks. Without a massive response to reduce greenhouse gas emissions, a significant warming of the atmosphere will happen with catastrophic consequences: it will be a global civilisation collapse.

Extinction Rebellion

Extinction Rebellion (XR) was established in the UK in May 2018. It is a global environmental movement with the aim of using non-violent, civil disobedience to compel government action to avoid tipping points in the climate system, biodiversity loss, and the risk of social and ecological collapse. Although criticised for its lack of diversity (mainly white and middle class) and even considered a threat alongside terrorist groups, it is a grassroots response to a sense (a paradigm) that we are heading towards extinction as a species.

The three demands of government that XR make, are quite simple to state: Tell the Truth, Act Now, Go Beyond Politics. The first of these: 'Government must tell the truth by declaring a climate and ecological emergency, working with other institutions to communicate the urgency for change',[3] contains a presumption that there is a truth out there, and that sufficient research evidence supports the prediction of grave danger and collapse unless action is taken. A second presumption is that government is not telling this truth and by denying it they are behaving unethically.

Deep Adaptation Forum (DAF)

The Deep Adaptation Forum was established in 2018 following the publication of a paper by Professor Jem Bendell. It explores how we prepare for and live with the idea of societal disruption and collapse. It assumes that we do not have sufficient resilience in our systems to cope with the changes arising from hitting the limits to growth and the effects of climate change and biodiversity loss.

Societal collapse means 'the uneven ending to our current means of sustenance, shelter, security, pleasure, identity and meaning' (Bendell, 2019).

People who consider that societal collapse is either inevitable, likely, or already unfolding, are using the term 'deep adaptation' to explore responses. This is a way of framing the current global situation that might refocus on what is important in life as collapse proceeds: 'we are finding new ways of being with ourselves and being together, no matter what happens' (Deep Adaptation Forum, 2022).

DAF encourages people to explore the emotional implications of collapse before looking outwards and acting in the world. They offer support to those who turn inward in grief at the coming loss. This movement is treading into an area of substantial, ethical debate – it is aware that some people may find this disturbing. The acceptance of the inevitability of collapse may be a tipping point for some people – but which way will they tip?

Prepper movement, delinkers and escape to Mars

> The world is reeling from multiple ongoing calamities, and only a tiny fraction has the luxury of retreating from it all. Those who lack that luxury are not going to be exactly happy about it. One way or the other, you will eventually have to pay for retreating from the world.
>
> Venkatesh Rao on Ribbonfarm

Survivalists have long considered the collapse of society as inevitable, but they are determined to survive it. Survivalism emerged in the 1980s and is now often called the Prepper movement. They are people who prepare for emergencies, including natural disasters, as well as disruptions to the social, political or economic order. The idea is to construct hardened shelters, stocked with food, goods and equipment (guns) to ride out the collapse. They are ready to defend themselves.

This drive for self-sufficiency is also a feature of some intentional communities, or eco-communities. These are planned to have a high degree of social cohesion and teamwork. The members of an intentional community typically hold a common social, political, religious or spiritual vision and often follow an alternative lifestyle. This is an expression of what is called 'delinking to relink later'. In a sense it means to set up a separate, independent community (delinking), to wait out a collapse in society, and then emerge to offer the world a new, better way without all the bad things that created the troubles we had (relinking).

For the wealthy, survivalism means escape to somewhere that will not be affected and where life can be lived out isolated from the troubles of Earth. There are stories of armed, gated compounds – robot-served, self-sufficient redoubts in places like New Zealand, sufficiently far away and difficult for the hordes to reach. And some rich people are even thinking of escaping to Mars to ensure the survival of humanity.

Existential risks

The drivers of societal collapse considered so far (in *Limits to Growth*, for example) tend to be of the stress type, slowly building problems, rather than sudden shocks. There are many institutions that conduct research into shocks, some of which could be catastrophic and world-ending (even in minutes). These existential level events are the plots for many disaster movies, in which we tend to survive by last-minute heroism or sacrifice. Research into these possibilities is included in this ethical examination just as much as any collapse by systemic failure or outgrowth of resources. The difference is in the 'act of god' nature – they are not human-induced shocks. They come with no blame attached: if an asteroid has Earth's name on it then there is no hope, if the Yellowstone National Park blows up as a super volcano, then it blows. Researching these possible natural events either induces anxiety or can reassure if the probability of occurrence is extremely low.

The Centre for the Study of Existential Risk[4] (CSER) at Cambridge University has a mission to evaluate low-probability/high-impact events. Their aim is to encourage responsible innovation among technologists and a safety culture among scientists. Many of the possibilities that are studied are of the type where humanity outgrows the resources of the planet, or we have damaged and polluted so much that ecosystems collapse, or we have altered the climate. However, there is another type of possibility that could lead to collapse and that is deliberate action such as war or terrorism. Collapse could occur through nuclear warfare, something we have lived with for over 70 years, but added to that danger are such actions as engineered pandemics, cyberwarfare, rogue artificial intelligence, geoengineering, genetically new life forms and other nasties.

There are three categories of collapse activators: (1) acts of God, (2) emergent issues from the collective activity of billions of people and (3) deliberately

Table 13.1: Risk assessment for societal collapse

	Low disruption	High disruption
High chance	Collapse seen as a naturally occurring cycle, and possibly beneficial for society to adjust to a resource constrained world.	This is the story of those who believe the collapse is inevitable – the response is to build the fortress or work out how to cope and adapt.
Slight chance	Rely on existing systems to adapt if the worse were to occur. No action needs to be taken just 'business as usual', with some prudence.	This is the problematic space. It is the asteroid impact, or malicious global scale terrorism. How much effort should go into preparing?

created issues – possibly accidental but also possibly engineered as a mechanism of war or conquest. Research action is different for each category – it will be some mix of assessing the dangers and warning, mitigating the danger, adapting to changes, discovering weakness and vulnerabilities, and preventing the implementation of harmful actions. But there is also considerable research that is about the world as it is rather than as it might be. Finding new knowledge about the beingness of things is also knowledge of becomingness (the two main ideas in ontology). Is it possible to distinguish ethical concerns about research into what things are from what they might generate? How can we do that when we do not know and cannot know the consequences?

In the world of resilience thinking we often consider the spectrum of chance and the spectrum of disruption. The business world evaluates its operations and plans by using a risk assessment. It imagines as many things as possible that may go wrong or negatively impact an organisation or a project and considers the probability of these happening and the scale of the impact. We can do this for the concept of societal collapse.

In each of the quadrants in Table 13.1, the effect on society and the individual varies from just taking precautions and building resilience to escaping or resignation to fate. Research activities continue to explore these spaces to understand the mechanisms and drivers, looking for causes, to identify ways to mitigate the issues, to improve resilience and the ability to recover, to improve our measure of the probabilities and the scale of the effects – the scale of climate change damage for example. It is fertile ground for researchers and is material for future studies and foresight to build scenarios of the outcomes – some of these are quite scary.

Power and control

> And always keep a-hold of Nurse
> For fear of finding something worse.
>
> Hilaire Belloc, 'Jim' (1907)

> The whole aim of practical politics is to keep the populace alarmed (and hence clamorous to be led to safety) by menacing it with an endless series of hobgoblins, all of them imaginary.
>
> H.L. Mencken, *In Defense of Women* (1918)

Using fear as a mechanism for power and control has a long history, predating humankind. Any social animal lives in a hierarchy and dominance is often fought over, to the death sometimes. More likely, though, is the fear of death.

The use of fear is subtle and pervasive, it affords power and control over others and is one tool in the box for political campaigning aided often by a press who are keen for any sensational story. It may lead to discrimination, marginalisation and the establishment of the 'other' – those to be eliminated or targeted for blame. Power and control through fear play to the ideologies of competition (for scarce resources) and conflict. And what better ways to create fear than invent scenarios of doom and gloom from which, of course, the politicians will rescue you, if you vote for them. Future studies and foresight provide the material for these scenarios and when backed up by reputable research, the fear is cemented.

The concept of what is called the future funnel or cone is a simple device to demonstrate how possibilities expand as you look further into the future; the imagined scenarios take on wider and wider ranges of possibility. The funnel, though, is often a way to demonstrate preferred scenarios or outcomes. However, the question must be asked: whose preferred outcome is it, and whose perspective is used to start looking into the future?

Disruption is also a tool of those seeking power and control. Naomi Klein's book *The Shock Doctrine* (2007) suggests that a deliberate strategy of inducing shocks is part of neoliberal free-market practices. The fear of future shocks is even cleverer in that it is all in the imagination. If research from a reputable source lends weight to these fears, then all the better for shaping attitudes and behaviour.

The dilemmas of future studies

> Science is properly more scrupulous than dogma. Dogma gives a charter to mistake, but the very breath of science is a contest with mistake and must keep the conscience alive.
>
> George Eliot, *Middlemarch* (1871)

Much of ethical thinking is about dilemmas and paradoxes. These arise through the superposition of two (or more) possible futures that are equally desirable (or undesirable). These futures determine action in the present, often a choice must be made between which path to tread, and both routes contain hazards and evils that are incommensurate. There is no set of scales that can indicate which is the better, and who should determine what 'better' means.

Another superposition occurs when we imagine a distant future that could trigger more than one possible course of action in the near future. If we imagine a societal collapse due to, say, global warming sometime towards the end of the 21st century, we may induce, as we speak today, a spirit of resistance and forthright action or we may trigger a fatalistic slide into depression, despair and lethargy. If we are pressing hard the sense of inevitability of societal collapse what is our ethical and social responsibility? What are the dangers? What could go wrong?

One of the first things to recognise is the paradox of future thinking. Once we have thought about the future, we make changes now, we prepare for what we imagine, which will produce a different imagined future, which we then prepare for and so on. This is a form of Zeno's paradox where we enter an infinite regression of future thinking. What are some of the ethical questions and dilemmas of which we should be aware?

Self-fulfilling prophecy

By constantly playing the same song, people come to believe you and their actions are influenced by this belief that what you project/envision is true and that their actions then produce that future or at least sway things towards it. There is also the sense that a prediction can lead to inaction that again contributes to the prediction being realised. If a significant amount of research points towards catastrophic climate change, then the temptation is to abandon attempts to reduce carbon emission as being futile and therefore catastrophe happens, when perhaps concerted action could have reduced the impacts. Imagination prompts research and the imagined is realised.

The Cassandra Matrix

When poor Cassandra, the daughter of King Priam and Queen Hecuba, did not return his love, the god Apollo placed a curse on her so that no one would ever believe her true predictions. Many people who spend their lives exploring environmental issues feel like a modern-day Cassandra. They predict catastrophic global warming or the end of civilisation as we run out of energy, but nobody seems to believe them – if they do then they find a way to pretend it is not true. If Cassandra told the truth but nobody would believe her then an anti-Cassandra is someone who always tells falsehoods, but everyone does believe them. Unfortunately, as the last few years have demonstrated, we now cynically consider most politicians

Table 13.2: The Cassandra Matrix

	Believe	**Disbelieve**
True	Where we want to be!	Cassandra: prophesying climate change disasters
False	Anti-Cassandra: politicians	Spam-scams

to be anti-Cassandras, we believe them usually because they are telling us what we want to hear.

If we consider a Cassandra Matrix of telling the truth or not against whether we are believed or not, then it would look like Table 13.2.
The bottom right-hand quadrant is occupied by those strange people who tell blatant falsehoods, and nobody believes them, like the classical email scams promising bank accounts bulging with lost millions. Well, we say no one, but some do.

We want to occupy the top left-hand quadrant – where the truth is told, and everyone believes it. As a scientist though, this is a dangerous place. In this matrix, we are applying classical logic where things are either true or false when, in practice, we should use fuzzy logic where things are to a degree true and false at the same time. We live in a world where most 'facts' and 'ideas' are something like 90 per cent right and 10 per cent wrong – but sometimes the other way around. We are continually exploring the world of uncertainty and experimenting with ways to cope. We are trying, through cycles of critical learning, to move towards that top left-hand quadrant while realising that we can never get there. Indeed, it would be unwise and dangerous to assume we can have scientifically proven facts; we must learn to live with just sufficiently high degrees of probability if we wish to make decisions and act. This makes us vulnerable to those opposed to change who wield the stick of certainty.

Cassandra's dilemma arises when she asks herself the self-fulfilling question: what if my telling the story makes it come true? She can predict the future, but should she just keep quiet because otherwise people will not believe her and so not do anything to change?

Crying wolf

This phenomenon is a combination of self-fulfilling prophecy and an exercise in how to become a Cassandra by giving false alarms repetitively so that people will always reject your predictions and calls for action even when your alarm call eventually happens to be true.In the Cassandra matrix the dynamics are that you start out in saying falsehoods that are believed but people get weary, so you move to saying falsehoods that are disbelieved but one day you say truth, but people are now fixed in disbelief of whatever you say – and the worst happens.

Parzival's dilemma

> Robert Frost wrote about 'two roads diverg[ing] in a wood' and taking 'the one less travelled.' But, in Molly's case, both roads continued to equally devastating destinations, even if the specifics were different. Which of the two paths would you choose if one went off a cliff and the other into quicksand?
>
> Ray Smith, *The Magnolia that Bloomed Unseen* (2017)

The complication in this quotation from Ray Smith lies in the assumption that we know the outcome in advance that both routes end in devastation. In future studies we can only imagine what the outcomes are – we can never be certain, and we can never measure anything to use as a choice-making instrument. This is at the heart of the ethical dilemmas in future studies.

> *The stakes were too high for error. If we are damned for our actions but don't know our action's result, then how dare we act? And yet, how dare we refrain?* (Kleiner, 1996)

The story of Parzival is told well by Art Kleiner in the *Age of the Heretics* (Kleiner, 1996). Parzival was a young knight of King Arthur's round table faced by a dilemma at a turning point in his life. He was brought up and trained to conform with a belief that he should refrain and hold back, it was improper to ask questions, and yet when his injured king could only be healed by the spontaneous act of a knight's noble heart, Parzival is locked in the dilemma and does nothing – and he is damned.

The dilemma for future studies is similar. It is a double bind. To rephrase: if we are damned for our prophecies but do not know the effects of our prophecies then how dare we prophesise? And yet, how dare we keep silent?

Precautionary principle

One way out is to adopt the precautionary principle. The precautionary principle is a form of the classical Hippocratic oath, simplified to the idea of 'first, do no harm'. If a course of action may lead to harm to humans or the environment, then we should take precautionary measures even if we do not know the full consequences. The proponent of any action should bear the responsibility to provide the proof of no harm.

The precautionary principle came to prominence in the environmental movement during the Earth Summit in 1992. Principle 15 of the Rio Declaration says: 'In order to protect the environment, the precautionary approach shall be widely applied by States. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation'.[5]

It is a negative approach that avoids everything not known to be safe. But should we retreat into ultra-safety and miss out on learning, sometimes, from mistakes? Better safe than sorry, as a principle, can fail us – by avoiding any potential damage, we fail to learn: a seemingly precautionary action can do more harm than good by freezing the state we are in, forbidding innovations and allows those with the worst imaginations to prevail.

In future studies the precautionary principle can become the target. Effort is placed on imagining possible damage, and research is directed towards this, with the aim of forestalling action. An example of this in climate science is geoengineering; any idea to directly affect the climate by, for example, solar radiation management, or seeding the atmosphere with chemicals, is consider a high-risk strategy. The uncertainty around the full consequences of disrupting natural systems is the focus of research. The opposite is to throw caution to the wind in the claim that we are addressing problems that are not being solved by direct means (by reducing carbon emissions) so we must adopt innovative ways – otherwise we will slide into crippling climate change and the potential collapse of societies.

The ideology of violation or security through obscurity

Should we tell?

Our third category of societal collapse activators (deliberately created issues – possibly accidental but also possibly deliberate, such as cyberwarfare) opens another horizon for dilemma. If, in our studies of the future and research, we discover or imagine and articulate a possible way to attack society, to disable vital systems, should we keep quiet or should we broadcast the idea?

The ideology of violation is a term devised by Jaron Lanier to describe the belief that discovering, and making public, ways to attack society will make society safer (Lanier, 2010). He proposes that we should indeed tell, but more than that, we should deliberately research into all possible ways to damage, disable or destroy our systems – we should violate what we consider holy to expose its vulnerabilities. The example he uses (2010: 65) is of a research team's efforts to use mobile phone technology to hack into a heart pacemaker to turn it off by remote control, to murder someone. Having discovered a way to do this, should it become public knowledge, or should this knowledge be controlled? A similar example of this dilemma is about publishing the 3D printing program to make a gun, having proved the point that this could be done – should we tell?

Those who disagree with the ideology of violation are said to subscribe to a 'fallacious' idea known as 'security through obscurity'. Society is safer if we do not tell and safer still if we do not imagine these possibilities and we do not conduct research into them.

What are the ethics of research to find vulnerabilities? Is uncovering vulnerabilities doing more harm than good? This is at the heart of resilience thinking. The concept of horizon scanning, and scenario development, is to dream up possibilities and to extrapolate weak signals to push the boundaries of possibility so that, through imagination, society can be better prepared for change than it is now.

Some arguments against the ideology of violation are:

- The diversion of effort: we are wasting resources looking for more ways that society might collapse when we should be researching into avoidance mechanisms for those concerns that we are sure about.
- Creation of division: having discovered or imagined a threat, we need new laws and controls, and these can lead to discrimination, marginalisation, conflict and polarisation in society.
- The arms race of looking for vulnerabilities: a technical term for this is 'schismogenesis' where our published discovery incites others to look for the vulnerabilities in our discovery, to look for new vulnerabilities, deeper and subtler in a never-ending battle of hacker versus protector.

This last argument comes from the world of software and system protection from viruses and their ilk. However, contrast this looking for holes in software with the aim of improving the system against looking for, or creating, novel lifeforms or diseases just so we can understand what some malicious agency might do, what they might create and how we might counter them or detect them doing it. In our pursuit of knowledge of what others might maliciously do, we may discover some new and highly dangerous knowledge or even technologies. If these get out – then what?

Adopt conspiracy?

Should we lie? The researcher and futurist can affect change by what they say. If they believe strongly enough that their work points to collapse and disintegration of society and that telling people so will make it happen or accelerate the process, then they could manipulate the message. Is it right to bend the truth or conspire to lose certain parts of the story for fear of alarm and shock? Or should the principles of honesty and truthfulness always prevail otherwise we descend into an assault on integrity and democracy, we enter the world of fake news, social media manipulations, constant lying and loss of trust.

From a deontological ethical perspective, we have this from Immanuel Kant, writing in 1797:

> Hence a lie defined merely as an intentionally untruthful declaration to another man does not require the additional condition that it must do harm to another … For a lie always harms another; if not some human being, then it nevertheless does harm to humanity in general, since it vitiates the very source of right. … All practical principles of right must contain rigorous truth. … This is because such exceptions would destroy the universality on account of which alone, they bear the name of principles. (Kant, 1898)

A consequential ethical position might argue differently, especially given the enormity of what societal collapse means.

Optimism and the Evel Knievel dilemma

Should the tone for future studies be pessimism (stop doing and retreat) or optimism (keep doing because we must get through this, we will get through this)?

In the face of societal collapse, and so that we can build resilience into our systems, there is a strong argument for an optimistic framework, otherwise we fall into the self-fulfilment of a negative story where we are looking to escape rather than looking for solutions. There is a growing school of positivity, seen in recent books by authors such as Stephen Pinker, Hans Rosling, Matt Ridley and Rutger Bregman. Their argument is that we are far better off materially and more equal and longer-lived than in any previous generation. This is due to trade, science and innovations that always seem to arise to overcome obstacles. Everything that seems to be driving us to the edge of the cliff is also making things better. Can that 'better' save us?

> Optimism is not about blindly ignoring the realities that surround us, that's foolishness. It is also not a naive faith that everything will take care of itself, even if we do nothing. That is irresponsibility. The optimism I am speaking of is not the result of an achievement, it is the necessary input to meeting a challenge. It is, in fact, the only way to increase our chance of success. (Figueres, 2020)

Perhaps we are facing what is called the Evel Knievel dilemma. Robert Craig Knievel (born 1938) professionally known as Evel Knievel, was an American stunt performer famous for ramp-to-ramp motorcycle jumps. One of his stunts was an attempt to jump the Snake River Canyon on a rocket-powered motorcycle. Imagine if he were halfway across and the thought occurred to him: 'this is polluting the atmosphere, I should stop this' and so he turns off the rocket engine. The analogy is that our drive for knowledge and innovation, although polluting as we go on, will get us to the other side where we have sufficient sophistication and technical abilities to build our utopia; but if we stop now, we will fail to construct what is necessary to survive and instead plunge into societal collapse.

Future generations

We have a duty to future generations that exist now or are most likely to exist – those born in, say, the next few decades. Do we have a duty to more distant, unborn, future generations? A belief in the inevitability of the collapse of civilisation would preclude concern for future generations in the sense of those who might have been born if collapse did not occur. The Deep Adaptation Forum concern is a compassionate one for those future generations that will be born before and during the collapse.

Thinking about 'future generations' stimulates thinking about intergenerational equity. The sustainability and climate action movements have long argued for considering future generations to embed the principles of long-term thinking into law. The 1997 UNESCO Declaration on the Responsibilities of the Present Generations Towards Future Generations is an international agreement which includes provisions related to the common heritage of humankind. Article 4 states:

> The present generations have the responsibility to bequeath to future generations an Earth which will not one day be irreversibly damaged by human activity. Each generation inheriting the Earth temporarily should take care to use natural resources reasonably and ensure that life is not prejudiced by harmful modifications of the ecosystems and that scientific and technological progress in all fields does not harm life on Earth.[6]

This declaration places a significant responsibility on researchers in future studies to fully explore and imagine the possibilities that might lead to harm to future generations. Thinking about the yet-to-be-born, the future of humanity and our purpose on this planet creates an immense context for ethical thinking. When we consider who is affected, who is neglected and who is rejected we must expand the framing of questions to include those who are yet to be conceived as far out as we possibly can.

The effect on the individual

> The brains of humans … contain a mechanism that is designed to give priority to bad news.
>
> Daniel Kahneman, *Thinking, Fast and Slow* (2012)

The World Future Society recently asked their members a question: how does a futurist mindset help us get through the day? It requires time to reflect on this negative question, especially on those deep existential enquiries that are with us always, never resolved and never clearly formulated – who am I, why am I here, why is anything here? We see our life as following one path, the only path we believe. We cut a singular track, a channel – canalised – the further we travel the harder it gets to escape; the more we want to escape the faster we travel. A reinforcing feedback that ends in a scream of despair.

The presumption in the question is that by thinking about the future we can somehow set a wider and longer context for our lives and perhaps see that we are in a rut. This view beyond and from above might offer ways to escape and reset our lives or at least direct ourselves away from the channel of hopelessness. It sounds appealing, it may work for some, not for others, and in certain cases it could lead the other way. Like the unknowns around the springtime suicide peak, there is a fear that encouraging a futurist mindset may just be one more unknown as to possible remedies for the trials of getting through the day and through one's life. It could be a powerful tool in some cases by removing the blinkers and the vision of just the ground in front – doing this with others by having future conversations is vital, doing it alone is nigh on impossible.

Should we have conversations that look forward and explore the future, develop a sense of 'wanting to know what happens next' – emotional messages to support rational arguments? It will not necessarily work and it may make things worse, but that is the condition with all attempts to help with existential crises.

There is no knowing. A future conversation framed by optimism, with a sense of conquering problems, can work for some; but for others it may push them further into depression if it reinforces their isolation from those who are building a bright future. Apocalyptic thinking shuts down initiatives and reduces the ability to think beyond the narrow horizon of horrific problems when what is needed is a trust that there are alternatives, nothing is certain and good outcomes can occur. Without some positive drivers we are condemned to a life of doom scrolling (searching ceaselessly on the social media sites for bad news coming towards us) – it is a downward spiral of negative thoughts leading to depression, anxiety, distress, suicidal thoughts, hopelessness.

For the individual, thinking about the future can bring great hope by transcending the noise of the bad news, driven by sensationalist media, to see the game from the outside rather than being caught up in a crossfire of fear and despair. It is the role of future studies and foresight (and the research that drives these) to take up the responsibility to present this game. The influence on the collective consciousness of future thinking could easily push societies down a pathway to Balkanisation, to ghettos, lack of action to mitigate – instead of stimulating action to adapt and protect, we withdraw as a group to erect the fortress. We polarise our communities as we protect 'us' by rejecting the 'other'.

Research, knowledge and future studies

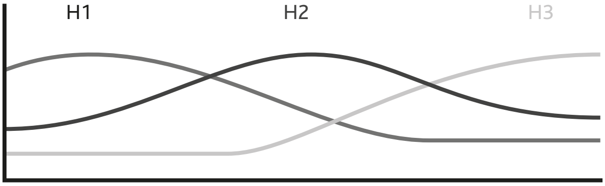
> Bad news sells because the amygdala is always looking for something to fear.
>
> Peter H. Diamandis and Steven Kotler, *Abundance: The Future Is Better Than You Think* (2012)

Many research projects produce results that are used as evidence to support prediction and to suggest change based on those predictions. The need for future projections can drive the research agenda. There are consequences, and ethical questions arise at many different levels. In the field of resilience, we need to shift our thinking from stasis to rhesis. We need to replace the idea of returning to a prior condition after a shock (homeostasis) with the idea of returning to the same processes or function (homeorhesis). The COVID-19 pandemic illustrates this in that health systems will not return to exactly what they were before but instead return to providing the services of healthcare – its function – with all kinds of new procedures and practices from what was learned in this awful period. The overall process of healthcare will return to what it was; the organisation and structures will be different.

When addressing societal challenges or issues of resilience, researchers need to think not only about problems – the knowledge of what the issues are – but also knowledge about how to overcome those problems. The concept of 'three types of knowledge' is used to formulate research questions and develop action plans. The concept first appeared in the 1990s and is explained by Buser and Schneider (2021). This has developed into a core underpinning of transdisciplinary research:

1. Systems knowledge, which is usually defined as knowledge about the current system or problem situation. It is propositional – analytical and descriptive – but also includes experiential and practical knowledge from our history.
2. Target knowledge, which is knowledge about the desired future and the values that determine the directions for action; this is now more presentational knowledge gained through describing possibilities and preferences – it talks about us, researchers, as well, in reflection.
3. Transformational knowledge, which is the process of change from the current to the desired situation. This is knowledge gained from experience and practice.

Figure 13.1: The three horizons model



Source: International Futures Forum (www.iffpraxis.com) and H3Uni (H3Uni.org)

The ethics we are a concerned with are mostly in the target and transformational types of knowledge. It is about the rights and wrongs of what we are aiming to create and how we go about creating.

The three horizons model developed by Bill Sharpe (2013) delves deeper into the transformation process and can be applied to the issues we are looking at – such as societal collapse (see Figure 13.1). Where we are today is described as horizon one (H1); it is the prevalent system, sometimes called 'business as usual'. It contains many serious problems and, as shown by exercises like the limits to growth, it is seemingly destined to fail. But it is where we live, and we need to keep it going. In contrast, horizon three (H3) is the imagined, target future – the one that we should desire, our preferred future out of many possibilities.

Horizon two (H2) is the transformation process, the disruptions that are necessary to shift systems that often have a huge momentum – business as usual is like a supertanker. H2 contains enabling mechanisms, innovations that allows H3 to emerge, these are called H2+ (plus). However, it can also contain blocking and stalling processes that try to prevent change and maintain the status quo; these are called H2− (minus) processes.

The third horizon is the long-term successor to business as usual, the danger that we are considering is that H2 activities cause a collapse of H1 with no chance of reaching H3. Perhaps we should call this horizon zero (H0).

'Three horizons' is an extremely rich model onto which we can layer ethical thinking about research work. Working with just an H1 perspective is about how to improve systems, in the existing paradigm of continuity and gradual change 'it is about fixing the failing system, innovating to maintain it, "keeping the lights on"' (Sharp, 2013). Research that extrapolates from this present and points out the failings and the potential for collapse highlight the possibilities of H0 – the obvious example is carbon emissions and consequent severe climate change leading to global societal breakdown.

Researcher as expert

The naive approach to research into the complex global issues – that might lead to collapse or generate action to improve resilience – would assume that clearly presented evidence, and its projection into the future, by itself would persuade people to act and make appropriate changes to forestall events or to adapt to them. That might work if the evidence fits existing world views but for many people it does not and, even more, the changes that are suggested as necessary are certainly not desirable. The evidence is rejected, and great efforts are made to rationalise why the evidence is wrong. Logic and rationality are not the most persuasive arguments. Exhorting people to 'listen to the science' is often counterproductive. Logos, persuading using reasoning, is limited to a conducive audience.

Complex global issues are transdisciplinary; they manifest as emergent problems created by the interconnection of many factors. It is impossible to simplify and present an argument like 'this cause produced that effect'. Reducing complexity in this way does a disservice – it denies the complexity and is open to shifting simplifications as fashions change. People feel unsure what 'science' really means, especially as science is built on a framework of contingent knowledge, it proclaims uncertainty and openness to refutation. Yet we all seek certainty and believe what science says and then lose faith when the ground shifts and something different is announced.

Pushing this 'listen to the science' message is not only counterproductive; it also opens the door to what is called 'SONKing' (the scientification of non-knowledge), where a simple claim about some complicated policy issues or moral question is made to appear scientific just to borrow the gravitas to make a point.

Science describes what the situation is and what might happen … but it still does not tell us what we should do. It is our values and what we care about that tell us what we should do. Ethos, persuading using ethics, is limited to those who care about values, especially honesty, truth and integrity. Science cannot be untangled from culture and politics, and it is directed towards benefit to society. It is not disinterested; it does not sit apart from the world. It comes as a surprise to some people when they discover that scientists are not detached but hoping to improve the world with their work.

Society does not benefit when people with questions about science get shut down too quickly. People have genuine concerns, and the most important thing is to hear them out. Scepticism over science increases whenever people feel they have been shut down when they tried to express a concern or have their say. Listening and finding common ground is a way to reduce polarisation of opinions. Pathos, persuading using emotion, is limited to those who have an emotional response – those with passion are more likely to understand if they are listened to and they see passion in science.

Research ethics

The UK's Economic and Social Research Council describes research ethics as: 'the moral principles and actions guiding and shaping research from its inception through to completion, the dissemination of findings and the archiving, future use, sharing and linking of data'.[7]

There is a long history of thought about ethics in research work of all kinds. The naive view of research as an activity conducted by men (sic) in white coats in a laboratory with test tubes and Bunsen burners barely suggests that we should ask what are they doing, why are they doing it, who is affected, is it good, fair, transparent? Yet these questions come into sharp focus when research is about people, nature and the environment – that is, when it is about us.

In the social sciences, ethics in research originated in 'protect the person' issues in medical practice but expanded to the benefits, risks and harms to all stakeholders affected by research and then to the social and environmental responsibilities of researchers. One approach to ethical thinking about conducting research on these issues is to consider seven categories:

1. Respect for persons – protecting those with diminished autonomy
2. Informed consent – openness
3. Confidentiality and data protection – personal security
4. Conflict of interest – no exploitation
5. Justice – fair and unbiased behaviour
6. Integrity and trust – a sense of reliability
7. Beneficence and non-maleficence – providing benefit

The first four of these are inner principles and are about how to research with people who are involved or directly affected by research activity. There is pressure to expand those affected beyond people to include non-human actors, animals obviously but also parts of nature such as river catchments, forests, marine reserves and more – to treat these as 'persons' with rights both moral and legal (de Toledo, 2020). The attempt to create a crime of ecocide is active and growing.[8]

The three last categories on the list are about doing good or at least not doing harm. These are outer principles directed at people generally and towards non-human actors and nature. This is where the distinction between ethics and morality gets blurred.

This virtue ethics approach, of not doing harm as a character trait, regardless of whether you might harm humans, animals or the environment, is encouraged by the concepts of justice and integrity and of course on adhering to the first four principles of correct professional behaviour.

The seventh principle is what is called the intersubjective community assessment of what is good, right or just for all. This morality principle covers societal norms about right and wrong. Applied to research this directs us to consider the outcomes and the consequences of our activity: the use of the knowledge created. Future studies add in the complications that so much is uncertain – there are multiple pathways forward all with different consequences.

Setting the research agenda

Research is not context free. It feeds into ideologies, and ideologies feed the agenda of those with power (money), and then research feeds off those with money. It is a self-reinforcing loop, that generates confirmation bias and the echo chambers of research. Where in this loop does responsibility lie? Where does ethics play a role?

The common perception of research is that:

- researchers are morally responsible for the consequences of the application of the knowledge or inventions that they produce; and
- researchers, individually and collectively, have a greater responsibility than other people for the use of knowledge – because they should have a better understanding of possible consequences.

But these responsibilities cannot all be laid at the doorstep of the researcher – they may not be able to appreciate what the consequences might be. In some areas this is clearly the case: dealing with complex environments, for example, is beyond anyone's ability to see the whole and, indeed, complexity implies the emergence of phenomena that no one could ever foresee. Researchers cannot predict exactly how their work may be abused or misused both in the near term and the distant future. This is especially true if the abuse arises from the combination of knowledge that produces something unimaginable when the research was started. For instance, laser and drone technologies both have benign uses, but put them together and frightening weapons emerge.

Another issue is that research is fragmented, and responsibilities are spread across partners and even countries – no one can control the application of research output. Everyone is nominally responsible, but no one person can be held responsible. The responsibility is institutionalised.

There is also a gradation of responsibility in research from those areas that are removed from day-to-day concerns – for example, fundamental physics – to those where the nature of the research is directly about constructing our world and the objectives are well known. These are mission-oriented areas like environment sciences, pharmaceuticals or defence. The application of the knowledge is understood and is thought through in advance – responsibility is surfaced.

The clients of research cannot avoid responsibility – whether that is in the private sector or government, by commissioning research and providing the facilities to conduct it there is complicity and researchers are obliged to follow and possibly subjugate any ethical concerns for fear of being dismissed or ostracised.

Types of ethics

When we refer to 'benefits, risks and harms' it suggests a 'consequentialist' form of moral reasoning where the consequences of research define ethical practice. For future studies this is doubly problematic. First, we cannot know exactly what the consequences might be in the present conditions. This is the predicament of all science and research, but we can explore the possibilities with the knowledge of our current systems. Second, for the future, the systems we are considering have not yet emerged within the complex interactions that will happen to create this future.

What other ethics can we apply to this concept of research feeding future studies? We have already discussed 'virtue ethics' as the original form for research with the perspective of concern with all aspects of a researcher's professional behaviour, emphasising a need for rigour in ethics reasoning about the circumstances and phases of the research work as it interacts with an external world. Virtue ethics is about character and reputation, which are vital for researchers to maintain.

We have also touched upon deontological ethics – doing what is right no matter the consequences. These approaches suggest that guided and rule-governed research practice can serve to fulfil moral obligations. In this framing, research that affects future studies can be considered right even if the imagined consequences are bad, as long we follow rules or moral law based on our values and principles – integrity, truth, justice.

Research that feeds future thinking can have both intrinsic value and extrinsic consequences, suggesting a combination of deontological and consequential ethics. The first may lead to great improvements that could go so far as mitigating the collapse of society, so preventing harm to future generations. The second would expose the potential harm to people's psychology both for the individual and for mass effects, creating group hysteria or activist movements like XR and so on. The consequences might be a mind shift to modes of fatalism and defeatism.

Is there a dilemma – which ethical frame do we use?

To resolve this perhaps we can follow Habermas and the concepts of discourse ethics. In this the issues and interests are discussed (freely, without any coercion) so that understanding is widespread and out of this emerges an agreement. There is a strong movement towards large group processes to determine courses of action such as citizen assemblies – randomly chosen people are presented with the issue and after rounds of discussion they produce recommendations. This is a democratic ethic; is it possible that future studies could operate in this environment of scrutiny?

Perhaps the concept of pragmatic ethics is a development of this in that moral direction and decisions – on what future studies should do – evolves like scientific knowledge: it is a social activity over the course of generations. In the same way that science proceeds through hypothesis and conjectures that are challenged and refuted only to be replace by new theories or paradigms, our ethical position on future studies could shift and change as future generations refine the ideas and replace them because of enquiry and discourse.

Another stream of ethical thinking, which is apparent in movements like Deep Adaptation Forum, is that of care ethics. As a branch of feminist thought, moral values emerge that stress empathy and compassion. An increasing theme in future studies is the need to cope with grief as we recognise the destruction of the natural world. The loss of biodiversity is a one-way street.

Future studies require a plurality of ethical thinking, there is no one way to determine what is right to research. There are many dangers as we see from considering the ideology of violation, the more we can freely talk about the issues the greater the chance that we decide how to proceed with research – this is a form of intuitive ethics, we just get better at knowing what is good or bad.

> Modern man … talks of a battle with nature, forgetting that, if he won the battle, he would find himself on the losing side.
>
> E.F. Schumacher, *Small Is Beautiful: A Study of Economics as if People Mattered* (1973)

In the same way that we are toying with endowing nature with legal rights, and the ideas of earth jurisprudence develops our understanding of the relevance of governance beyond humanity to the whole earth community, perhaps we can apply contractarian ethics? Is it time to produce a social contract (or should we say a peace treaty) where instead of the state we work with the higher authority of the whole earth? The Earth Charter movement attempts this: 'The EC is an ethical foundation for actions to build a more just, sustainable, and peaceful global society in the 21st century. It articulates a mindset of global interdependence and shared responsibility. It offers a vision of hope and a call to action'.[9]

Ethics in reporting on research

> 'The British nation is unique in this respect: they are the only people who like to be told how bad things are, who like to be told the worst.'
>
> Winston S. Churchill[10]

A discussion of ethics, research and futures studies cannot ignore the way that we communicate and the many different forms that takes. The media mediates between the expert and public. The dissemination of research takes a variety of forms. Professional reporting goes through the formal mechanism of peer review and publication in distinguished journals that is designed to maintain standards of procedure and practice. The Popperian concept of conjectures and refutations explains how knowledge grows through an unending process of trial and error. Virtue ethics is concerned with this rigour over the behaviour of the researcher. The consequences of the content of the research are included only in the way that projections are communicated – if this (our research) is true then that (the outcome) is likely. The outcome in question is seldom expanded to embrace the full consequences for society; these are left as implied or not relevant to the question in hand.

More populist reporting lies in the sensational articles in the press and, more likely, in social media feeds. Their focus is often on the dire consequences and the extrapolation of dangers and potential catastrophes although they are keen to broadcast any sensational breakthroughs and major advances. However, generating fear is good for business. This feeds into the discourse on a collapsing society, the deterioration of circumstances today as compared to yesteryear. 'Headlines, in a way, are what mislead you because bad news is a headline, and gradual improvement is not' (Bill Gates, quoted in Green, 2014: 32). Where we see a connection between research and future studies is in the way that research is reported in the media. The story of climate change has unfolded over the last few decades as research evidence accumulates to lend weight to the arguments that human activity is causing a dangerous increase in the greenhouse gas effect through emissions into the atmosphere of carbon dioxide, methane, nitrous oxide, water vapour, chlorofluorocarbons and ozone. The imagined consequences are volatile weather, ocean acidification, the destruction of ecosystems, sea level rise, melting glaciers and so on. All bad news that sells papers. The drip feed of imagined bad futures produces a constant state of anxiety about any future and a state of low-key existential crisis and depression. And the consequences of that are then expressed as more bad news – surveys on expectations for your children's future, generated by our bad news, becomes more bad news for us to report.

There are, of course, many channels of communication and there are serious attempts to create ones that explain research to the public in a balanced, professional and ethical way. Consider *The Conversation*, which is an independent source of news, analysis and expert opinion, written by academics and researchers, and delivered directly to the public. They have published an editorial charter, which includes these statements that are relevant to our discussion:

- Inform public debate with knowledge-based journalism that is responsible, ethical and supported by evidence.

- Unlock the knowledge of researchers and academics to provide the public with clarity and insight into society's biggest problems.
- Provide a fact-based and editorially independent forum, free of commercial or political bias.
- Ensure quality, diverse and intelligible content reaches the widest possible audience by employing experienced editors to curate the site.
- Set the standard in journalism best practice. Be open, transparent and accountable. Where errors occur correct them expeditiously.
- Work with our academic, business and government partners and our advisory board to ensure we are operating for the public good.[11]

These are all laudable aims (and they publish further important aims not shown here[12]); however, there is no definition of what the public good might be.

'The medium is the message' – what would Marshall McLuhan have made of social media? For McLuhan, it was the medium that controlled 'the scale and form of human association and action' (McLuhan, [1964] 2001: 7, 9). Could he have foreseen how vast and powerful that medium has become?

Conclusion

In thousands of years' time, if an alien people ask, 'how did the first, truly global, human civilisation collapse?', will the answer be because they talked themselves into it, because they researched their way into it, because they could not imagine an alternative to collapse?

Researchers and policymakers must reflect carefully on such large questions, and consider the many ethical issues that can arise during research that investigates them. In summary the ethical concerns that we have covered are:

- working with people on existential questions might trigger depression and suicide;
- the magnitude of the problems and the lack of political effort to address them may lead some towards political violence;
- studying the future affords power and control, or influence over others;
- whose perspectives and preferrable futures dominate – who gets marginalised?
- self-fulfilling prophesies – our foresight generates the future;
- we are damned if we do and damned if we do not;
- should we research into how vulnerable our systems are, or should we stay away?
- should we conspire to affect the future?

- being prepared is an ethical response to the stories of the future;
- should ethical concerns drive research;
- stirring up fears produces a reaction to deny they exist, which entrenches the behaviours that cause the fear. The reaction is stronger than the action;
- what ethic framing should we adopt for future studies? and
- how should we communicate ideas about the future?

There are many more deep imponderables about how we conduct future studies with the aim of building resilient communities and civilisation. Researchers and policymakers have an ethical responsibility to grapple with these challenging issues. If we believe that societal collapse is inevitable, what are the consequences for future studies, what duty do we have as researchers and to whom do we owe that duty – do we abandon future generations and focus on those alive or about to be born today? Does this line of thinking generate a new form of doomsayer that channels our compassion and humanity towards coping with the end – death counselling applied to society?

If instead we only believe societal collapse is possible but not inevitable then what are our duties and does preventing collapse take ethical priority over other concerns because the consequences are so great? If extinction is on the horizon, should we not rebel to chase it away? And research activity that is happening in thousands of institutions across the world carries on in its own compartments, feeding new knowledge and new hypotheses to those who are imaging the future – and feeding the ever-hungry media that demands sensation – true or not, it matters little.

Perhaps the role for research is to promote an 'ideology of concern', an ethic of care and respect that cherishes humanity as part of nature; one that is based on global optimism that a societal collapse will not happen because we are dedicated to finding ways to prevent it.

Notes

[1] See: https://www.weforum.org/strategic-intelligence

[2] The article was moved from Medium by the author; since January 2022 it has been hosted on Eudaimonia and Co, and its new details are given in the reference.

[3] See: https://extinctionrebellion.uk/the-truth/demands/

[4] See: https://www.cser.ac.uk

[5] See: https://www.jus.uio.no/lm/environmental.development.rio.declaration.1992/15.html

[6] See portal.unesco.org/en/ev.php-URL_ID=13178&URL_DO=DO_TOPIC&URL_SECTION=201.html

[7] See: https://www.ukri.org/councils/esrc/guidance-for-applicants/research-ethics-guidance/risk-and-benefit/

[8] See: https://www.stopecocide.earth/

[9] See: https://earthcharter.org/about-the-earth-charter/

[10] From a speech in the House of Commons, 10 June 1941.

[11] https://theconversation.com/uk/charter

[12] See: https://theconversation.com/uk/10-ways-we-are-different

References

Bendell, J. (2018) 'Deep adaptation: a map for navigating climate tragedy', Institute for Leadership and Sustainability (IFLAS) Occasional Paper no 2, Ambleside: University of Cumbria, available from: https://insight.cumbria.ac.uk/id/eprint/4166/ [accessed 9 February 2022].

Bendell, J. (2019) 'The love in Deep Adaptation – a philosophy for the forum', *Professor Jem Bendell* [blog], 17 March, available from: https://jembendell.com/2019/03/17/the-love-in-deep-adaptation-a-philosophy-for-the-forum/ [accessed 9 February 2022].

Buser, T. and Schneider, F. (2021) 'Three types of knowledge', *Integration and Implementation Insights* [blog], 11 February, available from: https://i2insights.org/2021/02/11/three-types-of-knowledge/ [accessed 8 February 2022].

Deep Adaptation Forum. (2022) 'Deep adaptation', available from: https://www.deepadaptation.info/about/ [accessed 10 March 2022].

de Toledo, N. (2020) 'What if nature became a legal person?', 19 May, World Economic Forum, available from: https://www.weforum.org/agenda/2020/05/nature-legal-personhood/ [accessed 10 February 2022].

Diamandis, P.H. and Kotler, S. (2012) *Abundance: The Future Is Better Than You Think*, New York: Free Press.

Diamond, J. (2005) *Collapse: How Societies Choose to Fail or Survive*, London: Penguin.

Eliot, G. (1861) *Silas Marner*, Edinburgh: William Blackwood.

Eliot, G. (1871) *Middlemarch*, Edinburgh: William Blackwood.

Figueres, C. (2020) 'The case for stubborn optimism on climate', October, available from: https://www.ted.com/talks/christiana_figueres_the_case_for_stubborn_optimism_on_climate/transcript?language=en [accessed 8 February 2022].

Foresight (2019) 'Foresight projects', GOV.UK, updated 10 November 2021, available from: https://www.gov.uk/government/collections/foresight-projects [accessed 8 February 2022].

Gibbon, E. (1776–89) *The Decline and Fall of the Roman Empire*, 6 vols, London: Strahan and Cadell.

Green, S. (2014) *Inside the Mind of Bill Gates*, Scotts Valley, CA: CreateSpace.

Haque, U. (2022) 'Why we need to fix the world', Eudaimonia and Co, 30 January, available from: https://eand.co/why-we-need-to-fix-the-world-3bc1c30f6eb4 [accessed 8 February 2022].

Hobbes, T. (1651) *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiasticall and Civil, Part 1*, London: Andrew Crooke.

Kahneman, D. (2012) *Thinking, Fast and Slow*, London: Penguin.

Kant, I. (1898) 'On a Supposed Right to Tell Lies from Benevolent Motives', In: T.K. Abbott (trans.), *Kant's Critique of Practical Reason and Other Works on the Theory of Ethics*. London: Longmans, Green and Co.

Klein, N. (2007) *The Shock Doctrine: The Rise of Disaster Capitalism*, London: Penguin.

Kleiner, A. (1996) *The Age of Heretics: Heroes, Outlaws and the Forerunners of Corporate Change*, London: Nicholas Brealey.

Kranzberg, M. (1986) 'Technology and history: "Kranzberg's laws"'. *Technology and Culture*, 27(3): 544–60.

Lanier, J. (2010) *You Are Not a Gadget: A Manifesto*, London: Penguin.

Lovelock, J. (2006) *The Revenge of Gaia: Why the Earth is Fighting Back – and How We Can Still Save Humanity*, London: Penguin.

Malthus, T. (1798) *An Essay on the Principle of Population, as It Affects the Future Improvement of Society*, London: Johnson.

McLuhan, M. ([1964] 2001) *Understanding Media: The Extensions of Man*, Abingdon: Routledge.

Meadows, D., Randers, J. and Meadows, D. (2004) *Limits to Growth: The 30-Year Update*, White River Junction, VT: Chelsea Green.

Meadows, D.H., Meadows, D.L., Randers, J. and Behrens, W.W. III (1972) *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*, New York: Universe Books.

Mencken, H.L. (1918) *In Defense of Women*, New York: Philip Goodman.

Rao, V. (2021) 'MJD 59,256', *Ribbonfarm* [blog], 11 February, available from: https://www.ribbonfarm.com/2021/02/11/mjd-59256/ [accessed 9 February 2022].

Meadows, D., Randers, J. and Meadows, L. (2004). *Limits to Growth: The 30 Year Update*. London: Routledge.

Rockström, J., Steffen, W. et al (2009) 'Planetary boundaries: exploring the safe operating space for humanity', *Ecology and Society*, 14(2): art 32, available from: https://www.ecologyandsociety.org/vol14/iss2/art32/ [accessed 8 February 2022].

Schumacher, E.F. (1973) *Small Is Beautiful: A Study of Economics as if People Mattered*, London: Blond and Briggs.

Sharpe, W. (2013) *Three Horizons: The Patterning of Hope*, Axminster: Triarchy Press.

Smith, R. (2017) *The Magnolia that Bloomed Unseen*, independently published.

Tainter, J.A. (1988) *The Collapse of Complex Societies*, Cambridge: Cambridge University Press.

Toynbee, A. (1934–39) *A Study of History*, Oxford: Oxford University Press.

UN (United Nations) (1992) *United Nations Framework Convention on Climate Change*, Bonn: United Nations.

UN (United Nations) General Assembly (1987) *Report of the World Commission on Environment and Development: Our Common Future*, 20 March, available from: https://digitallibrary.un.org/record/139811?ln=en [accessed 8 February 2022].

Wallace-Wells, D. (2019) *The Uninhabitable Earth: Life After Warming*, London: Penguin.

14

Research relationships and positionality: adult–child power relations

Katharina Miller and Nina J. Zugic

Introduction

This chapter addresses researcher positionality within the field of childhood studies in social sciences and humanities targeted by the European Commission Horizon 2020 Programme project. Path2Integrity is an educational project, funded under the Horizon 2020 *Science with and for Society* (SwafS) programme.

Path2Integrity's overall objective is on research integrity and education. The project has developed formal and informal pathways for teaching research integrity in a learner-centred manner. Both the formal and informal pathways are evaluated on an ongoing basis throughout the 36-month duration of the project. Consequently, this chapter's focus is on the formal and informal pathways evaluation, which includes the main ethical issues encountered. In particular, the Path2Integrity project aims are: (1) to equip teachers in high schools and educators in universities with cutting-edge, user-friendly tools to teach research integrity in a learner-centred manner, and (2) to raise awareness about the importance of research integrity in high schools, academia and in society (Inguaggiato and Lindemann, 2021).

An essential part of the Path2Integrity project is the evaluation of the training programme and of the campaign material used in informal learning exercises. Following the project's grant agreement guidelines, all researchers involved in the evaluation have been clearly informed that 'the evaluation will be designed as a formative mixed method evaluation combining qualitative (semi-structured interview, feedback sheets) and quantitative (structured questionnaire) elements to monitor the complete development and improvement process of the Path2Integrity 'Handbook of Instructions for learning pathways' (Prieß-Buchheit et al, 2020a: 4). Based on the reliable, objective and valid evaluation (Prieß-Buchheit et al, 2020b), the evaluation's main focus was on the short-term efficiency of the innovative methods and learning pathways in comparison to teacher-centred methods. The evaluation design was a quasi-experimental, non-randomised trial with participants and 'non-participants' or rather participants taught with teacher-centred methods.

Path2Integrity research includes over 300 participants, aged between 16 and 18, who are in a full-time schooling. Due to the involvement of participants who are aged 17 years and younger, one of the main ethics requirements to the Path2Integrity Consortium from the independent ethics experts within the European Commission (EC) ethics appraisal scheme, was to ensure the opinions/approvals by ethics committees and/or competent authorities for the research with humans (ethics approvals from research ethics committees (RECs) or institutional review boards (IRBs)). This basic requirement, turned out to be a big challenge for the Path2Integrity Consortium. The guidelines of the research institutions of the Path2Integrity researchers are very clear about when opinions/approvals by ethics committees have to be provided for. Prior to starting the research, each researcher was instructed to check in advance whether their research subject has fulfilled the applicable standards and whether any risks are foreseeable. Therefore, it was clear that an initial ethics assessment must have been conducted by the researcher themself. Should potential risks be identified in the risk assessment, a formal research ethics procedure will be conducted by the project's ethics committee.

In acknowledging the gap between the research ethics requirement from the EC, on the one hand, and the institutional guidelines of the Path2Integrity partners, which was responsible for the carrying out of the evaluation, on the other hand, the authors of this chapter (and also a partner in the Path2Integrity Consortium) understood the urgent need to specifically focus on the legal and educational aspects of research with children and young people in social sciences and humanities. By clearly addressing legal and educational aspects of research within the field of childhood studies in social sciences and humanities, the authors strongly believe that this chapter offers a critical perspective on researcher positionality.

Where a researcher involved in research with children and young people is solely responsible for ethical considerations without institutional support, the importance of guiding interactions and relations of power between adult researchers and younger participants, including ways to guide the researchers' actions at both national and international levels, is extremely compelling. The authors note that the researchers' power over the children and young people during the actual research potentially could lead to false or even misleading research outcomes and results, where a researcher, as the adult in the situation, is either unaware or does not recognise their own positionality prior to starting the research. By adopting a critical approach to 'researcher positionality', this chapter provides useful recommendations aiming at contributing to a better understanding of policymaking processes and, crucially, the relationships between adult researchers and children involved in the research. Hence, we drive to a conclusion offering some recommendations to policymakers and research institutions that can also support and inspire researchers to remain on an ethical track while conducting research with children and young people.

The three adult–child power relations

This chapter is structured according to main themes in Eija Sevón's (2015) research on power relations between adults and young children. Sevón clearly distinguishes three different kinds of adult–child power relations:

1. The power of parental authority, which requires children's obedience.
2. Negotiated power and agreements.
3. The power of resistance.

Sevón demonstrates the 'thin line' between researchers' responsibility while performing research with children and young people. Following Sevón's main distinctive research findings, the authors have divided this chapter along the three main points:

1. Ethics and positionality, including the new research terms in the adult–child power relations.
2. Ethics, positionality and the law, where the authors discuss how the adult–child power can be negotiated and possible requirements for such an agreement.
3. Ethics, positionality and education, as well as their historical connections.

Ethics and positionality in research with children

The research within Path2Integrity includes researchers working with over 300 young people aged between 16 and 18 years, who are in full-time schooling. The research itself consists of evaluations of the participants' experience in receiving the training programmes on research integrity and, also, of the campaign material used for informal learning. The young research participants were all asked to fill out a survey with various questions, both before and after the training programmes, and before and after seeing the campaign material on informal learning material. The survey is anonymous and takes each participant approximately 15–30 minutes to complete.

Consequently, the question of positionality has been raised here. According to Coghlan and Brydon-Miller (2014), positionality refers to the stance or positioning of the researcher in relation to the social and political context of the study – the community, the organisation or the participant group. The position adopted by a researcher affects every phase of the research process, from the way the question or problem is initially constructed and designed to how it is conducted. Hence, positionality is the social and political context that creates an identity in terms of race, class, gender, sexuality, age/generation and ability status. In addition, positionality also describes how identities influence, and create potential biases to, an understanding of and outlook on the world. This means that positionality influences how we come to see and understand the social world.

Perhaps the most challenging aspects were the ambition of the researchers to collect information on the gender of the participants and to be as inclusive as possible. Thus, the Path2Integrity researchers aimed to include these data. However, this needed to be consistent, and in accordance with the data protection policy within the Path2Integrity project. Hence, five gender options were offered to the participants to choose from: (1) Male, (2) Female, (3) Genderqueer, (4) Do not want to answer, or (5) Other (with a possibility to explain). Furthermore, the researchers made a significant effort to formulate their questions in an inclusive and intersectional way, so that all the children and young people could identify with these questions and to avoid any obvious biases or stereotypes.

On the other hand, it is worth mentioning that, for example, according to Shaw et al (2020), research is also influenced by institutional principles that shape the social and cultural context of research practice. These principles ensure that research participants feel safe, respected and listened to when they take part in research, and that the knowledge outcomes of the research process are disseminated and shared with relevant cultural groups. Shaw et al presented four case studies based on projects that discuss the researcher's positionality in relation to the ethical and emotional work involved in undertaking research on sensitive topics with individuals from vulnerable and marginal groups. In doing so, Shaw et al foreground the importance of articulating and managing emotion in research on sensitive topics, and suggest measures to ensure the well-being of researchers engaged in similar studies and related topics.

Within the Path2Integrity project, no sensitive topic in the sense of violence or abuse of children was involved in the research itself. The challenge of the research project was rather to reach out to as many different young people as possible and to be as inclusive as possible (such as reaching out to young people from economically disadvantaged areas). Being a European project, reaching out to as many participants as possible virtually, in particularly since the start and during the COVID-19 pandemic, and therefore making participants feel as data-protection safe as possible throughout the research, was a challenge. This was resolved, and successfully managed, by relying on, and using only, the contacts provided by the trusted sources within the Path2Integrity Consortium and its partners.

Last but not least, it is also worth mentioning that Pechurina (2014) discusses ethical decisions in the qualitative research of domestic sites, with particular focus on a situation, in which a researcher studies their own migrant community. In particular, in the article in which the researcher explores more general issues such as negotiating access and permission to photograph in participants' homes, the author discusses issues specific to community-based ethnographic studies among Russian migrants. Furthermore, the author discusses the importance of social positioning issues and the power distribution within the community under study, using examples from the study of Russian migrant housing in the UK. The complexity of the ethical decision-making at different stages of the research process, as well as the ever-changing relationship(s) between the cultural and social backgrounds, and the identities of researchers and participants, are reflected upon. The insider and outsider roles of the researcher are relative, with the constant need to balance them while creating difficult ethical dilemmas often bringing rich data to light, which in turn advances the whole research process.

In conclusion, the researchers within Path2Integrity have had significant difficulties in evaluating participants taking place in the training programmes on research integrity, in particular during the COVID-19 pandemic. Perhaps, one of the biggest ethical challenges during the Path2Integrity research was to actually encourage, but not obligate, participants to acknowledge the evaluation process as the part of the training, on which the research outcome heavily depends. Although challenging, this was successfully executed by clearly explaining in detail the importance of the evaluation to the research participants and the overall Path2Integrity research outcomes and lessons learned.

Ethics, positionality and the law

The relationship between ethics, positionality and the law at an international level remains complex. The experience of the authors of this chapter, along with some literature (Edwards et al, 2012), suggests that researchers in the realms of social sciences or humanities often cannot count on receiving the support of or approval by RECs and/or competent authorities (such as IRBs) for their research with humans. As described by Edwards et al (2012):

> Research ethics review practices vary considerably across countries and this variability poses a challenge for international research programmes. Although published guidelines exist, which describe underlying principles that should be considered and pragmatic approaches that could be followed in seeking ethics approval, most have roots in biomedical and clinical research. The result is that there is generally less clarity around institutional and/or country-level structures for ethics review of health policy and social sciences research. This is an

> important gap that needs to be addressed in order to ensure ethical practices in multi-country research programmes.

Although this was published in 2012, to the best knowledge of the authors, no changes have been made nor addressed since.

Another example of the importance of compliance with ethical standards can be seen in a chapter titled 'Research with child street traders: building trust and sensitivity', by Burch and Osaiyuwu (2021). The authors draw a rather important example from the experiences of children working as street traders in Nigeria, which is highly relevant to the overall importance of the research relationship and positionality. According to Burch and Osaiyuwu, some of the children who participated in the research were potentially vulnerable at several different levels. The children's work as street traders, for example, was extremely demanding, often causing conflict with their everyday education, entailing competing demands from school and family and, more importantly, operating around the margins of illegality. 'The research', as stated by Burch and Osaiyuwu 'with this group, thus, involved operating carefully and sensitively to establish trust across key processes' (2012: 122).

The Burch and Osaiyuwu case study clearly shows how the issues of ethics, positionality and law are dependent and how such a complex relationship is actually demonstrated in practice. Path2Integrity includes recourse to an international advisory board; one of its members is from Zambia and currently working in Ghana, where Path2Integrity also intended to conduct research with children and young people. That is why it is important for the Path2Integrity Consortium to take into consideration the complexity already outlined. Moreover, as argued by Parson (2019), there are also ethical concerns, and even potential methodological obstacles, when conducting similar research with under-represented, marginalised or minoritised groups, that is, research relationships and positionality are not exclusively present only in the adult–child power relations but also in an 'adult–adult' power relationship between the researcher and an adult participant of different ethnical, geographical and cultural background. At first glance, one might think that this is more about power relations than law. However, from the legal perspective, where there is a discrepancy in a power relationship, one could argue, in the last instance, that this imbalance could or even should be corrected by laws if there is no other way to correct the disproportion.

Therefore, it is fair to say that in such research, the actual 'power position', or to be more precise, positionality of the researcher, are of the same construct. In addition, also according to Parson, a researcher in most cases is a member of a 'dominant' research funded group and/or with a privileged outsider status, which can mean that a researcher is not necessarily aware of the potential obstacles and vulnerability of the research subjects. To overcome this imbalance in social sciences and humanities between the researchers' power relationship with children and/or young people, the activities of the researcher should be clarified prior to their contractual agreement between the researcher and their home institution and/or funders and, when necessary, the researchers' superior position of power towards their research subjects (especially in the case of children and/or young people) should be somehow restricted. This leads to the authors' first recommendation (see Recommendation 1).

From a legal point of view, research conducted in the realm of social sciences and humanities should be formalised by a set of guidelines according to existing legislation. Hence, such a document, which could benefit both policymakers and researchers when conducting the research with children and young people in social sciences and humanities, at both national and international level, should compiled in a 'Code of Research Conduct in Social Sciences and Humanities'. Thus far, and to the extensive knowledge of the authors, there is no specific regulation, nor instruction, on how to protect children and young people in a research activity in social sciences and humanities. The first international treaty on children's rights, the Convention on the Rights of the Child (UNCRC), was adopted in November 1989, entering into force on 2 September 1990 (OHCHR, 1989). The UNCRC has been ratified by all members of the UN with one exception: the US.

Article 32 of the UNCRC states that 'States Parties recognize the right of the child to be protected from economic exploitation and from performing any work that is likely to be hazardous or to interfere with the child's education, or to be harmful to the child's health or physical, mental, spiritual, moral or social development'. The remainder of Article 32, however, clarifies that this refers to formal (or informal) employment structures and not to work in a broader sense such as being part of a research project. Therefore, the UNCRC cannot oblige States Parties to approve legislation to protect children and young people in research in social sciences and humanities. As described with the Path2Integrity example and confirmed by Edwards et al (2012), some universities do not issue opinions/approvals by ethics committees and/or competent authorities for the research with humans (ethics approvals from IRBs) that aim to ensure research is not deemed harmful for children.

On the other hand, the European Union Agency for Fundamental Rights (FRA) provides regulations and guidelines regarding the participation of children in research within the European Union (EU), in particular concerning ethics approval and informed consent. The FRA (2014) notes that 'children and young people are increasingly involved in research and decision-making on different aspects of rights which relate to them'. The FRA itself does not suggest particular ethical principles. It only gives an overview of how the 27 EU Member States handle the context of legal framework and ethical codes of conduct and indicates a huge variety in how each state tackles these issues.

Among the EU states, in Bulgaria, Cyprus, Ireland and Portugal parental consent is mandatory for children under the age of 18 with respect to participating in research. Croatia, the Czech Republic and Poland, however, require parental consent only for children under the age of 14; whereas in Austria, Belgium, Denmark, Estonia, France, Italy, Romania, Slovakia and Spain there is no clear regulation about parental consent and age groups. The remaining Member States provide guidance on parental consent that varies depending on contexts. One could think that this is more appropriate rather than suggesting a blanket approach given the variability and diversity of childhood across cultures and nations. However, for researchers in social sciences and humanities, and especially researchers conducting research within more than one country, this variation can invoke many hurdles and might limit the researcher's willingness or abilities in doing a comparative or international research. In Germany, for example, 'there are large differences between [the 16] Federal States with some States requiring informed consent by parents up to the age of 18 and time-consuming procedures' (FRA, 2014). The lack of a homogenous regulation across just the EU makes it very difficult for researchers to comply with the existing legislations across the 27 EU Member States. This lack of homogeneity could become running the gauntlet especially for researchers in social sciences and humanities research fields, as well as medical research studies. For these reasons, we strongly recommend harmonising existing legislation, at least the existing legislation across the 27 EU Member States (see Recommendation 2).

Interestingly, the FRA does not provide any information as to the consent of children or young people in research. The EC requests from the researchers receiving funding within Horizon 2020 that the children and young people give their 'assent' to participating in research. Within Path2Integrity, this assent has been foreseen and integrated in the design of the research. It is important to note that the inclusion and respect of children and young people is also tackled in a guide issued by the UK's Nuffield Council on Bioethics (2015). Many international declarations, regulations and guidelines take different approaches to the extent to which children or young people should be involved in the decision to participate in a research project. Primarily, most regulations specify what information children and young people should receive. On one hand, it is necessary to involve children and young people in the assent process in a manner appropriate to their maturity. However, assent has many different meanings and implications. It can range from the nascent capacity to consent of a three-year-old to the knowing consent of young people. Our third recommendation is to harmonise this very important field of assent by children and young people and bring it into the centre of research with children and young people, in order to empower this group instead of seeing them simply as research subjects (Recommendation 3).

The European Commission (2019) advises researchers to involve children in research only if studies with consenting adults would not be effective. This does not mean research should be done with adults first. The intention of EC is rather to protect children and young adults and to fund and support researchers only in those cases that a researcher focuses on childhood studies. In the case of Path2Integrity, one of the requisites of the Horizon 2020 tender was to include young people as active participants in the research.

Furthermore, the EC requires that research participants be subject to only a minimal risk and burden. Finally, the results of all research should hold benefits for the individuals and groups represented by the participant(s). Relating to this last point, the EC tackles two aspects: first, for the EC, benefit-sharing is especially important if the research is conducted in low and/or lower-middle income countries according to the country classification by the World Bank. To underpin this requirement, useful documents to consult are, for example, the United Nations Environment Programme (UNEP, 1998), 'Synthesis of Case-Studies on Benefit Sharing', and the 'Global code of conduct for research in resource-poor settings' elaborated by the Horizon 2020 project TRUST (nd). Path2Integrity includes recourse to an international advisory board; one of its members is from Zambia and currently working in Ghana, where the Path2Integrity also intended to conduct research with children and young people. Both countries are classed by the World Bank (2021) as lower-middle income. None of documents mentioned, however, make a specific reference to children and young people in low- and/or lower-middle-income countries. It is also absolutely crucial that researchers comply especially with a specific reference to the children and young people in order to avoid a situation comparable to the 'forum shopping' in which litigants have their legal case heard in the court thought most likely to provide a favourable judgment. In the case of childhood research, this means discouraging researchers who find it too complicated, or simply cannot comply with the high hurdles of regulations and guidelines, from going to other countries where protection might not be as formalised as in many countries of the Global North. Therefore, our fourth recommendation is to include in the documents related to benefit-sharing a specific reference to children and young persons in these countries. The authors especially recommend making special reference to children and young people within the research projects financed by the EC and the United Nations Environment Programme (Recommendation 4).

Overall, research with children and young people should not be conducted only because a, or indeed any, research subject, is needed and/or because it has become fashionable to conduct research with children and young people. It has to be guaranteed that the research supports, helps and brings enlightenment to this potentially highly vulnerable group. As discussed, researchers, especially in social sciences and humanities, cannot always count on support by an institutional ethics review process that could help to prevent 'fashionable' research with children and young people. Therefore, the authors highly recommend to both the EC policymakers and research authorities at both national and international level that they include another requirement for more funding of research projects involving children and young people. This is Recommendation 5: any research needs to conscientiously avoid any stigmatisation, or indeed any type of negative influence on children or young people, while conducting research, and whenever possible, conduct the research that supports and encourages the insightfulness of children and young people.

Reflecting on Recommendation 5, we discuss further the three different approaches of the Society for Research in Child Development (SRCD, in the US), the National Institute for Health Research in the UK and National Association of Pediatric Nurse Practitioners in the US.

The SRCD (2007) has established 16 principles within their *Ethical Standards for Research with Children*. The first principle states that the researcher 'should use no research procedure that may harm the child either physically or psychologically'. The principles, however, do not establish an obligation to provide benefit sharing. Principle 5 states that:

> Incentives to participate in a research project must be fair and must not unduly exceed the range of incentives that the child normally experiences. Whatever incentives are used, the investigator should always keep in mind that the greater the possible effects of the investigation on the child, the greater is the obligation to protect the child's welfare and freedom.

It is clear that these two principles focus on the sanctions, 'just in case' of scientific or personal misconduct, which could, as a consequence, include the expulsion of the members. Having in mind researchers in social sciences and humanities, the intention of these principles does not primarily seem to have researchers focus on 'enlightenment' for children or young people.

The National Institute for Health Research (NIHR, 2016) provides tips and essential key issues for researchers when they involve children and young people in research. A benefit-sharing is foreseen by the NHS (Tip 2) and researchers are asked to show respect for the children's contributions, their inclusion and equality within the team (Tip 8). This factsheet also provides for more information on the reward, employment and benefit issues as well as on the ethics issues.

An important document was also published by the National Association of Pediatric Nurse Practitioners USA (NAPNAP, 2010), namely its 'Position

Statement on Protection of Children Involved in Research Studies', in which NAPNAP advocates for evidence-based practice in the clinical setting and recognises that continuing research, including research involving children, will be required to gather that evidence. NAPNAP advocates for 'direct benefit to the child' in case 'the research involves greater than minimal risk'.

The Path2Integrity researchers have consistently had the 'enlightenment' of young people as a lodestone. The design of the research has been done in such a way that young people would be able to acquire skills while participating in the research, to be better equipped as independent and critical thinking citizens and future researchers of this world. It is important to report that the Path2Integrity research subjects have not encountered any risks with respect to their own well-being.

Ethics, positionality and education

In planning or carrying out educational research involving children and young people, it is not unusual for the question to arise: is this research ethical? However, particularly in the social sciences and humanities, institutional review of the ethical aspects of a piece of research is often absent. Not surprisingly, this has been the case in the Path2Integrity project research execution.

The importance of ethical research has been raised by Edwards et al (2012) and by Breault (2017), who states that the ethical research claims in education and qualitative research is inherently relational. According to Breault, '*ethical research is shaped by the researcher and the researcher is shaped by it*'. Furthermore, qualitative research has made considerable progress in the past in perceiving and integrating researchers' positionality in relation to the power structures of race, gender, ethnicity and social class. However, less attention seems to have been paid to date as to how these positionalities have evolved and the role that sense and influence of place might play in researchers' approach, design, interpretation and use of research. This 'meta-synthesis' by Breault suggests a greater awareness and conscious consideration of place in the research, in order to increase the transparency of research, offer greater insight into research participants and generally increase the nuance and complexity of research findings.

With respect to the Path2Integrity project, due to the COVID-19 pandemic, the most challenging part of the research was the execution of the educational research integrity workshops. Path2Integrity includes training programmes designed as (external) workshops on research integrity and campaign material as an informal way of learning, which led to a number of challenges and problems, especially at the start of the COVID-19 pandemic and throughout the lockdowns. Consequently, the majority of the participants were challenged by the pandemic measures taken by their national governments, hence the execution of Path2Integrity and meeting the project's tight deadlines were at serious risk. It was down to the Path2Integrity partners' personal contact to 'convince' students to participate in the Path2Integrity training programmes on research integrity. Since then, the collaboration has been successfully reinstated and all deadlines met with minor timeline adjustments.

In addition, in the case of Path2Integrity, it was of great importance to find a workshop topic to engage the students, as an activity external and additional to their high schools' curricula. Before the COVID-19 pandemic, this connecting link was the citizenship education curriculum (Turner, 1999), which together with the SwafS programme has been instrumental in addressing the European societal challenges tackled by Horizon 2020 programme. This objective was clearly targeted as the Path2Integrity original project objective. This particular objective also played a crucial part while preparing the Path2Integrity project proposal due to the education for citizenship and the teaching of democracy already taught in the schools' curricula. In order to continue the research throughout the COVID-19 pandemic, the Path2Integrity researchers have changed the original workshop topics and have adapted and adopted the 'new emerging' topics, such as fake news and reliable science. These topics were welcomed by students, and were successfully connected to, and sometimes even embedded in, the schools' curricula. Such an effective and efficient approach to adapting the Path2Integrity objective did not just convince the schools' authorities, who let their students take part in the Path2Integrity research, but also made students enthusiastically participate in this educational research project, especially during the extremely difficult and challenging times throughout the COVID-19 pandemic.

According to the authors' findings to date, some ethical issues that may arise during an educational research activity such as Path2Integrity, are:

- How to be as inclusive as possible and also include participants from poor(er) neighbourhoods. It could possibly be that the Path2Integrity partners don't reach high school students from poor(er) neighbourhoods and that those do not see the necessity of participating in such activities.
- How to create both the training programmes on research integrity (plus the campaign material as informal learning material and the accompanying evaluation) with Gender Equality by Design. The topic of research integrity raises many questions in relation to gender equality and intersectionality. One question could be whether male and female researchers behave differently and have different attitudes regarding integrity issues (van den Hoven and Mol, 2021). This means that training programmes on research integrity have to include these many different aspects.

- How to get hold of the necessary completed evaluation sheets. Researchers and research integrity trainers cannot oblige the students to fill out the evaluation surveys. Therefore, the researchers have to find ways without coercing students to participate in the evaluation. One solution could be the detailed explanation why the participation in the evaluation is necessary and important for future research and policies.

Related to the Path2Integrity recent experience, it is also worth mentioning here that Dowling and Whiteman (2020) raised an extremely important point that focuses on an underexamined issue in the literature on educational research ethics. They ask whether an ethical authority can be established in educational research. They also address this from a perspective that disrupts naturalised approaches to ethical research, and argue that rather than seeking 'rights' or 'wrongs', researchers should always be tasked with constructing ethical stances, which is, in the light of the Path2Integrity experience, a point we would also like to raise.

Last but not least, it is also worth mentioning that within the education and learning context, positionality usually refers to the role of teachers and the goals, knowledge, strategies and other normative frames of reference as significant factors in the learning experience of students (Banks, 1995). Hence, teachers and researchers may come across the same ethical obstacles while either teaching or conducting research with children or young people, as clearly experienced while executing the Path2Integrity research. The intentions of the Path2Integrity training programmes on research integrity (together with the campaign material as informal learning material and the accompanying evaluation), are to promote (1) self-empowered thinking by students and (2) a dialogue on what is 'right' and 'wrong' in a research integrity situation. The intention of Path2Integrity researchers is to make both current and future researchers and citizens reflect on their own thinking and explain their arguments and perspectives. The Path2Integrity material gives the researchers, policymakers and citizens the tools to argue their views on these important statements.

Conclusion

In order to help researchers in the social sciences and humanities conduct research involving children and young people, harmonising legal research frameworks at national and international levels is recommended. Research must also be conducted in line with the conventions of the sociocultural context in which the study is being undertaken. A number of positionality issues have been identified especially within the research relationships and within adult–child power relations. The authors' major recommendations are:

Recommendation 1

Researchers, in general, appear to be members of a 'dominant group' and/or with 'privileged outsider status'. To overcome this issue, the imbalance in the power relationship with children and young people, the activities of researchers in human sciences should be clarified prior to any contractual agreement between the researcher and their home institution and/or funders, and when necessary, the researchers' position of superiority towards his or her research subjects (especially in the case of children and/or young people) should be somehow restricted.

Recommendation 2

The lack of a homogenous rules (not only) across EU could make it very difficult for researchers to comply with the existing legislations across the 27 EU Member States. This lack of homogeneity could, especially for researchers in social sciences and humanities research fields as well as medical research studies, become a running of the gauntlet. The authors strongly recommend that policymakers work to harmonise the existing legislations across the 27 EU Member States.

Recommendation 3

The authors' third recommendation is to harmonise the important field of obtaining consent (assent) from children and young people, and bring it into the centre of research with children and young people, in order to empower this group instead of seeing them as simple research subjects. This has already been argued comprehensively some years ago by the National Children's Bureau in the UK (McLaughlin, 2015). It's time to continue the lobbying at the EU level.

Recommendation 4

In the case of childhood researchers who find regulations and guidelines too complicated or simply cannot comply with the high hurdles regarding the participation of children in research, moving research to countries with lesser protections needs to be avoided. It is recommended to include in documents related to benefit-sharing, such as the UNEP's 'Synthesis of Case-Studies on Benefit Sharing' (UNEP, 1998) and TRUST's (nd) 'Global code of conduct for research in resource-poor settings' a specific reference to children and young people in these countries. The authors especially recommend making a special reference to children and young people within the research projects financed by the EC and UNEP. Policymakers are urged to take up this recommendation urgently.

Recommendation 5

It is suggested that EC policymakers and research authorities in general include the following requirement for funding research projects where children and young people are involved: the research must avoid any stigmatisation or indeed any type of negative influence on the children or young people while conducting research. In addition, if possible, research should be conducted that supports the 'enlightenment' of children and young people.

Recommendation 6

The authors' final recommendation is to encourage and enhance communication and collaboration on research and related policies that involve children and young people at all levels internationally and especially across the EU countries, the UK and the US, where quite clear sets of standards have already been established and maintained.

Acknowledgements

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824488. The authors want to especially thank the coordinator of this project, Prof. Dr Julia Priess-Buchheit, and researcher Nicolaus Wilder, for their very valuable input to this article.

References

Banks, J.A. (1995) 'Multicultural education and curriculum transformation', *Journal of Negro Education*, 64(4): 390–400.

Breault, R.A. (2017) 'Place and positionality in education research', paper presented at the annual meeting of the Midwest Educational Research Association, Evanston, IL, 10 October, available from: https://www.academia.edu/34933417/Place_and_Positionality_in_Education_Research [accessed 22 January 2021].

Burch, S. and Osaiyuwu, A.B. (2021) 'Research with child street traders: building trust and sensitivity', in G. Spencer (ed) *Ethics and Integrity in Research with Children and Young People* (Advances in Research Ethics and Integrity, Vol 7), Bingley: Emerald, pp 121–33.

Coghlan, D. and Brydon-Miller, M. (eds) (2014) 'Positionality', in *The SAGE Encyclopedia of Action Research, Volume 2*, London: Sage.

Dowling, P. and Whiteman, N. (2020) 'Authority and ethics: a case for estrangement in educational research and research education', *British Educational Research Journal*, 46(4): 770–85.

Edwards, N., Viehbeck, S., Hämäläinen, R.M., Rus, D., Skovgaard, T., van de Goor, I., Valente, A., Syed, A. and Aro, A.R. (2012) 'Challenges of ethical clearance in international health policy and social sciences research: experiences and recommendations from a multi-country research programme', *Public Health Reviews*, 34: art 11, available from: https://doi.org/10.1007/BF03391663 [accessed 10 February 2022].

European Commission (2019) 'Horizon 2020 programme: guidance – how to complete your ethics self-assessment, version 6.1', 4 February, available from: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf, p 9 [accessed 22 January 2021].

FRA (European Union Agency for Fundamental Rights) (2014) 'Child participation in research', available from: https://fra.europa.eu/en/publication/2014/child-participation-research [accessed 22 January 2021].

Inguaggiato, G. and Lindemann, L. (2021) 'Path2Integrity', The Embassy of Good Science, available from: https://embassy.science/wiki/Theme:1fc5c5b6-6c30-4400-a79b-8838b5a041cc [accessed 18 April 2021].

McLaughlin, H. (ed) (2015) *Involving Children and Young People in Policy, Practice and Research*, London: National Children's Bureau, available from: https://www.preparingforadulthood.org.uk/SiteAssets/Downloads/x21fm3yy636383196682430667.pdf [accessed 14 April 2021].

NAPNAP (National Association of Pediatric Nurse Practitioners USA) (2010) 'NAPNAP position statement on protection of children involved in research studies journal of pediatric health care', *Journal of Pediatric Health Care*, 24(1): A17–A18, available from: https://doi.org/10.1016/j.pedhc.2009.08.012 [accessed 22 January 2021].

NIHR (National Institute for Health Research) (2016) 'Involving children and young people in research: top tips and essential key issues for researchers', January, Eastleigh: INVOLVE, available from: http://www.invo.org.uk/wp-content/uploads/2016/01/involvingcyp-tips-keyissues-January2016.pdf [accessed 22 January 2021].

Nuffield Council on Bioethics (2015) *Children and Clinical Research: Ethical Issues*, London: Nuffield Council on Bioethics, May, available from: https://www.nuffieldbioethics.org/wp-content/uploads/Children-and-clinical-research-full-report.pdf [accessed 22 January 2021].

OHCHR (Office of the High Commissioner for Human Rights) (1989) *Convention on the Rights of the Child*, available from: https://www.ohchr.org/en/professionalinterest/pages/crc.aspx [accessed 22 January 2021].

Parson, L. (2019) 'Considering positionality: the ethics of conducting research with marginalized groups', in K. Strunk and L. Locke (eds) *Research Methods for Social Justice and Equity in Education*, Cham: Palgrave Macmillan, pp 15–32.

Pechurina, A. (2014) 'Positionality and ethics ien the qualitative research of migrants' homes', *Sociological Research Online*, 19(1): 111–20.

Prieß-Buchheit, J., Aro, A.R., Demirova, I., Lanzerath, D., Stoev, P. and Wilder, N. (2020a) 'Rotatory role-playing and role-models to enhance the research integrity culture', *Research Ideas and Outcomes*, 6: art e53921, available from: https://doi.org/10.3897/rio.6.e53921 [accessed 10 March 2021].

Prieß-Buchheit, J., Dwojak-Matras, A., Metodiev, T. and Miller, K. (2020b) *Path2Integrity Handbooks of Instructions: The P2ILC Programme*, Deliverable D3.2 EU Horizon 2020 Path2Integrity Project, grant agreement No 824488, available from: https://www.path2integrity.eu/documents [accessed 22 January 2021].

Sevón, E.M. (2015) 'Who's got the power? Young children's power and agency in the child-parent relationship', *International Journal of Child, Youth and Family Studies*, 6(4–1): 622–45.

Shaw, R.M., Howe, J., Beazer, J. and Carr, T. (2020) 'Ethics and positionality in qualitative research with vulnerable and marginal groups', *Qualitative Research*, 20(3): 277–93.

SRCD (Society for Research in Child Development) (2007) *Ethical Standards for Research with Children*, March, Washington, DC: SRCD, available from: https://web.archive.org/web/20201021102647/https://www.srcd.org/about-us/ethical-standards-research-children [accessed 22 January 2021].

The World Bank (2021) Convention on Bilogical Diversity. https://www.cbd.int/doc/meetings/cop/cop-04/information/cop-04-inf-07-en.pdf accsessed 22 January 2021.

TRUST (nd) 'Global code of conduct for research in resource-poor settings', Horizon 2020 Project, available from: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/coc_research-resource-poor-settings_en.pdf [accessed 22 January 2021].

Turner, P. (1999) 'Education for citizenship and the teaching of democracy in schools', *Teaching Public Administration*, 19(2): 46–9.

UNEP (United Nations Environment Programme) (1998) *Convention on Biological Diversity (UNEP/CBD/COP/4/Inf.7): Synthesis of Case-Studies on Benefit Sharing*, 4 May, available from: https://www.cbd.int/doc/meetings/cop/cop-04/information/cop-04-inf-07-en.pdf [accessed 22 January 2021].

Van den Hoven, M. and Mol, H. (2021) 'Do male and female researchers behave differently and have different attitudes regarding integrity issues? That is the question', *Op-ed, H2020 Integrity* [blog], 3 March, available from: http://h2020integrity.eu/male-female-researchers-behave-differently-different-attitudes-regarding-integrity/ [accessed 6 March 2021].

World Bank (2021) 'The world by income and region', available from: https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html [accessed 10 February 2022].

PART III

Case studies

15

What is good evidence?

Helen Kara

> A philosopher and a logician take their child on a train from London to Glasgow. As they cross the border into Scotland they see a brown cow standing alone in a field.
>
> 'Look,' says the child, 'the cows in Scotland are brown.'
>
> 'No,' says the philosopher, 'there are cows in Scotland, and at least one of them is brown.'
>
> 'That's not right either,' says the logician, 'there is at least one cow in Scotland, of which one side appears to be brown.'
>
> (Anonymous popular anecdote)

Introduction

The evidence before our own eyes can be the subject of debate: what is the evidence we are seeing, and how should we interpret that evidence? These questions lead to more questions: what counts as evidence, and when? Who decides? And what is 'good' evidence? The answers to these questions partly depend on the context in which they are being asked. Here I am considering evidence in the context of policymaking and that will shape my argument.

Policy can form a bridge between knowledge and action. It is almost unthinkable that policymakers would not use evidence in their work (Hammersley, 2009). Yet there is no simple linear relationship between identifying a problem that requires a policy, finding and using evidence to make that policy, and implementing the policy to solve the problem (Cairney, 2016). The term 'evidence-based policy' came into widespread use in the late 20th century, yet policy is – at best – likely to be informed by, rather than based on, evidence (Mayne et al, 2018; Sohn, 2018).

This chapter considers the meaning of 'evidence' in policymaking and 'good' in good evidence. Then it discusses what good evidence might be, its limitations and how it should be used. The chapter concludes with recommendations for policymakers.

What do we mean by 'evidence'?

In policymaking, evidence may be anything that can contribute to a decision being made. There are many types of evidence, which can be broadly divided into informal and formal evidence. Informal evidence is often used in personal decision-making. If you need to find a tradesperson for a domestic repair, you might ask friends and neighbours for a recommendation, or search reviews online. This is informal evidence, whereas formal evidence – such as legal evidence or research evidence – is collected more systematically, using tried-and-tested techniques. You might think policymakers would only use formal evidence but, in fact, they use a wide range of evidence, including informal evidence from colleagues, superiors, the public and other stakeholders (Hammersley, 2009; Sohn, 2018).

Formal and informal evidence are not discrete categories defined by the nature of the evidence itself. Whether evidence is defined as formal or informal depends on how the evidence is gathered and the wider context. Anecdotes are an interesting case in point: they often form part of informal evidence, yet are not generally regarded as credible for formal evidence. However, an anecdote provided under oath in a witness box becomes formal testimonial evidence for a case in law. Also, anecdotes that are gathered methodically may become a data set that can be used in research to generate formal evidence.

This chapter is mostly concerned with evidence generated by research, which is always formal evidence. Research evidence should be produced rigorously, with careful attention to ethical considerations, and reported in an accessible and transparent way. Helping people to understand how and why research evidence was generated is important because then users of that evidence can form a view about its integrity. Conversely, the less a user knows about the genesis of research evidence, the less they can trust that evidence.

What do we mean by 'good'?

There are different types of research evidence and there is a great deal of debate about their relative merits, though one principle holds for all sides. This principle is that good evidence comes from good research. Some government and clinical researchers espouse the idea of a 'hierarchy of evidence', though this is actually a hierarchy of research methods that often prioritises quantitative research over qualitative, mixed-methods, evaluation and other forms of research. It usually has randomised controlled trials (RCTs) at the top, then quasi-experimental or non-randomised studies, with qualitative research in third place and evaluation research trailing behind. In this hierarchy, quantitative post-positivist research is seen as 'good' and all other types of research as less good, poor or bad.

Other researchers see the hierarchy of evidence the other way round. For example, social policy researchers were found to rate in-depth interviews and longitudinal research most highly, and RCTs and quasi-experimental designs least highly (Sempik, 2007). This may be because quantitative research only answers a very limited range of questions, such as whether an intervention is effective and in what circumstances. Policymakers often have other questions, such as who will benefit and what are the barriers, and qualitative research may be better placed to answer these kinds of questions (Thompson, 2017).

Part of the backdrop to this debate is the contest between the relative merits of quantitative and qualitative research. That contest in turn is underpinned by a conflict of philosophies, focusing on the epistemological question of what we regard as knowledge. For quantitative researchers, broadly speaking, knowledge is fact-based; derived from observation and measurement; supported by logic and the principle that variables can be controlled. For qualitative researchers, broadly speaking, knowledge is based on both facts and values; derived from interaction, experience and thought; supported by theory and the principle that complexity should be acknowledged.

The indicators of quality in quantitative and qualitative research highlight these differences. Here, too, there is debate rather than consensus (Sempik, 2007), though this time the debate exists within the different approaches as well as between them. In quantitative research, indicators of quality are usually held to be reliability, replicability and validity (Bryman, 2016). Reliability refers to the stability of a measure, such as a weighing scale which, if well balanced, will weigh a perfect gram or kilo every time. A measure devised to assess a human attribute, such as depression or self-esteem, needs to be thoroughly tested before it can be confirmed as reliable. Replicability is a variant of reliability, where if an experiment is repeated, albeit with different participants or in a different location, the results will be the same. Validity focuses on whether the research results are what they appear to be, and is usually divided into internal and external validity. Internal validity is about whether the research demonstrates a relationship between cause and effect. External validity is about how far the research results can be generalised beyond the research context. Some people argue for other kinds of validity, such as measurement or construct validity (whether a measure devised to assess a human attribute actually tests what it sets out to test), ecological validity (whether research results are relevant to real life) or testimonial validity (whether participants regard results as valid) (Kara, 2020).

Qualitative researchers tried to use reliability, replicability and validity to check the quality of their own work, but soon realised they needed different indicators. Some qualitative researchers resisted the use of quality indicators altogether, seeing them as too regulatory and inflexible for qualitative work (Tracy, 2010). Others developed indicators they thought would be suitable. In the mid-1980s Lincoln and Guba put forward indicators of confirmability, dependability, credibility and transferability (Lincoln and Guba [1985], cited in Bryman, 2016). Confirmability refers to the extent to which the researcher has allowed their own values to influence the research. Dependability refers to the extent to which the research results could apply at other times. Credibility refers to how believable the results are. Transferability refers to the extent to which the results could apply to other contexts. Confirmability is a counterpoint to the quantitative researcher's underpinning value of objectivity. Dependability, credibility and transferability are, respectively, counterpoints to the quantitative research quality indicators of reliability, internal validity and external validity.

In 2010, Tracy conducted a qualitative analysis of the literature on research quality, and proposed eight indicators for qualitative research. These were designed to be accessible, flexible and universal, and to support dialogue and learning. Tracy's indicators are:

1. Worthy – relevant, timely, significant, interesting
2. Rigorous – suitable theoretical basis, appropriate methods, enough data
3. Sincere – transparent and reflexive
4. Credible – enough detail and explanation, drawing on different perspectives, trustworthy results
5. Resonant – presented in an evocative, aesthetic style; transferable results
6. Significant – contributing on different levels, such as practical, theoretical, methodological and/or ethical
7. Ethical – with a holistic rather than a tick-box approach
8. Coherent – doing what it claims, using appropriate methods, and making meaningful connections between literature, research topics, results and interpretations (Tracy, 2010)

Tracy points out that these indicators are not rules to be followed without thought. In fact, they may conflict. Full transparency could compromise participant anonymity. Presentation that is evocative and aesthetic can increase accessibility at the expense of detail and explanation (Pickering and Kara, 2017). In such situations, Tracy's view is that a researcher's primary obligation is to be truthful, both with themselves and with their audiences (Tracy, 2010).

Others have suggested quality indicators for other types of research, such as mixed-methods research (Klassen et al, 2012) and arts-based research (Lafrenière and Cox, 2013; Leavy, 2018). The quality indicators for quantitative research have been stable for some decades now, but suggestions have been made to change them, such as by including reflexivity (Shimp, 2007) or identifying relevance to practice or policy. Argyrous (2012) suggests that the principles of transparency and accountability are particularly important quality indicators for any evidence to be used in policymaking. This shows that none of these indicators are actually final or definitive.

Researchers from different professions may have different ideas about what constitutes research quality. Sempik et al (2007) surveyed 251 social policy researchers about the quality indicators they perceived to be most important. The researchers used a list of 35 quality indicators and found that 66 per cent or more of respondents valued seven indicators most highly. Here are those seven, in order of preference, from the most valued of all:

1. Research reported in ways that are accessible to the relevant audiences
2. Research design clearly addresses research question(s)
3. Data collected and analysed in transparent ways
4. Explicit account of research design, methods and data analysis
5. Research contributes to knowledge
6. Informed consent given
7. Safety of participants assured (Sempik, 2007)

Sempik et al's work does not appear to have been replicated with researchers from other professions. However, it does not seem unreasonable to suggest that, say, engineering researchers or theatre researchers might have a different order of preference for quality indicators. Indeed, another piece of empirical work around policymaking in the field of environmental sustainability found that evidence was used when it was credible (scientifically adequate), salient (relevant to policymakers' needs) and legitimate (unbiased and fair) (Cash et al, 2003). This framework has been used extensively by sustainability researchers around the world in the last two decades and has proved consistently useful in these rapidly changing times (Cash and Belloy, 2020).

As Cash's criterion of legitimacy suggests, when we talk about 'good research', we may also mean research that is ethical. As with quality indicators, there is no consensus about what constitutes ethical research, nor any agreed definition of research ethics. To some extent this is inevitable because whether an action is ethical can depend so heavily on context. Throwing a solid ball in the direction of another human being, as hard as you can, is generally deemed unethical because of the risk of injury. But if you're a bowler in cricket or a pitcher in baseball, it is expected as part of your role in the game; it would be unethical to do anything else.

Some people regard research as ethical if its design has been scrutinised and approved by a research ethics committee or institutional review board. Others argue that this does not go far enough because no committee or board can be fully informed about all the different contexts that may prevail during a research project. Researchers need the knowledge and skills to think and act ethically throughout their research work (Colnerud, 2014; Kara, 2018).

There is a clear link between ethical practice and research quality (Tangen, 2014). So, I would argue that research is 'good' if it is demonstrably ethical and meets reputable and relevant quality indicators. The question of which quality indicators are appropriate can only be decided in the light of the specific policy to be made (Parkhurst, 2017). Parkhurst asserts that quality should only be judged when the evidence has been assembled, but I disagree with him there. In my view, quality indicators should be established and then potential evidence should be assessed against those indicators before deciding whether or not to use that evidence. It is important to note that this chapter is written by a researcher, and that researchers and policymakers have different views of evidence (Cairney, 2016; Sohn, 2018). This is why I am not suggesting specific quality indicators for policymakers to use, but instead setting out a range of options.

So what is good evidence?

Good evidence comes from good research, but that is a starting point for our definition of good evidence rather than being the whole story. Whether or not evidence is 'good' in policymaking depends on what policymakers want to know, when and why they want to know it, and how and where they envisage using that evidence (Hammersley, 2009; Nutley et al, 2012; Bache, 2019).

Overall, policymakers want information to help them make decisions about the future, yet there is no evidence of the future (Wolf, 2017). Research does not always offer the most useful evidence for future-focused work because it is inevitably retrospective. Also, policymaking works on a tight schedule. Problems need swift solutions, so evidence that is good for policymaking must be available immediately or soon (Donnelly et al, 2018; Mirzoev et al, 2019).

Research is unlikely to provide all the information policymakers need (Hammersley, 2009; Sohn, 2018). However, research evidence may be useful for policymakers who need to build an argument or a case. Here, good evidence is evidence which can help build that argument or case. No matter the quality of the evidence, it is no good unless it is also relevant to the policymaker's concerns (Parkhurst, 2017; see for example the discussions in Chapters 11, 12, 13 and 16 in this volume).

Policymakers may need to build an argument or case for a way to address a social problem, or to support a governance network, or amend or maintain local or national institutions (Wolf, 2017). Any policy needed will have a variety of stakeholders, with different views about the problem to be solved and how a solution should be implemented (Donnelly et al, 2018). Therefore, research evidence will only ever be part of the evidence used by policymakers. We have seen that policymakers are not a homogeneous group, and this means their approach to evidence will vary. Other types of evidence include theoretical evidence (ideas, concepts, models), expert advice, political and professional knowledge, and experiential evidence or testimony (Glasby, 2011b; Nutley et al, 2012; Sohn, 2018; Bache, 2019). These may also be good evidence, for some policymakers, in some contexts – sometimes one or another will be given more importance than research evidence (Sohn, 2018; Bache, 2019: 1165). So research evidence, however high its quality, has to compete with other types of evidence for policymakers' attention and use. This is the arena in which 'good evidence' is accessible, well designed and fully explained, and contributes to knowledge (Sempik et al, 2007; Bache, 2019). Parkhurst (2017) takes this further, and suggests that good evidence for policymakers is focused on their current policy concerns, is constructed in ways that help to address those concerns, and is applicable to the local context.

Assuming people plan to use evidence, how they envisage doing this is important because of the scope for misuse of evidence, whether accidental or deliberate. Whatever the inherent quality of the evidence, if it is misused it is no longer good evidence. Policymaking is a political business and political people have agendas. Where these agendas dominate, policymakers may succumb to the temptation to manipulate evidence, conceal evidence and cherry-pick from evidence (Parkhurst, 2017).

What are the limitations of good evidence?

Numerous commentators have written about a 'rational model' of policymaking, in which policymakers 'identify a problem, search the evidence, identify what works and implement the findings in practice' (Glasby, 2011a: 4). Policymakers have increasing quantities of competing evidence at their fingertips, although this evidence is very unlikely to be either perfect or complete (Parkhurst, 2017). Also, good policy cannot be created using research findings alone (Hammersley, 2009; Sohn, 2018). In reality, policymakers have to assess different types of evidence, presented by people and organisations with a range of agendas, and use their judgements, expertise and experience (Hammersley, 2009) in creating policy that may or may not then be implemented in the way they intended (Glasby, 2011a).

Policymaking requires a lot of decision-making. We all like to think our decisions are logical and rational, but that is not necessarily the case. Consider mass vaccination programmes. Some people decide to be vaccinated because of the evidence that vaccination can, or does, prevent the spread of disease. Some people decide not to be vaccinated because of the evidence that vaccination can, or does, cause illness. And of course there are more stances than two: some people accept some vaccinations and not others; some people think mass vaccination programmes have no effect on health and only exist to make money for pharmaceutical companies. Each side advocates evidence in support of their view and pours scorn on other evidence, and everyone thinks they are making logical, rational decisions.

Decision-making does not only draw on conscious or even cognitive processes (Smerek, 2014). There are other factors, too, that affect decision-making, such as people's stress levels, emotional state and blood sugar levels. The vaccination example suggests that many decisions have an emotional component. Indeed, emotions can rule decision-making at times, though increasing our awareness of emotional responses can help to improve our judgement (Harford, 2020). Decision-making can also be affected by unconscious biases (Agarwal, 2020). We are all subject to these biases and there are a lot of them, such as availability bias (we place more importance on evidence that is easily available or familiar) and confirmation bias (we are more inclined to pay attention to, and trust, evidence that supports our existing views). There are actions we can take to become more aware of and so mitigate these biases, but we cannot eradicate them completely (Agarwal, 2020).

Evidence is dynamic (Shaxson, 2005). As knowledge develops and social priorities change, even good evidence may become obsolete. This can happen surprisingly fast so it is important for evidence to be as up to date as possible.

Good evidence is worth little unless it is used. Using evidence in policymaking is clearly not a straightforward process. In particular, research findings that challenge a prevalent paradigm of governance are not likely to be used, no matter how high quality the evidence may be (Monaghan and Ingold, 2019). Also, the producers of evidence may seek to influence or even persuade policymakers by framing and presenting their evidence in appealing ways (Mayne et al, 2018; Sohn, 2018). While it is worth striving to produce and use good evidence, it is also important to recognise that the quality of evidence alone cannot determine its fate. It is perhaps inevitable that policymakers will ignore some good evidence, and give credence to some evidence that is inadequate or misleading.

Cairney (2016) identified a number of barriers to the use of evidence in health-related and other types of policymaking. These included: research that is not available or accessible; findings that are not clear, relevant or reliable; evidence that arrives at the wrong time; some policymakers having poor research skills. Rickinson and his colleagues (2016) identified some barriers to the use of evidence in policymaking for education. These included: over-reliance on easily accessible sources; ineffective knowledge management processes; variable abilities of policymakers to find and use good quality evidence; and less use of evidence in later than earlier stages of the policymaking process. No doubt barriers vary depending on factors such as context, policy topic and location in time. Nevertheless, the overlaps here are interesting, with accessibility and policymakers' skills featuring in both studies.

How should good evidence be used?

As we have seen, a single piece of evidence – however good it might be – will not lead to the making of a policy. Sohn (2018) develops this thought by identifying two stages in the policymaking process. The first stage involves deciding whether to establish a particular policy. If a positive decision is made, then the second stage involves decisions about how to formulate that policy. Evidence is used differently in each stage: informal evidence is more prominent in the first stage, and formal evidence in the second stage (Sohn, 2018).

Clearly research evidence is not the only kind of evidence policymakers will use (Hammersley, 2009; Bache, 2019). However, policymakers seek to influence government ministers, and the use of sound evidence is a key part of this process (Bache, 2019). Policymakers have to collate and synthesise evidence from many disciplines and sources, which may include qualitative and quantitative research, academic and grey literature, and expert and public opinion (Donnelly et al, 2018). Some research methods, such as systematic reviews or meta-analyses, also involve collation and synthesis. These types of research can be particularly valuable for policymakers, though they are also time-consuming and so may not fit with policymaking schedules (Donnelly et al, 2018). Also, even a systematic review or meta-analysis is unlikely to form the basis of a policy on its own. Yet the evidence gathered by policymakers, whether or not it includes other collated work(s), will need to be collated and synthesised accurately, concisely and with as little bias as possible.

Donnelly and her colleagues are a group of academics, policymakers and other stakeholders in research funding and publishing. They have produced a set of four practical and ethical principles to help policymakers collate and synthesise good evidence (Donnelly et al, 2018). These are:

1. Inclusive – of a variety of skills and people including policymakers themselves, plus different types and sources of evidence.
2. Rigorous – using as much good evidence as possible, with bias recognised and minimised, and independently reviewed for quality assurance.
3. Transparent – clear description of: research question(s), methods, sources of evidence and quality assurance process; complexities and areas of contention; assumptions, limitations, uncertainties and any evidence gaps; personal, political and organisational interests; and any conflicts of interest.
4. Accessible – written in plain language, timely and freely available online.

Inclusive collation and synthesis can help policymakers to draw on the full range of relevant evidence types, sources and expertise. Rigorous work helps to avoid mistakes and minimise bias. Transparency produces work that is more credible, replicable and useful, and easier to keep up to date. Accessible work is much more likely to be used, and has the potential to create benefit beyond its immediate audiences (Donnelly et al, 2018).

Conclusion

The work of policymakers is complex and demanding, political and practical, contested and continually changing. Evidence should be important in policymaking, but evidence alone, however well framed, cannot tell policymakers what to do (Cairney, 2016; Sohn, 2018). Policymakers have to take other factors into consideration, such as political priorities, public perceptions and cost (Nutley et al, 2012). Sohn, in her detailed study of policymaking for the Ontario Poverty Reduction Strategy, found that policymakers were 'influenced by a complex interplay of their own psychology, heuristics, and moral agendas, as well as the political and economic climate, their advisors, the evidence presented to them, and the roles and responsibilities of their particular positions' (Sohn, 2018: 7). She concluded that policy emerges from 'a complex interplay of evidence being used in different ways' by different people 'throughout the process' (Sohn, 2018: 11).

As Parkhurst (2017: 8) asserts, 'an explicit desire to improve the use of evidence in policy will require the establishment of new principles of what would constitute good evidence to inform policy, as well as considering what constitutes the good use of evidence within a policy process'. Parkhurst calls for the establishment of evidence advisory systems to promote good uses of good evidence in policymaking – or, in his definition, 'working to ensure that rigorous, systematic and technically valid pieces of evidence are used within decision-making processes that are inclusive of, representative of and accountable to the multiple social interests of the population served' (2017: 8). I can see some merit in this suggestion, but I am not sure the world needs more systems, and anyway I think it is very unlikely to happen in our messy reality.

Donnelly et al (2018) articulate a different goal: 'to create an effective marketplace for synthesis in which policymakers and commentators always seek the best evidence because they know it will be available, and researchers synthesize evidence because they know it will make a difference'. This is also an ideal but seems a more practical one as instead of setting up new systems, it aims to improve policymaking culture by building on the systems we already have.

People value different types of evidence depending on their views of research and ethics, and that is difficult to change (Sohn, 2018; Bache, 2019). However, agreement on appropriate quality indicators for evidence and principles for evidence use, *before* evidence is selected, can help to mitigate the effect of these differences in practice.

Perhaps it would be useful to determine indicators for the quality of evidence use, as well as indicators for the quality of the evidence itself (Monaghan and Ingold, 2019: 365). For the moment, in the interests of making better use of existing systems, I would recommend that policymakers identify appropriate quality indicators for their work, use those indicators to assess potential evidence, and make ethical use of the evidence that they judge to be good quality, in accordance with the four principles put forward by Donnelly and her colleagues. This approach would give policymakers confidence that the policies they create are built on sound evidential foundations.

References

Agarwal, P. (2020) *Sway: Unravelling Unconscious Bias*, London: Bloomsbury Sigma.

Argyrous, G. (2012) 'Evidence based policy: principles of transparency and accountability', *Australian Journal of Public Administration*, 71(4): 457–68.

Bache, I. (2019) 'How does evidence matter? Understanding "what works" for wellbeing', *Social Indicators Research*, 142(3): 1153–73.

Bryman, A. (2016) *Social Research Methods* (5th edn), Oxford: Oxford University Press.

Cairney, P. (2016) *The Politics of Evidence-Based Policy Making*, London: Palgrave Macmillan.

Cash, D.W. and Belloy, P.G. (2020) 'Salience, credibility and legitimacy in a rapidly shifting world of knowledge and action', *Sustainability*, 12(18): art 7376, available from: https://doi.org/10.3390/su12187376 [accessed 11 February 2022].

Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Guston, D.H., Jäger, J. and Mitchell, R.B. (2003) 'Knowledge systems for sustainable development', *Proceedings of the National Academy of Sciences*, 100(14): 8086–91.

Colnerud, G. (2014) 'Ethical dilemmas in research in relation to ethical review: an empirical study', *Research Ethics*, 10(4): 238–53.

Donnelly, C., Boyd, I., Campbell, P., Craig, C., Vallance, P., Walport, M., Whitty, C., Woods, E. and Wormald, C. (2018) 'Four principles to make evidence synthesis more useful for policy', *Nature*, 558: 361–4, available from: https://www.nature.com/articles/d41586-018-05414-4 [accessed 11 February 2022].

Glasby, J. (2011a) 'Introduction', in J. Glasby (ed) *Evidence, Policy and Practice: Critical Perspectives in Health and Social Care*, Bristol: Policy Press, pp 1–10.

Glasby, J. (2011b) 'From evidence-based to knowledge-based policy and practice', in J. Glasby (ed) *Evidence, Policy and Practice: Critical Perspectives in Health and Social Care*, Bristol: Policy Press, pp 85–98.

Hammersley, M. (2009) 'What is evidence for evidence-based practice?', in R. St. Clair (ed) *Education Science: Critical Perspectives*, Leiden: Brill Sense, pp 101–11.

Harford, T. (2020) 'Facts v feelings: how to stop our emotions misleading us', *The Guardian*, 10 September, available from: https://www.theguardian.com/science/2020/sep/10/facts-v-feelings-how-to-stop-emotions-misleading-us [accessed 14 September 2020].

Kara, H. (2018) *Research Ethics in the Real World: Euro-Western and Indigenous Perspectives*, Bristol: Policy Press.

Kara, H. (2020) *Creative Research Methods: A Practical Guide* (2nd edn), Bristol: Policy Press.

Klassen, A., Creswell, J., Plano Clark, V., Clegg Smith, K. and Meissner, H.I. (2012) 'Best practices in mixed methods for quality of life research', *Quality of Life Research*, 21(3): 377–80.

Lafrenière, D. and Cox, S.M. (2013) '"If you can call it a poem": toward a framework for the assessment of arts-based works', *Qualitative Research*, 13(3): 318–36.

Leavy, P. (2018) 'Fiction-based research', in P. Leavy (ed) *Handbook of Arts-Based Research*, New York: Guilford Press, pp 190–207.

Mayne, R., Green, D., Guijt, I., Walsh, M., English, R. and Cairney, P. (2018) 'Using evidence to influence policy: Oxfam's experience', *Palgrave Communications*, 4: art 122, available from: https://doi.org/10.1057/s41599-018-0176-7 [accessed 11 February 2022].

Mirzoev, T., Poudel, A.N., Gissing, S., Doan, T.T.D., Ferdous, T., Regmi, S., Duong, M.D., Baral, S., Chand, O., Huque, R., Hoang, V.M. and Elsey, H. (2019) 'Is evidence-informed urban health planning a myth or reality? Lessons from a qualitative assessment in three Asian cities', *Health Policy and Planning*, 34(10): 773–83.

Monaghan, M. and Ingold, J. (2019) 'Policy practitioners' accounts of evidence-based policy making: the case of Universal Credit', *Journal of Social Policy*, 48(2): 351–68.

Nutley, S., Powell, A. and Davies, H. (2012) *What Counts as Good Evidence? Provocation Paper for the Alliance for Useful Evidence*, London: Alliance of Useful Evidence, available from: https://www.nesta.org.uk/report/what-counts-good-evidence/ [accessed 11 February 2022].

Parkhurst, J. (2017) *The Politics of Evidence: From Evidence-Based Policy to the Good Governance of Evidence*, Abingdon: Routledge.

Pickering, L. and Kara, H. (2017) 'Presenting and representing others: towards an ethics of engagement', *International Journal of Social Research Methodology*, 20(3): 299–309.

Rickinson, M., de Bruin, C., Walsh, L. and Hall, M. (2016) The Use of Evidence in Educational Policy: A Pilot Study in Victoria, Melbourne: Department of Education and Training.

Sempik, J., Becker, S. and Bryman, A. (2007) 'The quality of research evidence in social policy: consensus and dissension among researchers', *Evidence & Policy*, 3(3): 407–23.

Shaxson, L. (2005) 'Is your evidence robust enough? Questions for policy makers and practitioners', *Evidence & Policy*, 1(1): 101–12.

Shimp, C.P. (2007) 'Quantitative behavior analysis and human values', *Behavioural Processes*, 75(2): 146–55.

Smerek, R.E. (2014) 'Why people think deeply: meta-cognitive cues, task characteristics and thinking dispositions', in M. Sinclair (ed) *Handbook of Research Methods on Intuition*, Cheltenham: Edward Elgar, pp 3–14.

Sohn, J. (2018) 'Navigating the politics of evidence-informed policymaking: strategies of influential policy actors in Ontario', *Palgrave Communications*, 4: art 49, available from: https://doi.org/10.1057/s41599-018-0098-4 [accessed 11 February 2022].

Tangen, R. (2014) 'Balancing ethics and quality in educational research: the ethical matrix method', *Scandinavian Journal of Educational Research*, 58(6): 678–94.

Thompson, S. (2017) 'Qualitative methods and public policy', *Journal of Positive Psychology*, 12(3): 321–2.

Tracy, S.J. (2010) 'Qualitative quality: eight "big-tent" criteria for excellent qualitative research', *Qualitative Inquiry*, 16(10): 837–51.

Wolf, A. (2017) 'Seeing ahead – relationally', presentation to 3rd International Conference on Public Policy (ICPP3), 28 June, Singapore, available from: https://www.ippapublicpolicy.org/file/paper/593b7157d8be2.pdf [accessed 11 February 2022].

16

Black Lives Matter: reconsidering systemic racism

Sally Burns

Introduction

'Structural and institutional inequality is real and it's everywhere. Racism is lived and experienced by our friends, neighbours, colleagues, teachers, doctors and in all our communities' (Zammit, 2020). This was written by a young friend of mine and it inspired me to reflect deeply on these issues and to write about them.

Racism is alive and well in 2020, the time of this writing. Ethnic and racial discrimination and inequality not only exist but are thriving in our global society. How do we know? If your skin is black, you know. As Delpit said many years ago, 'Those with power are frequently least aware of – or least willing to acknowledge – its existence [and] those with less power are often most aware of its existence' (Delpit, 1988). Her words are truer than ever today. Who has the power? Many would say power is in the hands of privileged Whites and to a large extent, this is true, as discussion in this chapter will show. However, examples from history also illustrate that racism is not just confined to colour. It boils down to a formula to which I was introduced many years ago: prejudice + power = discrimination.

Racism is not new. There have been innumerable examples throughout the ages. In our lifetimes many of us have witnessed instances of extreme racism in the form of genocide. Six million Jews and up to half a million Romani were exterminated during the Holocaust in Germany between 1941 and 1945 (Wikipedia, 2020c). In Rwanda in 1994 it is estimated that 800,000 Tutsis were slaughtered by Hutus (BBC, 2019). More recently, in 2017, the persecution of the Rohingya Muslims in Myanmar resulted in the death of 6,700 Rohingyas (BBC, 2020).

It was known that systemic racism was a problem in the 1970s. In response, the British government passed the Race Relations Act in 1976 focusing particularly on discrimination and victimisation and defining a number of terms such as 'racial grounds' and 'racial group' (Race Relations Act, 1976). However, by the end of the century there was a recognition that this was not working. The MacPherson Enquiry, in 1997, triggered largely by the murder of Black teenager Stephen Lawrence in 1993, identified systemic racism in the UK police force as well as inherent in the justice system, the BBC, government and education. There were 70 proposals to address racism in the police force alone. According to Ben Quinn of *The Guardian*, 'Some 67 of the report's recommendations led to specific changes in practice or the law within two years of its publication' (Quinn, 2019).

Was this enough? Not according to Dal Babu, a former chief superintendent in the Metropolitan police: 'The continued disproportionate use of stop and search and the vanishingly low numbers of stops that result in a substantive charge, never mind conviction, cements in young black consciousness an underlying enmity – a feeling that the police are "other"' (Babu, 2019). He feels that while outwardly there may have been changes and lip-service paid to anti-discrimination, what he calls the *canteen culture* 'still has subtle ways of excluding BAME [Black, Asian and Minority Ethnic] staff as well as LGBT officers' (Babu, 2019).

Racism in schools, universities and government

Racism is endemic in educational systems as well as at government level. 'Black students are excluded at triple the rate of White children', averred Anton Chisolm in a recent Diverse Ed event (Wilson, 2020). This was corroborated by an earlier piece on the BBC News website (Richardson, 2020). It is sadly a fact that Black teenagers are more often stopped and searched or arrested than their White counterparts. An example of this was reported in the Canadian news in June 2019. Jamal Koumiye-Boyce, a student at Ottawa University, was detained by police while heading for class on his skateboard. He stated on social media that he was put in handcuffs and interrogated for two hours. White students on their skateboards were not stopped (Ramlakhan, 2019) As this example illustrates, racism remains an issue in universities today. A revealing article by Murugesu (2020) outlines a number of case studies in the UK, US and Canada where Black students have been victimised on the grounds of their colour.

Nor are teachers immune to racism. As a newly qualified teacher, Ebanie Xavier-Cope reported racist name-calling by students to the head of her school. The latter said he did not know anything about it and seemed unwilling to take action. Only when Ebanie threatened to take it further did the school administration react. A student was suspended but the teacher never received an apology of any sort (Wilson, 2020). An American teacher (Statz, see next section and Kingkade, 2020) who taught a lesson about Black Lives Matter and was subsequently lambasted by parents was likewise not supported by her school. One mother in Burlington, Wisconsin, the town where the incident happened, felt the school administrators were not doing enough to stamp out racism which was making her daughter's and other children's lives a misery. However, a lawyer working for the school said the issue was nothing to do with race. This reaction seems to be common (Darnisha Garbade cited in Kingkade, 2020).

Sara Bafo of Goldsmiths University of London suggests students don't report racism because they feel nothing is done about it or the tribunal-style process is intimidating and may even harm the plaintiff, who is likely to be pushed into signing a non-disclosure agreement. It would seem, according to Bafo, that, 'Senior management take the university's reputation more seriously than combating racism' (Murugesu, 2020). Cited in the same article is a statistic from a Goldsmiths, University of London, report that, 'The majority of BME [Black and Minority Ethnic] survey respondents do not know where to report a hate crime at Goldsmiths College – a total of 79%' (Akel, 2019). Sara Khan, former liberation and access officer at the University of Manchester, said, 'I don't think a single case resulted in a satisfactory outcome' (Khan in Murugesu, 2020).

The UK's Equality and Human Rights Commission found that, 'Around a quarter of students from an ethnic minority background (24%), and 9% of White students, said they had experienced racial harassment since starting their course. This equates to 13% of all students' (EHRC, 2019: 2). However, the 159 UK state universities logged only 559 complaints of racial harassment during 2015–19. The report notes, 'Our call for evidence found that fewer than 4 in 10 student complaints of racial harassment, and around 1 in 6 staff complaints, were upheld and offered some kind of redress' (EHRC, 2019: 6).

According to Murugesu, another factor is that the tribunals rarely have Black and ethnic minority members and there is a sense that those who are appointed to senior management positions in universities are likely to be people who prefer to uphold the status quo. This again tallies with Delpit's observation so many years ago ('Those with power …') and with many recent comments recorded later in the section on 'Representation'.

Perhaps the most significant factor is the dearth of data to demonstrate the issue. While a vast number of students cited racist incidents on Instagram accounts in the US and Canada using the account @BlackatHarvardLaw, this is not recognised as hard, reliable data. The lack of 'proof' of racism allows universities to hide behind their impressive-sounding policies and to avoid addressing the racism (Murugesu, 2020). This was also stated by Hannah Wilson who said that there is a significant lack of data on racist incidents particularly relating to staff (Wilson, 2020).

Racist incidents are often obfuscated as Kingkade (2020) notes:

> While the district reported zero cases of racial discrimination to the state from fall 2016 to spring 2019, it documented 21 incidents of bullying based on race during the same period in reports to the school board,

> records show. District data also show Black students were disciplined at a rate nearly five times higher than their white peers.

The district described these as 'student-to-student microaggressions that may or may not be intended as racist but inflict harm' (Kingkade, 2020).

Black Lives Matter

As young Mathilda said in a video, 'When we say black lives matter, it doesn't mean all lives does not matter, it simply means black lives are the ones in danger at the moment' (The Black Curriculum, 2020). The Black Lives Matter campaign, first developed through social media in 2014 (Wikipedia, 2020b; Black Lives Matter, 2020), has brought racism out in the open and highlighted systemic racism in police forces throughout the world. The death of George Floyd in Minneapolis on 25 May 2020 caused thousands to take to the streets in protest. It was the latest of many such deaths and acts of aggression against the Black, Asian and Minority Ethnic (BAME) community in the US and UK.

Two incidents in 2020 provide further recent evidence of other racist attacks. On 23 February, Ahmaud Arbery was shot and killed while out jogging in Brunswick in the south-east of the US (Griffith, 2020). In August of the same year, in Kenosha, Wisconsin, Jacob Blake was shot multiple times in the back by police (Givetash et al, 2020). These two deaths not only provide recent evidence of other racist attacks but also serve to illustrate the range of reactions the Black Lives Matter campaign has unleashed, as the next example shows.

In Burlington, Wisconsin, a fourth-grade teacher named Melissa Statz happened to overhear students wanting to know more about protests in nearby Kenosha following the shooting of Jacob Blake. Statz responded by teaching a lesson on Black Lives Matter. Two Black students personally thanked her for teaching about racism; however, her action unleased an unprecedented backlash, notably on Facebook, creating a furore in the town that has divided its inhabitants. The school board did not discipline Statz but nor did they initially offer her support. It was only later, when there were two incidents of graffiti on school property using the n-word that the school administration spoke out against the attacks and openly recognised them as racist. Reporting on Statz's story, Tyler Kingkade (2020) wrote, 'The words "Black Lives Matter" have become a Rorschach test, splitting people who see it as a political slogan inappropriate for school, and those who consider it a statement vital to ensuring students of color feel safe and valued.' It is interesting to note that recently, a study by the Pew Research Center has shown that support for Black Lives Matter has declined in the US, notably among the White population. Whereas in June some 60 per cent of Whites surveyed said they supported the campaign, by November the figure was only 45 per cent; support from the BAME community has changed little (Thomas and Horowitz, 2020). It would seem racism is an issue for Whites following the well-publicised murder of a Black person but it quickly slips off the agenda.

To conclude this section, it should be noted that Black Lives Matter is by no means a simple response to racism. Even within the movement there are differences of opinion and concerns about its finances, leadership and direction (Eligon, 2021). What it has done is bring systematic racism to the attention of a worldwide audience.

COVID-19

The outbreak of COVID-19 in 2020 has further served to highlight racial inequality with many countries exhibiting a disproportionate number of cases in Black and minority ethnic communities as compared to White communities. In the UK, there has been a higher incidence of cases and deaths among BAME communities than in White communities. Citing a survey in June 2020, Nick Stripe of the UK Office for National Statistics states, 'The raised risk of death involving COVID-19 for people of Black ethnic background of all ages together was 2.0 times greater for males and 1.4 times greater for females compared with those of White ethnic background' (Stripe, 2020). The report suggests the risk of death from COVID-19 is greater for Black people than for other ethnic minorities. Godin (2020), citing National Health Service data, maintains, 'England's Black, Asian and Ethnic Minority groups are two to three times more likely to die from COVID-19 compared to the general population'. This is confirmed by a government report (PHE, 2020b). Moncrieffe (2020) maintains the pandemic has further highlighted the spectre of White privilege in Britain.

There has also been a higher incidence of cases and deaths among BAME communities in the US and France. In the US, 'African Americans have died from the disease at almost three times the rate of white people' (Pilkington, 2020). In Chicago alone, nearly two thirds of the COVID-19 deaths were among the Black population who make up just under one third of the city's population (Lutz, 2020). Based on statistics from the National Institute of Statistics and Economic Studies in France (INSEE), 'Death rates among immigrants from sub-Saharan Africa doubled in France and tripled in the Paris region at the height of France's coronavirus outbreak' (quoted in Associated Press, 2020).

It is important to note that the imbalance is not only evident in BAME patients but is clear among BAME healthcare professionals. The NHS reported, 'We know there is evidence of disproportionate mortality and morbidity among black, Asian and minority ethnic (BAME) people, including our NHS staff, who have contracted COVID-19' (NHS, nd). According to a report on the Sky News website, while just under half of the UK National Health Service (NHS) medical personnel are from a BAME background, 72 per cent of the healthcare staff who have died from COVID-19 are from Black and ethnic minorities (Rashid, 2020).

While studies do acknowledge that the relationship between health and ethnicity can be complicated by factors such as endemic diseases and countries of birth which might explain some of the discrepancy, the figures are nonetheless striking.

A report by Public Health England suggests that COVID-19 was in itself not the cause of the inequality, 'but rather it exposed and exacerbated longstanding inequalities affecting BAME groups in the UK' (PHE, 2020a: 27). It goes on to propose that the crisis should be seen as a chance for urgent action to address these inequalities. Among the recommendations in the report, perhaps the most significant in the context of this chapter is, 'Tackling workplace bullying, racism and discrimination to create environments that allow workers to express and address concerns about risk' (PHE, 2020a: 33).

Moncrieffe (2020) argues that there is a difference in the quality of care between BAME and White patients but this is difficult to prove. Tello goes some way to doing this. Citing *Health United States 2015*, Tello (2017) writes, 'It is well-established that blacks and other minority groups in the U.S. experience more illness, worse outcomes, and premature death compared with whites.' Tello's opinions are in part based on a 2004 report that states there are, 'systematic racial and ethnic differences in the receipt of a broad spectrum of therapeutic interventions' (NRC, 2004: 97). Why this should be so is not clear. Tello thinks that, in the main, doctors are not overtly racist but that the system they are working in is. Not only does this affect healthcare, in her opinion, but all social, political and economic spheres.

Mental health is another area which has been under scrutiny. In the UK a Bradford NHS (2016) survey found that 'Young black men are six times more likely than young white men to be sectioned for compulsory treatment under the Mental Health Act.' More recently Dr Sami Timimi, a visiting professor at the University of Lincoln said, 'The label of schizophrenia … is disproportionately applied to black people, particularly black men' (Timimi, cited in Siddique, 2020). The Royal College of Psychiatrists decided to address this directly. In an open letter to its new president, Dr Adrian James calls for action given that this is a 'once-in-a-generation opportunity to put psychiatry at the forefront in tackling systemic racism' (James, cited in Siddique, 2020). At the same time, it is important to note that, 'the stress associated with being discriminated against based on race/ethnicity affects mental and physical health' (PHE, 2020a: 38).

At government level, the Windrush Scandal

One of the biggest scandals of the last 20 years demonstrating systemic racism in the UK is the Windrush Scandal. Wendy Williams defines the Windrush generation as: 'a racial group by reference to nationality and national origin, deriving from the Caribbean and having entered the UK, or their ascendant relatives having entered the UK, in the window between 1948 and 1973, and who almost all are black' (Williams, 2020: 11).

The issue came to light in 2001 when a British passport holder went to Jamaica for a holiday and, on return, was told his passport was not valid and he had no right of abode in the UK. He died in Jamaica nine years later. Legislation had been passed in 1988 which changed the status of thousands of people like this man. The National Audit Office reckoned the position of half a million people in the UK could be affected (NAO cited in Williams, 2020: 25). Williams believes the crisis was brought about by a plethora of legislative amendments and policy changes in the 30-year period from the 1960s to the 1980s exacerbated by poor documentation. She says of the victims of this scandal, 'Their history was institutionally forgotten.' As to whether this was an act of racism Williams concludes:

> While I am unable to make a definitive finding of institutional racism within the department, I have serious concerns that these failings demonstrate an institutional ignorance and thoughtlessness towards the issue of race and the history of the Windrush generation within the department, which are consistent with some elements of the definition of institutional racism. (Williams, 2020: 7)

Williams certainly seems to think race was a factor in these events (2020: 13).

It is clear then that racism is inherent in education, police forces, health services and politics. It would seem reasonable from this to extrapolate to other systems too. The problem is real, is present and has been highlighted internationally by the Black Lives Matter campaign. The question now is what can we do about it? I would like to begin answering this with my own personal experiences in education and the recent compelling question I have been asking myself, 'Where did we go wrong?'

Where did we go wrong?

In the 1980s there were efforts to counter racism through multicultural and anti-racist initiatives within the UK education system. Multicultural projects were set up all over the country, often combined with Section 11 funding[1] which provided for the teaching of English as a Second Language (ESL[2]). Language centres in cities such as Leicester, Birmingham and Coventry produced excellent resources to promote multiculturalism and anti-racism. Many development education centres sought to raise awareness of global inequality by developing students' concepts, attitudes, skills and knowledge. In Bury, north Manchester, the project I was working on in the late 1980s was aiming to bring multicultural, anti-racist training to the borough including the 'white highlands': areas where the population was predominantly White. There were projects similar to ours all over the UK.

By the 1990s I was working freelance with education advisors, with teachers and with students in schools, running workshops to recognise racism and promote anti-racism. I was one of many engaged in such activities. Backed by a multi-faith advisory group, I co-wrote *Values and Visions: A Handbook for Spiritual Development and Global Awareness* (Burns and Lamont, 1995) setting out ways of working which celebrate diversity, tackle difficult subjects and enable students to envisage a world at peace. We thought we were doing a good job.

Thus, in the summer of 2015, when my co-author and I discussed rewriting *Values and Visions*, our first thoughts were that we no longer needed an emphasis on global awareness; 20 years on, things had changed. I personally had spent most of those 20 years in an international school overseas with a student body of over 70 nationalities and more than 50 languages. The many different cultures interacted respectfully and racist incidents were rare. We had no school rules but everyone was charged with observing the school's mission: 'To provide an international education of the highest quality to enable students to be confident, responsible, caring, life-long learners'. The emphasis on responsibility and caring in that multiracial, diverse community, combined with the fact that bullying in any form was not tolerated, worked. Sadly, I soon came to realise that not all schools were like the one I worked in. I was in a bubble.

That same summer I learned of the lack of cultural awareness among teachers on an in-service training day in the UK. I learned of cultural insensitivity among mature students studying for a doctorate in education in the US. Things had not changed outside my bubble. They were at best static, at worst worse as the incidents of racism already described indicate. It seemed that, despite the efforts of countless educators and other professionals around the world for decades, racism was still prevalent. Jane Elliot, famous for her controversial blue eyes/brown eyes experiment in 1968 (Bloom, 2005), cannot believe that over half a century on, racist attitudes are still so prevalent: 'We are repeating the blue-eyed/brown-eyed exercise on a daily basis' (NPR, 2020). Professor Vini Lander recently stated, 'I've been in this for 30-odd years and we are still talking about the same issues' (Wilson, 2020). Where did we go wrong?

Returning to the materials predominantly being used in schools in the 1980s and 1990s might offer some insights. I did my master's degree in

British Columbia in 1988. I focused on images of First Nations people in social studies textbooks being used in the province. During the course of my research I was naively surprised to learn that there were monopolies in the publishing industry. In the US, for example, the largest educational publisher was based in Texas. They supplied books throughout the States. If a book was adopted for schools in Texas, it was likely to be used in the majority of states in the country, representing huge revenue for the publishers. The content had to be approved, of course, and this is where things got interesting.

The content favoured a White, middle-class market. There were numerous examples in the social studies textbooks being used which showed the White saviour versus Black savage/noble savage: Englishmen in finery handing beads to the natives. Less blatant and more interesting was what was not in the books. It is a fact that many First Nations people died of smallpox because of the arrival of the White settlers. This was not evident in the texts. Similarly, bison were described as having died out when in fact they were slaughtered in their thousands by White settlers who wanted their skins and their tongues, which were considered a delicacy (Wikipedia, 2020a). It was also part of the government's plan to cut off a vital food supply of the Native Americans. '"Buffalo" Bill Cody, who was hired to kill bison, slaughtered more than 4,000 bison in two years' (PBS, nd). As Michael Apple noted, 'The not said of a work is as powerful as the said since "ideology is present in the text in the form of its eloquent silences"' (Apple, 1982). A distorted view of history appeared to be presented to students in North America (and no doubt elsewhere) in the 1980s.

Christopher Richards maintains this is still the case in the field of English Language Teaching (ELT) publishing. According to him, three major UK publishers control the textbooks used for teaching ELT in Madrid, Spain (Wilson, 2020). These depict predominantly White characters in stereotypical/traditional roles: a White heterosexual couple getting married in a church for example. How many other such monopolies are there?

So there were two things going on in education in the latter part of the 20th century: on one hand, movements to address racism and promote multiculturalism and diversity and on the other, teaching resources clearly presenting a distorted image of society. Pioneers such as the late Rachel Evans, with whom I collaborated, produced a database of nearly one thousand teaching resources for primary, secondary and tertiary education covering all subject areas. We sourced the texts from small, minority and niche publishers. Our aim was to ensure that books in schools represented a balance of colour, race, gender, sexual orientation and disability. Rachel and I promoted small publishers such as Tamarind Press set up by author Verna Wilkins, selling titles like *Dave and the Tooth Fairy* where a young Black boy surprises the tooth fairy and is amazed to find she is Black or

Board Games from Around the World, a great resource for mathematical investigations and *The Real McCoy*, a celebration of the life of the African American inventor.

This was a mammoth task and not everyone saw that there was a problem to be addressed. Indeed the attitude was often, 'No problem here', to quote the title of Chris Gaine's book (Gaine, 1987).[3] The attitude identified by Gaine was probably not malicious. Many educators genuinely believed things were all well and good. In fact, as Vini Lander recently pointed out, it might have seemed a logical position to take, particularly if a school was located in a predominantly White area. It was the colour-blind approach, the 'Well, they're very well integrated' response (Wilson, 2020). I was stunned to find recently that educators 30 years on are still talking about under-representation of BAME characters in books. According to Lisa Stephenson of the Story Makers Company, 'Only 4% of children's lit. had black & ethnic minority characters' (Wilson, 2020). As a student she cited said, 'It's not about characters that look like me. It's about characters that are like me.'

What it boils down to is that, 'Issues of power are enacted in classrooms' (Delpit, 1988). A teacher is in a position of power vis-à-vis their students, a White teacher in a class of Black students even more so.

White privilege

The underlying issue involves presumptions about White privilege and what counts as the norm. I have been in a racial minority for much of my life (see The Values and Visions Foundation, 2020). I have been a 'hawaja' and a 'mzungu'. I was the only White female in a remote Sudanese village. I was a White female living in Oman for over 21 years. I have been welcomed as a White family member in rural North and East Africa. In none of those settings was I taunted and mistreated for the colour of my skin. My dress, lifestyle and alien ways were often a matter of discussion and laughter rather than a cause for taunts and derision. I never feared for my safety, I was never mistreated by the forces of law and order for my colour. I was treated with curiosity and respect. As I result, I have been complacent because I have not been afraid. A co-educator from the 1980s and co-contributor to my blog summed this up as follows: 'A white person on their own in a situation brings a legacy of power and authority whereas a black person in a similar situation is vulnerable, seen as other and, at worst, victimised and attacked' (The Values and Visions Foundation, 2020).

The issue of White privilege rears its ugly head. Cuncic (2020) gives this definition: 'White privilege is a concept that highlights the unfair societal advantages that white people have over non-white people. It is something that is pervasive throughout society and exists in all of the major systems and institutions that operate in society, as well as on an interpersonal level.'

McIntosh (1989) wrote of privilege as 'a favored state' often determined by sex or race. She said there is 'a pattern running through the matrix of white privilege, a pattern of assumptions which were passed on to me as a white person' (McIntosh, 1989: 3). McIntosh's fascinating essay lists no fewer than 50 attributes of being a White person: attributes which we White people never think about and unquestioningly accept. McIntosh cites Elizabeth Minnich, 'Whites are taught to think of their lives as morally neutral, normative, and average, and also ideal, so that when we work to benefit others, this is seen as work that will allow "them" to be more like "us"' (McIntosh, 1989: 4). Thirty years on things appear not to have changed. Moncrieffe states, 'Society's default perspective for history is framed by the majority ethnic group, and in Britain that is the privileged "White" perspective' (Moncrieffe, 2020). He cites Robin Di Angelo who echoes McIntosh and Minnich saying White people are brought up to view themselves as 'universal human beings', in other words, the norm.

Thinking about what went wrong, it is perhaps that we well-meaning, predominantly White educators, were unaware of or failed to take sufficiently into account our White privilege. An excerpt from McIntosh's essay expresses this eloquently:

> There was one main piece of cultural turf; it was my own turf, and I was among those who could control the turf. My skin color was an asset for any move I was educated to want to make. I could think of myself as belonging in major ways, and of making social systems work for me. I could freely disparage, fear, neglect, or be oblivious to anything outside of the dominant cultural forms. Being of the main culture, I could also criticize it fairly freely. (McIntosh, 1989: 3)

McIntosh came to these understandings through her work on gender relations, realising men work from a position of privilege of which they are not aware.

At the time McIntosh wrote and on into the 1990s we taught *about* racism. We had students feel what it was like in simulations such as *Spots*.[4] We encouraged them to relate to their own experiences of prejudice, for example name-calling based on hair colour, being overweight or wearing glasses. We were unconsciously perpetuating the 'us' and 'them' syndrome. Racism was something that others experienced, that put 'them' at a disadvantage. We were not addressing the core underlying issue of White privilege.

McIntosh (I wish I had read her work at the time!) put into words what we were trying to do: 'I was taught to think that racism could end if white individuals changed their attitude' (McIntosh, 1989: 4). This is what I myself felt and the premise on which others working alongside me based our training. It is part of the equation but not enough, which I now see.

What should be done about it?

The question is what should we be doing differently now from what we know did not work in the past? What policies should be put in place to ensure racism is eliminated for ever? Where better place to start than with the adults of the future?

Education

According to McIntosh, we need to differentiate between what should be the norm for everyone and what is founded on arrogance; what she calls 'earned strength' versus 'unearned power' (McIntosh, 1989: 3). She says, 'My schooling gave me no training in seeing myself as an oppressor, as an unfairly advantaged person, or as a participant in a damaged culture' (McIntosh, 1989: 1).

According to Moncrieffe (2020), 'The biggest challenge to achieving a genuine sense of equality in society is through the transforming of fixed Eurocentric mindsets held by the majority of White-British teachers – the default position from which they begin to think about teaching and learning through curriculum.'

Countless teachers like Statz (cited earlier) support teaching about racism: 'An EdWeek Research Center survey in June found the vast majority of educators, 81 percent, supported Black Lives Matter, and only 16 percent were unwilling to teach or support an anti-racist curriculum' (Kingkade, 2020).

However, it would seem then that while we do need to teach about racism and about Black Lives Matter, it must be in the context of White privilege. Students should not only be explicitly taught about colonialism but also about White people's acceptance of Whiteness as the norm. To pick up on what McIntosh said, we need to give White students training in seeing themselves as oppressors. This can be done in the classroom in the form of simulations followed by deep, frank discussions. If a school does not have Black staff, parents or members of the Black community need to be invited in to speak. Ironically, this is exactly the recommendation I made in my master's dissertation back in 1989: if you are teaching about First Nations people, get First Nations people to do the teaching. However, individual teachers in individual classrooms doing this sort of work is not enough, as has been shown. There need to be clear educational policies which embed teaching about colonialism in the context of Black History. Teaching about racism when someone has been murdered is not enough.

Curriculum

Black History Month has been observed in several countries since it was officially inaugurated by US President Ford in 1976 and it is celebrated in the US in February each year. In the UK, October is the designated month. However, the view expressed by many, for example those participating in Diverse Ed (Wilson, 2020), was that Black History should not just be taught in Black History Month. It should be a core element in the history national curriculum.

Black History is currently not an essential part of the curriculum in England. In June 2020 a government petition entitled *Teach Britain's colonial past as part of the UK's compulsory curriculum* was set up by Esmie Jikiemi-Pearson because, 'Currently, it is not compulsory for primary or secondary school students to be educated on Britain's role in colonisation, or the transatlantic slave trade' (Government Petition, 2020). The petition was signed by over 267,300 people and was therefore eligible for discussion in a parliamentary debate on 20 October 2020. The most recent government response (30 July 2020) to the petition is considered by many to be unsatisfactory.

> Within the history curriculum there is already a statutory theme at Key Stage 3 titled "ideas, political power, industry and empire: Britain, 1745-1901", as such we do not believe there is a need to take this action as the option to teach this topic exists within this compulsory theme.

In other words, it is going to be left up to individual schools and teachers to decide which, in practice, means many will not tackle the issue directly.

Creating a policy which ensures that Black History is embedded in the UK national curriculum (as well as in other countries' curricula) is an essential step in the right direction. Interestingly, in March 2021, the Welsh government voted to introduce Black History into the Welsh curriculum with effect from 2022. 'The education minister, said the changes would help children and young people become "ethical and informed citizens of Wales and the world"' (Lock, 2021).

Young Black students want to learn about Black History and they find it inspirational. Some powerful examples can be found in comments in a video on the Black Curriculum website (theblackcurriculum.com). The teaching of Black History is described as 'vital'. One young woman says, 'Not only does Black History bring representation in our society, but also it has the power to inspire those of us who cannot aspire' (The Black Curriculum, 2020). The UK Ambassador for Happiness, @ActionJackson, believes we must learn more about other people's history in general. 'We must respect, understand and nurture' (EduKit, 2020).

However, it is not only history that needs to be more diverse. Other curriculum subjects should be re-examined to ensure they include role models and examples from a wide variety of cultures, not just White, European or North American. Sufian Sadiq recently made the point that content in geography can be extremely Eurocentric. Instead of teaching about the Alps, the same concepts could be covered by teaching about the Himalayas, for example. He stressed that ticking the 'Inclusivity' box is not enough; 'Inclusivity must add value to the curriculum' (Wilson, 2020). The teaching of English too has long been challenged for focusing on 'dead White males' and, as a result, there have been improvements in this field over the last years with novels like *Things Fall Apart* (Achebe, 1994), *Half a Yellow Sun* (Adichie, 2007) and *Hullabaloo in the Guava Orchard* (Desai, 1999) becoming set texts on examination curricula. Music has traditionally cited predominantly male, White, European classical composers. Art has taken its examples from the West as well. Science and technology have largely ignored the huge contribution of BAME scientists in science and medicine. In short, the national curriculum of the UK, like many European and North American curricula, has celebrated the achievements of White people. Only in physical education have the sporting achievements of Black people been highlighted. While the feminist movement has led to more recognition of women in the sciences and the arts, multiracial representation is still not present.

Penny Rabiger of Lyfta[5] recently asserted that a diverse curriculum is about human storytelling. Students need to know about their fourth-generation ancestors. It is important to do this as it can help youngsters make sense of their many links and connections. She cautions that an intense look at the history curriculum and the textbooks associated with it is essential. Rabiger echoed Apple from many years earlier (already cited) when she asked, 'Who is being silenced? Who is being seen?' (Wilson, 2020). In a similar vein, Laila El-Metoui has stressed the need to be aware of the language used, for example, describing countries as being 'invaded' rather than 'conquered'. This recalls my own findings about social studies textbooks (already discussed).

Language is key in a bias-free, inclusive curriculum and is a major element of cultural competency. Concepts exist in one language that do not exist in another. Raising awareness of these can lead to empathy, critical thinking and compassion. It is part of cultural competency and goes hand-in-hand with diversifying the curriculum.

Representation

Representation matters. Improving representation and ensuring diversity at all levels of administration in statutory bodies would make a huge difference. 'Diversity should be the norm' as Yamina Bibi stated (Wilson, 2020). The UK statistics for BAME teachers are revealing. According to Javay Welter, a teacher in Birmingham, UK, there are only 0.7 per cent Black male teachers in schools in the UK (EduKit, 2020). Bibi, who describes herself as a visibly Muslim headteacher, believes the lack of diversity in staff is a question of representation at leadership level. Xavier-Cope stated that 96.5 per cent of headteachers in UK are White and 93.9 per cent of deputy heads. This is corroborated by the UK Department for Education website which recorded that in 2018, 'White British people accounted for 92.9% of headteachers, 89.7% of deputy or assistant headteachers and 85.1% of classroom teachers' (DfE, 2020). Murugesu (2020) likewise made this point about tribunals (academic disciplinary bodies) in universities.

Hannah Wilson, summing up the Diverse Ed session, wrote of the need for 'visibility of diversity' and questioned how we can increase diverse role models and amplify who has a voice. This is key because diversity is not only vital in achieving equality, it is also crucial for young people to have visible role models. A refrain echoing through the Diverse Ed session was, 'You cannot be who you cannot see' (Wilson, 2020). How can this be achieved? What policies need to be put in place to ensure there is as wide a representation of BAME people as possible? Proportionally, BAME people constitute nearly 13 per cent of the population of the UK. 'According to the Office for National Statistics, there were approximately 64.6 million people living in the UK in mid-2014. Of these, 56.2 million (87.2 per cent) were White British' (IRR, 2020). Realistically it could be difficult to replicate this 6:1 ratio across the country but there are ways to ensure diversity even in the 'White highlands'.

A point raised in a Q&A session recently was that rural schools have trouble attracting staff from BAME backgrounds. One solution that has been tried is to work in partnerships. In the UK, there are clusters of schools called multi-academy trusts. Within these, tiny rural schools may be partnered with inner city schools and can thus draw on a broader representation of staff. Another initiative being explored in Britain is the pioneering diverse governance series of webinars. Diverse Ed[6] are working in collaboration with four national organisations responsible for school governance in order to amplify diverse representation. This is definitely a step in the right direction and with more and more people presenting online, training in diversity and inclusion can reach a much wider audience in more remote areas. Diversity could become the norm. However, ensuring diversity may not be easy, even if measures are in place. A commenter in the chat box of a Black Lives Matter seminar made the point, 'Sometimes it's hard to have diversity in leadership roles … I know many people that shy away from the responsibility and express concerns about being treated equally if they are promoted' (native Eastender in EduKit, 2020). The policies for diversity have to go hand-in-hand with building respect and mutuality so BAME people can go for promotion without fear of retribution.

Another sector where policies requiring diversity could be effective is in the police force. As has been shown, it is felt that the MacPherson Enquiry of 1997 did not go far enough (Quinn, 2019); while overt racism had probably been reduced, covert racism and prejudiced attitudes in the police force are still common (Babu, 2019). According to Francis, Black boys are nine times more likely to be stopped than their peers (EduKit, 2020). Given the disproportionate number of young Black people who are stopped and searched as compared with White youths this would seem to be the case. So what can be done about it?

Having more diversity in police forces and recruiting more BAME police officers is highly likely to make a difference. Radical measures in this direction are being considered in some cities, for example Minneapolis in the US. Following the death of George Floyd, the city pledged to dismantle its police force and rebuild it (Searcey and Eligon, 2020). This is certainly one way to tackle the issue. However, it is not an easy task and this particular initiative now looks likely to founder. Herndon (2020) wrote, 'After a summer that challenged society's commitment to racial equality and raised the prospect of sweeping political change, a cool autumn reality is settling in'. Councillors are reneging on the pledges they made or saying the wording was not clear. There is disillusion and anger and a fear of the consequences of this indecision. Miski Noor of the Black Visions Collective is cited as saying, 'When these decisions are made on a political level, they have human consequences' (Herndon, 2020). The Minneapolis case illustrates just how difficult it can be to implement radical measures.

There is another way to tackle the challenge and that is to work with young Black people. Charles Francis, Safer Schools Officer with the London police force, has set out to address how we make sure policing is fair. He proposes trying to change Black youths' mindsets about how they view the police. One area he focuses on is people's reactions to stop and search. He points out that an officer must identify who they are and why they are stopping and searching the person. He encourages young people to challenge if they feel they are being stopped and searched unfairly. The incident must be on camera and usable as evidence; the police must be accountable. This measure could certainly be enshrined in policy. Another measure is to encourage students to join the police in order to effect change from within. He believes empowerment is key: 'You can be anything you want to be – aspire to leadership roles' (EduKit, 2020).

The situation within health services is perhaps more complex. What is needed, according to Public Health England, is 'a clear commitment for increasing diverse leadership at all levels in health and care system, reflecting the communities being served' (PHE, 2020a: 39). They also cite the need for cultural competency defining it as: 'the ability of providers and organisations to effectively deliver services that meet the social, faith, cultural, and linguistic needs of service users' (PHE, 2020a: 39). Headteacher Amardeep Panesar, who has been developing this in her school, describes it as enabling people to interact effectively by being aware of their own world view (Wilson, 2020). Policies that require personnel to have training in this area would go a long way to addressing the racism currently inherent in many sectors as well as at government level.

Conclusion

Policies need to be drafted and implemented with extreme care, taking into account the evidence available and the lessons learned. The Windrush Scandal, for example, has taught us that policies and legislation can lead to human tragedy. What can be done about it? If policies were partly to blame in the first place, what can be learned? In her report, Williams (2020) recommends that the Home Office first openly acknowledge and admit the injustice they have perpetrated. A commitment to transparency helps to encourage such open honesty. Second, there must be closer scrutiny of Home Office policies and legislation. When policies are implemented, mechanisms for their monitoring and evaluation should be included. Only then will evidence be collected to show what is working and what is not. In this way, researchers can work with policymakers to generate relevant evidence. And when evidence becomes available about policies, whether showing them to be effective or not, an ethical obligation exists to incorporate that evidence into policymaking.

Finally, and probably the most telling, if obvious, recommendation is the need to remember that, 'Migration and wider Home Office policy is about people and, whatever its objective, should be rooted in *humanity*' (Williams, 2020: 136, emphasis added). This latter point is crucial for policymaking and is a fitting place to end this chapter. It is where racism, White privilege and ingrained colonial attitudes must be cast aside in favour of recognising that we are all human, whatever our skin colour and origin. All policy should be rooted in humanity.

Notes

[1] In 1966 the British government made Section 11 Funding available to local authorities who, in the Secretary of State's opinion, 'are required to make special provision in the exercise of any of their functions in consequence of the presence within their areas of substantial numbers of immigrants from the Commonwealth whose language or customs differ from those of the community' (Local Government Act 1966).

[2] ESL was in itself a racist assumption, suggesting people learning English had only one other language. It was later replaced by EAL (English as an Additional Language) and ESOL (English to Speakers of Other Languages).

[3] Gaine went on to write *Still No Problem Here* in 1995 – clearly pointing out there was! – as well as *We're All White Thanks: The Persisting Myth About 'White' Schools* (2005).

[4] A simulation game similar to the *Blue eyes/brown eyes* experiment (described in Bloom, 2005) where participants are forced into power roles

[5] See https://www.lyfta.com/

[6] https://www.diverseeducators.co.uk/

References

Achebe, C. (1994) *Things Fall Apart*, New York: Anchor Books.

Adichie, C.N. (2007) *Half of a Yellow Sun*, New York: Anchor Books.

Akel, S. (2019) Insider-Outsider: The Role of Race in Shaping the Experiences of Black and Minority Ethnic Students, London: Goldsmiths University of London, available from: https://www.gold.ac.uk/media/docs/reports/Insider-Outsider-Report-191008.pdf [accessed 13 February 2022].

Apple, M. (1982) *Education and Power*, London: Routledge & Kegan Paul.

Associated Press (2020) 'French coronavirus study finds black immigrant deaths doubled at peak', The Guardian, 7 July, available from: https://www.theguardian.com/world/2020/jul/07/french-coronavirus-study-finds-black-immigrant-deaths-doubled-at-peak [accessed 13 February 2022].

Babu, D. (2019) 'I served in the Met. The lack of progress on diversity is disgusting', *The Guardian*, 21 February, available from: https://www.theguardian.com/commentisfree/2019/feb/21/served-met-diverse-force-discrimination-police-london [accessed 13 February 2022].

BBC (2019) 'Rwanda genocide: 100 days of slaughter', BBC News, 4 April, available from: https://www.bbc.com/news/world-africa-26875506 [accessed 13 February 2022].

BBC (2020) 'Myanmar Rohingya: what you need to know about the crisis', BBC News, 23 January, available from: https://www.bbc.com/news/world-asia-41566561 [accessed 13 February 2022].

Black Lives Matter (2020) 'About', Black Lives Matter, available from: https://blacklivesmatter.com/about/ [accessed 13 February 2022].

Bloom, S. (2005). Lesson of a lifetime. https://www.smithsonianmag.com/science-nature/lesson-of-a-lifetime-72754306/

Bradford NHS (2016) Race: Evidence of Health Inequalities Affecting Black and Minority Ethnic People, available from: https://www.bdct.nhs.uk/wp-content/uploads/2016/12/Race.pdf [accessed 13 February 2022].

Burns, S. and Lamont, G. (1995) *Values and Visions: A Handbook for Spiritual Development and Global Awareness*, Trowbridge: Hodder & Stoughton.

Cuncic, A. (2020) 'What is white privilege?' *verywellmind*, 25 August, available from: https://www.verywellmind.com/what-is-white-privilege-5070460 [accessed 20 May 2022].

Delpit, L.D. (1988) 'The silenced dialogue: power and pedagogy in educating other people's children', *Harvard Educational Review*, 58: 280–98, available from: https://newlearningonline.com/new-learning/chapter-8/lisa-delpit-on-power-and-pedagogy [accessed 15 July 2020].

Desai, K. (1999) *Hullabaloo in the Guava Orchard*, New York: Anchor Books.

DfE (Department for Education) (2020) 'School teacher workforce', GOV.UK, available from: https://web.archive.org/web/20201124174417/https://www.ethnicity-facts-figures.service.gov.uk/workforce-and-business/workforce-diversity/school-teacher-workforce/latest [accessed 13 February 2022].

EduKit (2020) 'EduKit 6th Annual TeachMeet: Black Lives Matter', [video] YouTube, 13 October, available from: https://www.youtube.com/watch?v=qEMRTYAYWAE&ab_channel=EduKit [accessed 13 February 2022].

EHRC (Equality and Human Rights Commission) (2019) *Tackling Racial Harassment: Universities Challenged – Executive Summary*, available from: https://www.equalityhumanrights.com/sites/default/files/tackling-racial-harassment-universities-challenged-executive-summary.pdf [accessed 13 February 2022].

Eligon, J. (2021) 'Black Lives Matter has grown more powerful, and more divided', *New York Times*, 4 June, available from: https://www.nytimes.com/2021/06/04/us/black-lives-matter.html [accessed 13 February 2022].

Gaine, C. (1987) *No Problem Here: A Practical Approach to Education and 'Race' in White Schools*, London: Hutchinson.

Gaine, C. (1995) *Still No Problem Here*, Stoke on Trent: Trentham Books.

Gaine, C. (2005) *We're All White Thanks: The Persisting Myth About 'White' Schools*, Stoke on Trent: Trentham Books.

Givetash, L., Radnofsky, C. and Fieldstadt, E. (2020) 'Video shows police in Kenosha, Wisconsin, shooting Black man in back', NBC News, 24 August https://www.nbcnews.com/news/us-news/video-shows-police-kenosha-wis-shooting-black-man-back-n1237819 [accessed 13 February 2022].

Godin, M. (2020) 'Black and Asian people are 2 to 3 times more likely to die of COVID-19, U.K. study finds', *Time*, 6 May, available from: https://time.com/5832807/coronavirus-race-analysis-uk/ [accessed 13 February 2022].

Government Petition (2020) 'Teach Britain's colonial past as part of the UK's compulsory curriculum', set up by Esmie Jikiemi-Pearson, London: UK Government and Parliament, available from: https://petition.parliament.uk/petitions/324092 [accessed 13 February 2022].

Griffith, J. (2020) 'Ahmaud Arbery shooting: a timeline of the case', NBC News, 12 May, available from: https://www.nbcnews.com/news/us-news/ahmaud-arbery-shooting-timeline-case-n1204306 [accessed 13 February 2022].

Herndon, A.W. (2020) 'How a pledge to dismantle the Minneapolis police collapsed', *New York Times*, 26 September, available from: https://www.nytimes.com/2020/09/26/us/politics/minneapolis-defund-police.html [accessed 13 February 2022].

IRR (Institute of Race Relations) (2020) 'Ethnicity and religion statistics', available from: https://irr.org.uk/research/statistics/ethnicity-and-religion/ [accessed 13 February 2022].

Kingkade, T. (2020) 'How one teacher's Black Lives Matter lesson divided a small Wisconsin town', NBC News, 24 October, available from: https://www.nbcnews.com/news/us-news/how-one-teacher-s-black-lives-matter-lesson-divided-small-n1244566 [accessed 13 February 2022].

Local Government Act 1966, available from: https://www.legislation.gov.uk/ukpga/1966/42/section/11/enacted [accessed 13 February 2022].

Lock, H. (2021) 'Wales makes Black History lessons compulsory in new school curriculum', *Global Citizen*, 19 March, available from: https://www.globalcitizen.org/en/content/wales-black-history-lessons-compulsory-school/?template=next [accessed 13 February 2022].

Lutz, E. (2020) 'Chicago: Black people are 30% of its population but 60% of its Covid-19 deaths', *The Guardian*, 24 May, available from: https://www.theguardian.com/world/2020/may/24/chicago-black-coronavirus-fatalities-us [accessed 13 February 2022].

McIntosh, P. (1989) *White Privilege: Unpacking the Invisible Knapsack*, available from: https://nationalseedproject.org/Key-SEED-Texts/white-privilege-unpacking-the-invisible-knapsack [accessed 13 February 2022].

Moncrieffe, M. (2020) '#BlackLivesMatter in education', Brighton Educational Research Association (BERA) [blog], 11 June, available from: https://www.bera.ac.uk/blog/blacklivesmatter-in-education [accessed 13 February 2022].

Murugesu, J. (2020) 'Universities are failing to address racism on campus', *Times Higher Education*, 1 October, available from: https://www.timeshighereducation.com/features/universities-are-failing-address-racism-campus [accessed 13 February 2022].

NHS (nd) Addressing impact of COVID-19 on BAME staff in the NHS, available from: https://www.england.nhs.uk/coronavirus/workforce/addressing-impact-of-covid-19-on-bame-staff-in-the-nhs/ [accessed 13 February 2022].

NPR (2020) 'We are repeating the discrimination experiment every day, says educator Jane Elliott', 8 July, available from: https://www.npr.org/sections/live-updates-protests-for-racial-justice/2020/07/08/888846330/ [accessed 13 February 2022].

NRC (National Research Council) (2004) *Understanding Racial and Ethnic Differences in Health in Late Life: A Research Agenda*, Washington, DC: National Academies Press.

PBS (nd) 'The buffalo: yesterday and today', PBS, available from: https://www.pbs.org/buffalowar/buffalo.html [accessed 13 February 2022].

PHE (Public Health England) (2020a) *Beyond the Data: Understanding the Impact of COVID-19 on BAME Groups*, gateway number: GW-1307, London: PHE.

PHE (Public Health England) (2020b) *Disparities in the Risk and Outcomes of COVID-19*, gateway number: GW-1447, London: PHE.

Pilkington, E. (2020) 'Black Americans dying of Covid-19 at three times the rate of white people', *The Guardian*, 20 May, available from: https://www.theguardian.com/world/2020/may/20/black-americans-death-rate-covid-19-coronavirus [accessed 13 February 2022].

Quinn, B. (2019) 'Macpherson report: what was it and what impact did it have?', *The Guardian*, 22 February, available from: https://www.theguardian.com/uk-news/2019/feb/22/macpherson-report-what-was-it-and-what-impact-did-it-have [accessed 13 February 2022].

Race Relations Act (1976) Chapter 74, available from: https://www.legislation.gov.uk/ukpga/1976/74/enacted [accessed 13 February 2022].

Ramlakhan, K. (2019) 'Black student carded, cuffed at University of Ottawa, prompting review', CBC, 14 June, available from: https://www.cbc.ca/news/canada/ottawa/university-human-rights-office-student-twitter-carding-1.5175864 [accessed 13 February 2022].

Rashid, I. (2020) 'Coronavirus: BAME people make up 72% of all NHS and carer deaths with COVID-19', Sky News, 22 April, available from: https://news.sky.com/story/coronavirus-bame-people-make-up-72-off-all-nhs-and-carer-deaths-with-covid-19-11977263 [accessed 13 February 2022].

Richardson, H. (2020) 'Black pupils face trebled exclusion rate in some areas of England', BBC News, 30 July, available from: https://www.bbc.com/news/education-53516009 [accessed 13 February 2022].

Searcey, D. and Eligon, J. (2020) 'Minneapolis will dismantle its police force, council members pledge', *New York Times*, 7 June, available from: https://www.nytimes.com/2020/06/07/us/minneapolis-police-abolish.html [accessed 13 February 2022].

Siddique, H. (2020) 'Psychiatrists urge new leader to rid profession of institutional racism', *The Guardian*, 1 July, available from: https://www.theguardian.com/society/2020/jul/01/psychiatrists-urge-new-leader-to-rid-profession-of-institutional-racism [accessed 13 February 2022].

Stripe, N. (2020) 'Coronavirus (COVID-19) related deaths by ethnic group, England and Wales: 2 March 2020 to 15 May 2020', London: Office for National Statistics, available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/coronaviruscovid19relateddeathsbyethnicgroupenglandandwales/2march2020to15may2020 [accessed 13 February 2022].

Tello, M. (2017) 'Racism and discrimination in health care: providers and patients', Harvard Health Blog, 16 January, available from: https://www.health.harvard.edu/blog/racism-discrimination-health-care-providers-patients-2017011611015 [accessed 13 February 2022].

The Black Curriculum (2020) 'Why Black British History Is Important', [video] YouTube, 11 June, available from: https://www.youtube.com/watch?v=kEhcbKOeQG8&feature=youtu.be [accessed 13 February 2022].

The Values and Visions Foundation (2020) 'Where did we go wrong? Three voices, shared ideas', Values and Visions Foundation, 15 June, available from: https://values-and-visions.com/2020/06/15/where-did-we-go-wrong-three-voices-shared-ideas/ [accessed 13 February 2022].

Thomas, D. and Horowitz, J.M. (2020) 'Support for Black Lives Matter has decreased since June but remains strong among Black Americans', Pew Research Center, 16 September, available from: https://www.pewresearch.org/fact-tank/2020/09/16/support-for-black-lives-matter-has-decreased-since-june-but-remains-strong-among-black-americans/ [accessed 13 February 2022].

Wikipedia (2020a) 'Bison hunting', available from: https://en.wikipedia.org/wiki/Bison_hunting [accessed 20 October 2020].

Wikipedia (2020b) 'Black Lives Matter', available from: https://en.wikipedia.org/wiki/Black_Lives_Matter [accessed 5 November 2020].

Wikipedia (2020c) 'The Holocaust', available from: https://en.wikipedia.org/wiki/The_Holocaust [accessed 23 November 2020].

Williams, W. (2020) *Windrush Lessons Learned Review*, Home Office, March, available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/876336/6.5577_HO_Windrush_Lessons_Learned_Review_LoResFinal.pdf [accessed 13 February 2022].

Wilson, H. (2020) 'Diverse Ed Live 17th October 2020', [video] YouTube, 17 October, available from: https://youtu.be/c98WrDvy1wM [accessed 13 February 2022].

Zammit, J.R. (2020) 'I haven't posted anything about black lives matter yet, but have spent a lot of time following links, new people', Facebook status update, 6 June, available from: https://www.facebook.com/profile.php?id=501553908 [accessed 13 February 2022].

17

The state of scientific policy advice: reflections on strategies and challenges of foreign policy think tanks in times of crisis

Nicolas Lux

Introduction

Recent years have been profoundly challenging for scientific research institutions, like think tanks, to provide policy advice to political elites. Many of the crises in the international order that policymakers as well as scientific policy advisors have been facing for several years are further deepened by the COVID-19 pandemic. The outcome of the 2016 Brexit referendum in the UK and the erratic administration of President Donald Trump in the US from 2017 to 2021 have called into question long-held foreign policy assumptions about ever-advancing globalisation and rules-based multilateralism. Unsettled by these developments as well as by the spread of 'fake news', conspiracy myths and populism in Western democracies, foreign policy think tanks have begun to discuss what effects are discernible on scientific policy advice. In the following, strategies will be discussed that reflect an ongoing debate in the field of foreign policy think tanks. This debate revolves around questions regarding the challenges think tanks see for their advisory work in view of an increasingly polarised political environment; how they should position themselves vis-à-vis the public and political sphere; and how they can maintain their independence and scientific integrity in these uncertain times.

There will first be a brief overview of and introduction to the landscape of foreign policy think tanks with a focus on Anglo-Saxon and European institutes. Next, current challenges that face think tanks will be considered, followed by an outline of possible forward strategies that are currently discussed by think tankers for their own professional future. These discussions are based on a selection of recent contributions by active senior think tank scholars (or 'think tankers') from Anglo-Saxon and European policy advising institutions via journal articles, blog posts or (recorded) conference speeches starting from 2017.[1] The conclusion summarises the main aspects of the ongoing reflection process that takes place within and among think tanks and will offer an outlook on how think tanks can be more resilient in view of current difficulties and future challenges.

An overview over the field of foreign policy think tanks

The understanding of what think tanks are and are supposed to do varies widely, as no uniform definition of them yet exists. Legally and practically, any organisation can decide to bear that label (Zuleeg, 2020: 5). However, what the role of a think tank usually should be, particularly in foreign affairs, can be generally understood as follows: think tanks conduct original, independent and scientific research, analysis and policy advice that aims to be relevant, multidisciplinary, application-oriented and fact-based, whether in publications, briefings or other formats. They primarily offer their work to governments, parliaments and other political decision-makers. Consequently, they try to help politicians gain a better understanding of international relations and security affairs and thus make informed policy decisions (for example Perthes, 2011; Brozus and Maull, 2017: 5; de Montbrial, 2019; Lux, 2021).

As at June 2021, a total of 12,664 different institutions were registered in the Transparency Register of the European Union (EU), which is jointly managed by the European Commission and the European Parliament.[2] Of these, 590 institutions identified themselves as 'think tanks and research institutions'. Although this self-classification appears to be relatively broad, this data query correctly indicates that the field of think tanks, among which exist many foreign policy institutes, presents a relatively crowded field for scientific policy advice. Moreover, the number of think tanks has continuously increased as shown for the case of Germany: between 2000 and 2020 the number of this kind of institute in foreign and security affairs doubled from 12 to 24 in Berlin alone (Bertram and Hoffmann, 2020: 14). The scene of scientific policy advice differs from country to country and from system to system regarding its structure and composition of its actors; it also ranges from think tanks and ad hoc commissions to governmental strategic planning units and the like (Hustedt and Veit, 2017: 42).

In this environment, many organisations labelled 'think tanks' compete with other knowledge providers, such as private consultancies, firms, banks and so on for the attention of decision-makers in proximity to their respective government and parliament as well as other influential actors (Drezner, 2017a). Consequently, the work of foreign policy think tanks is also about maintaining their own relevance in a competitive 'knowledge market' through a high degree of networking with important political players. Historical traditions present one part of the variations observed in these scenes across borders. Whereas, for instance, foreign policy think tanks from the

US and the UK can look back on a long tradition (especially after the end of the First World War), think tanks in proximity to the EU institutions, foremost in Brussels, have been able to establish themselves as relevant actors for politics only in recent decades, in the context of the deepening European integration process (Abelson, 2014: 136; Niblett, 2018: 1415). While according to the mainstream literature, US think tanks are much more involved in the process of policy formulation and are perceived as agents of social and political change (Niblett, 2018: 1414), think tanks in the EU context fill the comparatively more passive role of knowledge brokers and providers of (confidential) knowledge exchanges (Abelson, 2014; Kelstrup and Dialer, 2019).

Comparing the literature on think tanks working in the field of foreign and security affairs, several different approaches to their categorisation can be found. Often, the degree of independence or autonomy a think tank has from external funding and/or control is used to distinguish them from each other. McGann offers an approach to group think tanks into seven different categories by primarily considering the degree of a think tank's independence from governmental or political control – and which he uses within the framework of the annual *Global Go To Think Tank Index Report* at the University of Pennsylvania (McGann, 2021: 13–14; for example Balfour, 2017; Brozus and Maull, 2017: 5–7). This report creates a ranking to measure the performance of think tanks worldwide in a number of different categories. Criticism of this kind of distinction includes a US focus on the understanding of autonomy, which risks ignoring experiences with public versus private funding of think tanks from other parts of the world (Brozus and Maull, 2017).[3]

Other authors, such as Perez (2014: 329–30), Brozus and Maull (2017: 11–12) or Bertram and Hoffmann (2020: 12–17) rather focus on the function of a think tank as a distinguishing criterion. For instance, Perez classifies think tanks according to their function as academic, advocacy or so-called platform think tanks. This categorisation already makes clear that think tanks are advising politics in very different ways – be it through extensive original research, through advocacy work (so-called activist or do tanks) or as mere platforms or providers of (confidential) forums for knowledge exchange and networking. Moreover, those institutes with a considerable and public convening power can be also labelled part-time think tanks, as most of the resources are dedicated to networking events rather than to original research (Bertram and Hoffmann, 2020: 17). In addition to their function, foreign policy think tanks can also be grouped according to their geographical origin, as can be well illustrated by the example of the Brussels think tank landscape. In the EU's principal host city and the seat of NATO, four functional-geographic categories of think tanks can be identified: (1) EU policy think tanks with headquarters in Brussels; (2) 'national' think tanks, mainly from EU Member States with a Brussels liaison office; (3) globally active, US-based think tanks with a presence in Brussels; and (4) so-called 'in-house' think tanks of EU institutions. The influence and role of think tanks in Brussels differ from those, for example, in the US capital, Washington DC, as described by some scholars (Abelson, 2014: 136; Brozus and Maull, 2017: 11; Rastrick, 2017: 7-11; Bertram and Hoffmann, 2020: 25-7). Since the political system of the EU is profoundly different from that of the US – the latter has regular, partisan changes of government and as a result a high turnover of personnel in the administration after an election (the so-called 'revolving door' culture, Rastrick 2017: 4) – Brussels think tanks on the contrary usually do not function as a reservoir of senior experts being attached to one specific partisan bloc and from which the EU executive branch regularly would recruit new officials.

In order to show how the landscape of foreign policy think tanks within one political environment, here Brussels and the EU, vary in its organisational set-up, size and funding structure, Table 17.1 presents an overview of two think tanks for each of the four described categories with current key figures. For reasons of space, this overview cannot include a larger number of institutes. Thus, the selected think tanks each represent institutions that either continuously offer scientific contributions to foreign and security policy debates, facilitate the exchange between experts and decision-makers in this topic area through both public and confidential discussion formats, or whose research activities are sufficiently noted by other institutions as relevant. In this respect, the mention in the monthly 'Think Tank Review' of the General Secretariat of the European Council and the Council of the EU serves as a potential indicator of relevant attention that the scientific output of a think tank receives.[4] This selection is therefore considered important because of the Council's prominent role in the Common Foreign and Security Policy and Security and Defence Policy – an intergovernmental EU policy area – and as the Council's functioning is much less transparent and thus less accessible to think tanks compared to other EU institutions and policy fields (Shapovalova, 2019; Bertram and Hoffmann, 2020: 26-7). Another selection criterion is whether the respective think tank is registered in the EU Transparency Register. An entry in this database can make it easier for a think tank to gain access to the European Commission and the European Parliament, which usually should be in the interest of almost all active Brussels think tanks.

Times of crisis: what are the challenges for the work of foreign policy think tanks?

Researchers in foreign policy think tanks in Western countries have entered a process of reflection about their own role in the democratic policymaking

Table 17.1: Selection of Brussels think tanks in the field of foreign and security policy: current key figures until 2020

Category	Think tank's name and foundation year	Official no. employees in Brussels (year)	Annual budget (year)	Main sources of funding (rounded)	Subscription to the EU-Transparency Register	Number of mentions in the Council's Think Tank Review, 2019
1. EU policy think tanks based in Brussels	**European Policy Centre** (EPC), 1996	35* (2018)	€2,788,471 (2018)*	**Diverse funding:**+ 1. Cooperation and sponsoring partners: 55% 2. Membership fees: 24% 3. EU: 9%	Yes, 2009	26
	Centre for European Policy Studies (CEPS), 1983	78+ (2020)	€5,861,681 (2017)*	**Diverse funding:**+ 1. EU: 41% 2. Membership fees: 21% 3. Privately funded projects: 13%	Yes, 2011	44
2. Think tanks from EU Member States with presence in Brussels	**Stiftung Wissenschaft und Politik** (SWP), 1962 (Brussels Office since 2009)	3 (in total 172, 2018)*	€197,000 (for Brussels budget; total budget: €16,908,499; 2018)*	**Public funding:***+ 1. German federal budget: 82% 2. External project funding: 18%	Yes, 2014	30
	Egmont Royal Institute, 1947	21* (2018)	€1,152,749 (2018)*	**Public funding:***+ Belgian Federal Public Service Foreign Affairs: ca. 86 %	Yes, 2016	19
3. Global think tanks with presence in Brussels	**Carnegie Europe**, 2007	16* (2019)	€1,033,431 (2018/2019)*	**Foundations:*** 1. Carnegie Endow.: 40% 2. Foundation / private sector: 32%	Yes, 2015	8[a]
	International Crisis Group (ICG), 1995	44* (2019)	ca. €950,000 (for Brussels budget; total budget: €17,512,784; 2018/2019)*	**Diverse funding:**+ 1. Governments: 43% 2. Foundations: 31% 3. Private funding: 22%	Yes, 2009	12

Table 17.1: Selection of Brussels think tanks in the field of foreign and security policy: current key figures until 2020 (continued)

Category	Think tank's name and foundation year	Official no. employees in Brussels (year)	Annual budget (year)	Main sources of funding (rounded)	Subscription to the EU-Transparency Register	Number of mentions in the Council's Think Tank Review, 2019
4. In-house think tanks of EU institutions in Brussels	**European Political Strategy Centre** (EPSC, European Commission), 2014–19	56[+] (2018)	€704,435 (2018)	**Public funding:**[+] admin. budget,[b] European Commission	No (EU institution)	4
	European Parliamentary Research Service (EPRS, European Parliament), 2013	314[+] (2018)	€9,300,000 (2018)[+]	**Public funding:**[+] admin. budget,[c] European Parliament	No (EU institution)	25

Source: author's illustration.

Notes:

This research was originally conducted by the author as part of the project 'Research and Advice in Foreign and Security Policy. An Analysis of the German Think Tank Landscape' commissioned by Stiftung Mercator and Robert Bosch Stiftung (Bertram and Hoffmann, 2020).

a Of these, 3 at 'Carnegie Europe' and 5 at 'Carnegie Endowment for International Peace'.

b Without staff budget.

c Without staff budget.

* Data taken from the EU Transparency Register.

+ Data taken from the website of the respective think tank (all accessed: 10 February 2020), available from:

EPC: http://www.epc.eu/en/financing

CEPS: https://www.ceps.eu/wp-content/uploads/2019/06/CEPS_Annual-Report_2019.pdf

SWP: https://www.swp-berlin.org/ueber-uns/finanzierung/

Egmont: https://diplomatie.belgium.be/fr/Newsroom/actualites/communiques_de_presse/affaires_etrangeres/2014/09/ni_170914_egmont_institute

ICG: https://www.crisisgroup.org/how-we-work/financials

EPSC: https://ec.europa.eu/info/sites/info/files/epsc_aar_2018_final_0.pdf

EPRS: www.europarl.europa.eu/RegData/etudes/STUD/2019/637954/EPRS_STU(2019)637954_EN.pdf cycle. These thoughts, as showcased in a growing number of think tank publications, conferences or research projects, have intensified in the light of a crisis of the state of the international order and an increasing elite and science scepticism (for example Schenuit, 2017; Maull, 2018; Niblett, 2018). Lastly, beginning in early 2020, the COVID-19 pandemic has begun to ripple politics and societies globally as economies started to shrink, severe social restrictions were imposed and hence, also the work of think tanks has been greatly impacted (Biggins, 2020; Mendizabal, 2020; Zuleeg, 2020: 6).

A look at media contributions that covered the role of independent scientific political advice reveals the increasingly difficult position that many think tanks are faced with, as the following selection exemplarily demonstrates: *Politico* in 2017: 'Are think tanks doomed?' (Drezner, 2017b); *Deutschlandfunk Kultur* in 2019: 'The role of experts in the post-factual age' [in German: 'Die Rolle der Experten im postfaktischen Zeitalter'] (Gerstenberg, 2019); *The Economist* in 2019: 'Can think-tanks survive a post-fact world?' (R.C., 2019); and *The Guardian* in 2020: 'There's no such thing as just "following the science" – coronavirus advice is political' (Bacevic, 2020).

Several senior think tankers from European and Anglo-Saxon institutes have tried to pinpoint the current and future challenges that independent policy advice sees itself confronted with.

The crises in international relations amid rising populism and political polarisation

Many scholars see an increasingly difficult, even hostile, political and societal work environment for foreign policy think tanks that is a consequence of changing fundamentals of the state of international relations and the mainstream views think tanks had on them.

As pointed out openly by several authors, think tanks find themselves confronted with public distrust that stems from their past, often uncritical support of an ever-advancing globalisation amid an increasingly polarised political landscape and rising social-economic inequalities in Western countries (Niblett, 2018: 1419; Gomart, 2019: 58). In this respect, think tankers have only been rarely critical of potential negative repercussions of the effects of a globalised world, in recent decades. Consequently, political shifts intruded that were in stark opposition to the majoritarian 'internationalist' opinions of experts on the state of international relations. Two of these symptoms can be found, first, in the Brexit referendum of 2016 leading to the UK's withdrawal from the EU in 2020 and, second, in the 2016 election victory and subsequent presidency of Donald Trump in the US. Populism in political debates and a growing science scepticism that goes hand in hand with a perceived politicisation of the role of scientific research have thus far changed the work of foreign policy think tanks, within just a few years (Rastrick, 2017; Haass, 2019; Brozus, 2020a). Nowadays, think tankers are confronted at home with new questions that revolve around the real benefits of globalised societies and economies, and around cultural identity and social participation (Niblett, 2018: 1420).

A changing, more competitive 'knowledge market'

Several authors have focused on how foreign policy think tanks compete with other (private) actors in a crowded 'knowledge market', respectively the 'market [place] of ideas' (Husted and Veit, 2017: 42; Balfour, 2017: 8–9; Gomart, 2019; Lux 2021: 2).

This market has been changed profoundly as Gomart of the French Institute for International Relations (IFRI) has pointed out (2019: 58–60). However, technological shifts (such as in the use of big data) are not to be seen as the main drivers of change of the marketplace. Rather, it is the observed polarisation of the think tanks' work environment that is now increasingly occupied by non-scientific knowledge providers and by the use of facts and truth as mere social constructs (Gomart, 2019: 59). The latter development invites opinion leaders to use the spread of conspiracy theories and the manipulation of facts, such as fake news (Balfour, 2017: 6). Since think tanks, according to Balfour, then with the German Marshall Fund of the US in Brussels, do not have their own political-social constituency that could provide them with a certain legitimacy, they run the risk of being defenceless against criticism from the populist side. There, think tanks would be portrayed as closed echo chambers that contribute to foreign policy remaining a purely elite affair (2017: 8). This marketplace is further polarised by economic factors, as financial resources are distributed unevenly, benefiting established think tanks with international influence (Gomart, 2019: 60). Another factor negatively affecting the think tank community is that authoritarian regimes began to deliberately establish and fund their own think tanks in order to build up soft power – like China – while their counterparts in Western democracies all too often receive less financial support from their own governments. In this respect, Beijing has recently invested large resources in the creation of own research institutes with the goal of competing, especially with American think tanks (2019: 56–8).

Further reasons for concern for think tanks are their own deficiencies as perceived by their target groups or by themselves. Think tank–based scholars have scrutinised common allegations against their own industry that address a lack of diversity among think tank staff (Balfour, 2017; Islam, 2018), as well as of integrity and transparency in terms of funding and potential (foreign) influence (Niblett, 2018; Gomart, 2019).

A changed relationship vis-à-vis politics and the public

Amid a seemingly difficult state of the political international order and a changing, more challenging knowledge market, foreign policy think tanks face an altered relationship between them and political actors and the society.

A profoundly changed relationship between think tankers and their target groups – first, politics, but also academia, the media and the public – has been identified by numerous scholars – albeit with a different focus (Niblett, 2018; Gomart, 2019; Brozus, 2020a). Niblett (2018: 1418) of Chatham House in the UK sees the emergence of new communication technologies and thus habits as one key changer of the think tank–politics relationship. In a period of excessive supply of news and opinions in real time, analyses by think tanks seem to receive less attention from their target group if they do not focus their output on daily developments. The strength of a think tank, which is to analyse larger contexts through extensive research, seems no longer sufficiently relevant. Moreover, Brozus (2020a) of the German Institute for International and Security Affairs (SWP) and Richard N. Haass (2019) of the Council on Foreign Relations in the US see the danger of a politicisation of the think tank's work. Whereas the former outlines concerning trends such as 'weaponising science', a process in which the scientific work is being politicised in partisan opinion battles, the latter takes aim at the pre-commitment of many think tankers whether or not to question long-held first-order issues of international relations. Also, the recruiting of new think tank personnel seems to be all too often tailored to existing policy beliefs that in turn could nourish the impression of a politicised policy advice industry in the public and the media alike, as Haass points out.

In addition, the risk of a decreasing relevance of think tanks in the political system can be attributed to an over-specialised thematic narrowing of think tank research, as other scholars, such as Morillas (2019) from the Barcelona Centre for International Affairs (CIDOB), argue. Morillas perceives two risks in the current political environment for the sector's relevance, simultaneously: a disconnection of policy advice from society and from politics alike. This puts the very purpose of think tanks at risk, as he understands it, to serve as an interface between expert knowledge and political decision-making processes.

How to move forward: current discussions on strategies vary widely

Questions addressing the current problems and challenges that face foreign policy think tanks lead many scholars to reflect about long-term strategies on how to improve and better adapt their profession for the future. At the centre of the discussion on forward strategies and possible solutions, contributions see the relationship of those institutes with political and the societal actors as important aspects that they need to address.

Think tankers such as Islam (2018), then with Friends of Europe in Brussels, advocate for better communication and outreach – with an emphasis on social media where illiberal narratives and fake news currently dominate. This is in line with the recommendation of others, like Brockmeier of the Global Public Policy Institute (GPPi) from Germany and Nitzschke (2017: 32) with the Planning Unit of the German Federal Foreign Office (in private capacity), who not only see a need for using clear language when think tankers disseminate their work, but who in more general terms advocate for a popularisation of the debates on foreign policy – a demand shared by other scholars such as Haass (2019) and Balfour (2017: 10). The latter argues for an end to foreign policy debates being regarded as 'elite's business' only. Think tanks in the field of international relations need to acknowledge the rising interest in these debates among the general population that has been observed in surveys (Brockmeier and Nitzschke, 2017: 31). Hence, think tanks should leave the 'expert echo chambers' of traditional foreign policy circles that are perceived as closed shops and rather embrace the benefits of putting controversial and normative foreign policy issues up for discussion with citizens (2017: 31, 35). Consequently, think tankers need to reconsider their role profile as they should no longer be just opinion makers, but also facilitators of politico-societal debates in order to improve the acceptance of their profession. In respect to increased acceptance and being more responsive to social changes, many scholars agree on the need for more ethnic, social and educational diversity among the staff of foreign policy think tanks (Balfour, 2017: 12; Islam, 2018; Niblett, 2018: 1425). Balfour (2017: 10) also emphasises the importance of career opportunities, especially for younger colleagues, as well as the use of innovative research methods such as collaborative and interdisciplinary approaches that also considers an extended target group. This plea goes in line with the recommendations offered by Kefferpütz of the Mercator Institute for China Studies (MERICS) and Leifert of the German Council on Foreign Relations (DGAP; Kefferpütz and Leifert, 2021). They argue that think tankers, who are in the political process uniquely positioned at an 'intersection of academic knowledge, political power, and civil society interests', can better maintain their relevance by promoting innovation and including non-traditional stakeholders.[5]

The question of the interdisciplinarity of scientific policy advice through think tanks is of relevance for other scholars such as Thränert (2020) of the Center for Security Studies at ETH Zurich – albeit with a different focus concerning the target group. Against the backdrop of the early days of the COVID-19 pandemic, Thränert suggests that foreign policy think tanks should establish closer links to their political clients. Although think tanks are not democratically legitimised themselves, the democratic decision-making in politics is in constant need of scientific and interdisciplinary advice from outside. To improve this function, think tankers ought to position themselves closer to politicians and even to work directly for them at times as was traditionally the case in the US due to the 'revolving door' culture (*Drehtürkultur*). A desired outcome of a closer link to politics should be that both sides can develop a deeper understanding of each other which in turn could help think tankers to better tailor their advisory work as it is needed in political decision-making. To achieve this outcome amid an unforeseen pandemic, where think tankers have at the beginning only partly been able to provide useful policy advice, more practical training of researchers as well as more interdisciplinarity between natural, social and political sciences should be achieved.

Contrary to previous contributions advocating for a popularisation of foreign policy debates by think tanks or arguing for a more intermeshed relationship between them and political decision-makers, some rather plead for keeping a certain distance while being more conscious about the own role. Brozus (2020a) argues that an intensified exchange with the public or the establishment of a revolving door between think tanks and politics could be rather counterproductive in the age of political polarisation. Instead, think tankers should better reflect about their political role in the current system that sees an increasing politicisation of scientific arguments in political opinion battles as seen clearly in the US during the COVID-19 pandemic. Other prominent think tankers such as Haass (2019) also see a benefit in keeping a certain distance from politics as busy government schedules usually do not allow officials to thoroughly consider think tank policy advice, in the first place. Think tankers should rather invest more resources in conceptual research that aims to address the next generation of decision-makers and this way potentially influencing the path of future policy debates in international relations. Furthermore, he agrees with Bertram (former director of SWP) and Hoffmann (2020: 33) (journalist with *Der Spiegel*), as the latter call for resisting uniformity in questions of a think tank's research agenda or the recruitment of personnel. Think tanks should constantly review their own research priorities and encourage alternative approaches (Bertram and Hoffmann, 2020) as well as continue to question and explain first-order issues of international relations anew (such as why NATO and free trade are something 'good', Haass, 2019). This way the acceptance of a think tanker's policy advice can be increased among the respective target group. A further politicisation of the work of think tanks could thus be avoided. To support that goal, Brozus (2020a) concludes that think tankers must guarantee transparency of funding, enforce rigid quality measures and promote multiperspectivity in their research agenda. Additionally, they should learn to be honest and clear about their own (perceived) shortcomings and errors in order to eventually protect their institutional and individual independence, credibility and relevance.

The importance of adhering to a clear ethical framework for think tanks is discussed by Zuleeg (2020) of the Brussels-based European Policy Centre (EPC). Embedded in the 'Promoting Ethics and Integrity in Non-Medical Research' (PRO-RES) project,[6] his research advocates for such a normative framework that – unlike in traditional academic research – does not yet exist for non-academic research as it happens in think tanks (2020: 4). As in general, think tanks seek to influence democratic policymaking processes by offering their scientific advice, the potential danger that unethically acting advisors could pose – be it through manipulation of research for political or other reasons – to the legitimacy of the whole sector is evident (2020: 5). In order to preserve the legitimacy and integrity of policy advice through think tanks, Zuleeg identifies a possible normative incentive of the sector by jointly commencing a process that aims to formulate an accepted set of ethical standards for research and work through following certain basic principles: first, a think tank should be independent in its research and be able to demonstrate it; second, it should provide transparency about their operations and funding; third, it should have an adequate (good) governance structure that enables it to implement its ethical commitments; and lastly, it should follow a multi-stakeholder approach that helps the think tank to ultimately act as a bridge between different political and societal actors, 'including those that might struggle to engage effectively without the facilitation of think tanks' (Zuleeg, 2020: 8). The framework itself should be neither rigid nor legalistic, but rather provide incentives for think tankers to strive for ethically sound practices in the work of scientific advice and research. As a consequence, think tanks would agree to only use the research that was undertaken ethically. Said framework should be jointly developed from the bottom up, possibly within an institutionalised setting like a 'European Alliance of Independent Think Tanks'. Such an association could eventually be developed into a capacity-building effort for the whole sector, including financial support tools through improved joint access to external funding (Zuleeg, 2020: 10). The latter aspect could prove valuable for an industry that is most likely to be negatively affected due to the COVID-19 crisis. Unethical behaviour could ultimately be sanctioned by excluding actors from their rank. In conclusion, the think tank sector could – even amid politically as well as economically challenging times – preserve its legitimacy through following ethical guidelines which in turn should 'provide as few points of attacks to their new political opponents as possible' (2020: 8).

Conclusion

How are we to achieve more resilience in an age of crises? Through ethical standards, integrity and constant reflection about our own role.

The landscape of foreign policy think tank work is diverse – in its set-up as already shown, but also regarding the forward strategies discussed in this chapter. The perception of the current problems, however, such as the spread of fake news and conspiracy myths in political debates, a polarisation and politicisation of the field of scientific advice, the rippling effects of the COVID-19 pandemic, but also its own (perceived) shortcomings such as a lack of transparency, and internal and topical diversity, have been assessed quite similarly by the scholars mentioned in this chapter. However, what is lacking in these contributions – despite a common view of an increasingly difficult state of international relations and domestic politics – are objective criteria to measure and operationalise the difficulty level and thus make it possible for think tankers to better compare and to analyse the challenges in different regional, national or policy contexts. Consequently, some questions remain open: is the situation for all or most think tanks as difficult given the international context; if not, why? How and why might they actually perceive differences concerning the state of international relations and the political system they are working in? And how and to what extent do structural variations, for example different funding models and mission statements, come into play, when think tanks try to find the right answers to the problems that were discussed?

Looking at the discrepancies of possible solutions for think tanks in dealing with multiple crises, noticeable differences come to the foreground. The discussed forward strategies prominently address, first, the question of whether or not promoting a popularisation of the foreign policy debate versus striving for a closer relationship to political elites or, instead, keeping a certain distance from politics and society. Second, the question of legitimacy, integrity and ethical conduct are seen as central by many think tank scholars while reflecting about maintaining (and improving) the relevance of their own sector in a challenging socio-economic environment. This concurs with the chapters in this volume, which consistently identify ethical commitments as central to the use of evidence in policymaking. Think tanks play an important role in encouraging and supporting policymakers to base policy on ethically generated evidence and avoid the use of fake or distorted claims. Related to the questions of the legitimate role of think tanks in a democratic policymaking process is the real or perceived impact and the performance of think tanks in foreign affairs. Numerous publications in the mainstream literature have already addressed this issue with similar conclusions as Abelson (2014: 142), who – focusing on foreign policy think tanks in North America and Europe – attested to the lack of meaningful data to actually and objectively measure their influence.[7]

The deeper-running issue regarding a think tank's legitimacy in the political system, however, pertains to the acceptance and relevance of their work that is attributed by their peers and target groups. Calls from within the sector to diversify the workforce in order to make it more responsive and sensible to changes in society as well as a demand for full transparency about their funding and working operations could certainly help them to build up and maintain their credibility in times of politicised opinion battles. However, whether a think tank is actually supposed to contribute to a further popularisation (or democratisation?) of foreign policy debates or whether it should rather intensify its (closed shop?) relations to political elites, can only be decided on a case-by-case basis, given the heterogeneity of the field of independent think tanks. For some, only one way might be in line with their own understanding of how a think tank should provide research and analysis in foreign affairs, yet others might come to different conclusion(s). Self-consciousness and a transparent communication among each individual think tank about the mission and the understanding of their own role in politics and society could prove valuable, so expectations from within and from outside the think tank community can be better managed and eventually matched.

More fundamentally however, the legitimacy of the think tank sector needs to be further debated in regard to questions of scientific and ethical integrity. This ties into the general call of this volume for 'ethical evidence'. As discussed by Zuleeg (Chapter 6 in this volume), a drive for a common understanding and implementation of a normative binding set of ethical standards for the work of think tanks could crucially contribute to the sector's reputation as well. The discussion within and among think tanks has just begun. Whether during the course or at the end of this process a sort of 'ethical quality label' or an institutionalised set-up like a European alliance can be the right answer, certainly deserves more thorough consideration.

In other aspects of their work, such as their research methods, think tanks should consistently reflect and where necessary evolve. Innovative analysis tools such as *Foresight* exercises might help think tankers and policymakers alike be better prepared for the next '(un)expected crisis' (for example Brozus, 2020b). A thorough and honest stocktaking that reflects on real or possible methodological and operational shortcomings cannot be limited to a one-off exercise 'until the current crisis is over'. A process of constant self-evaluation within the institutes, with peers from the think tank community and with outsiders who are relevant to the work of scientific policy advice is necessary for think tankers to adapt to new challenges and sustain and improve their research and analysis.

To that end, a number of thought-provoking practical exercises have already begun to enrich the discussion process among think tankers. Aside from the ideas outlined by Zuleeg on an ethical framework, a comprehensive stocktaking of a national think tank landscape in foreign and security affairs was recently conducted by Bertram and Hoffmann (2020) for Germany that could be replicated internationally, and for other countries and regions as well. Moreover, multi-annual projects dedicated to serve as reflection and innovation platforms, such as the German 'Think Tank Lab'[8] starting in 2021, could offer creative space for training, exchange and discussions about new solutions and best practices for think tankers in the future.

These efforts through scientific articles and opinion pieces as well as through ad hoc or permanent discussion forums can further incentivise think tankers to continuously reflect on their work and impact in politics and society. This involves a commitment to working with integrity in how evidence is selected and communicated to policymakers, along with a commitment to take new paths if current methods seem no longer sufficient. Necessary changes within a think tank could be facilitated through a more diverse and thus, possibly, more creative think tank staff. By maintaining good scientific and ethical standards, by encouraging policymaking based on such ethical evidence, by being more responsive to social developments that take place outside the (closed) foreign policy circles, and by welcoming innovative solutions in policy advice, think tanks can continue to play their scientific advisory role with integrity in a democracy, in uncertain times.

Notes

[1] Transparency disclosure: parts of the underlying research for this chapter stem from an original *SWP Journal Review* written by the author (Lux, 2021).

[2] Available from: https://ec.europa.eu/transparencyregister/public/homePage.do?redir=false&locale=en#en [accessed: 28 June 2021].

[3] The latest ranking for 2020, for example, counts globally 11,175 'think tanks' (McGann, 2021: 43). Unresolved methodological and procedural issues raised by individual think tanks limit the scientific validity of this survey as claimed by some (SWP Berlin, 2 February 2021: 'Think Tank Index: SWP abstained from evaluation', available from: https://www.swp-berlin.org/en/media/detail/swp-hat-sich-abermals-nicht-am-global-go-to-think-tank-index-beteiligt [accessed: 16 May 2021].

[4] Available from: https://www.consilium.europa.eu/de/documents-publications/library/library-blog/think-tank-review/ [accessed: 16 May 2021].

[5] Yet other scholars argue for the need to rethink the theoretical models of the politics–think tank relationship, for example, by applying the model of the *Science and Technology Studies* (Schenuit, 2017).

[6] PRO-RES is a European Commission-funded project aiming to PROmote ethics and integrity in non-medical RESearch. see: https://prores-project.eu/ [accessed: 16 May 2021].

[7] However, as he added on an anecdotal basis, probably every director of a think tank would describe the impact of their own institute as quite large in most cases.

[8] This project is supported by the DGAP and MERICS, established in 2021, available from: https://dgap.org/de/forschung/programme/impact-innovation-lab/think-tank-lab [accessed: 16 May 2021].

References

Abelson, D.E. (2014) 'Old world, new world: the evolution and influence of foreign affairs think-tanks', *International Affairs*, 90(1): 125–42.

Bacevic, J. (2020) 'There's no such thing as just "following the science" – coronavirus advice is political', *The Guardian*, 28 April, available from: https://www.theguardian.com/commentisfree/2020/apr/28/theres-no-such-thing-just-following-the-science-coronavirus-advice-political [accessed: 24 June 2021].

Balfour, R. (2017) 'What are think tanks for? Policy research in the age of anti-expertise: what role can and should think tanks play in a 'post-truth' world?' Strategic Update 17.7, December, available from: https://www.lse.ac.uk/ideas/publications/updates/think-tanks [accessed: 16 May 2021].

Bertram, C. and Hoffmann, C. (2020) 'Research and advice in foreign and security policy: an analysis of the German think tank landscape', Stuttgart: Robert Bosch Stiftung GmbH and Essen: Stiftung Mercator GmbH, available from: https://www.bosch-stiftung.de/en/publication/research-and-advice-foreign-and-security-policy [accessed: 16 May 2021].

Biggins, E. (2020) 'The future of think tanks in a post-Covid-19 world: Looking to (social) science for the answers', On Think Tanks: Opinion, 18 May, available from: https://onthinktanks.org/articles/the-future-of-think-tanks-in-a-post-covid-19-world-looking-to-social-science-for-the-answers/ [accessed: 16 May 2021].

Brockmeier, S. and Nitzschke, H. (2017) 'Mehr Marktplatz, weniger Papier: Think Tanks sollten stärker den Dialog mit der breiteren Öffentlichkeit suchen', *Internationale Politik*, 6: 30–5, available from: https://internationalepolitik.de/system/files/article_pdfs/ip_06-2017_brockmeier-nitzschke_s-30-35_o.pdf [accessed: 16 May 2021].

Brozus, L. (2020a) 'Politikberatung: nicht unpolitisch, aber distanziert', SWP: Kurz Gesagt, 7 September, available from: https://www.swp-berlin.org/publikation/politikberatung-nicht-unpolitisch-aber-distanziert/ [accessed: 16 May 2021].

Brozus, L. (2020b) 'Preparing for the crises after COVID-19: systematic foresight as a tool for evidence-informed precautionary policy', SWP 2020/C 51, 6 November, available from: https://www.swp-berlin.org/en/publication/preparing-for-the-crises-after-covid-19/ [accessed: 16 May 2021].

Brozus, L. and Maull, H.W. (2017) 'Think tanks and foreign policy', in W.R. Thompson (ed) *Oxford Research Encyclopedias*, Oxford: Oxford University Press, available from: https://doi.org/10.1093/acrefore/9780190228637.013.385 [accessed: 16 May 2021].

de Montbrial, T. (2019) 'Quel rôle pour les think tanks?', Radio France Internationale: Géopolitique, le débat, 19 April, available from https://www.rfi.fr/fr/emission/20190421-think-tanks-role-technologies-ifri-chine-moyen-orient-climat-conflits-menaces [accessed 14 February 2022].

Drezner, D.W. (2017a) 'The traditional think tank is withering. In its place? Bankers and consultants. How the private sector is taking over policymaking', Washington Post, 4 June, available from: https://www.washingtonpost.com/posteverything/wp/2017/04/06/the-traditional-think-tank-is-withering-in-its-place-bankers-and-consultants/ [accessed: 4 February 2021].

Drezner, D.W. (2017b) 'Are think tanks doomed?' Politico, 30 August, available from: https://www.politico.com/magazine/story/2017/08/30/are-think-tanks-doomed-215560/ [accessed: 24 June 2021].

Gerstenberg, R. (2019) 'Die Rolle der Experten im postfaktischen Zeitalter', Deutschlandfunk Kultur, 19 August, available from: https://www.deutschlandfunkkultur.de/wissenschaftler-als-politikberater-die-rolle-der-experten.976.de.html?dram:article_id=456651 [accessed: 24 June 2021].

Gomart, T. (2019) 'Changes in the think tank industry', in T. de Montbrial and T. Gomart, *What Is a Think Tank? A French Perspective*, Paris: IFRI, pp 55–61.

Haass, R.N. (2019) 'A century of think tanks', Council on Foreign Relations, 6 May, available from: https://www.cfr.org/event/century-think-tanks [accessed: 16 May 2021].

Hustedt, T. and Veit, S. (2017) 'Policy advisory systems: change dynamics and sources of variation', *Policy Sciences*, 50(1): 41–6.

Islam, S. (2018) 'Why think tanks matter in the era of digital and political disruptions', Bruegel, 30 January, available from: https://www.bruegel.org/events/why-think-tanks-matter-in-the-era-of-digital-and-political-disruptions/ [accessed: 16 May 2021].

Kefferpütz, R. and Leifert, C.L. (2021) 'Why German think tanks have to change the way they work', DGAP Online Kommentar, 21 June, available from: https://dgap.org/de/forschung/publikationen/why-german-think-tanks-have-change-way-they-work [accessed: 30 June 2021].

Kelstrup, J.D. and Dialer, D. (2019) 'EU think tank lobbying strategies: between theoretical advice and practice', in D. Dialer and M. Richter (eds) *Lobbying in the European Union: Strategies, Dynamics and Trends*, Cham: Springer, pp 445–59.

Lux, N. (2021) 'Foreign policy think tanks in times of crisis: how do they see themselves and their relevance? Contributions to the debate since 2017', *SWP Journal Review*, 1, available from: https://www.swp-berlin.org/10.18449/2021JR01/ [accessed: 16 May 2021].

Maull, H.W. (ed) (2018) *The Rise and Decline of the Post-Cold War International Order*, Oxford: Oxford University Press.

McGann, J.G. (2021) *2020 Global Go To Think Tank Index Report*, Think Tanks and Civil Societies Program (TTCSP), 28 January, available from: https://repository.upenn.edu/think_tanks/18/ [accessed: 16 May 2021].

Mendizabal, E. (2020) 'How COVID-19 is affecting funding, and what think tanks and donors can do about it', On Think Tanks: Opinion, 6 May available from: https://onthinktanks.org/articles/how-covid-19-is-affecting-funding-and-what-think-tanks-and-donors-can-do-about-it/ [accessed 14 February 2022].

Morillas, P. (2019) 'Grappling with dissent: how can think tanks cope with politicisation?', CIDOB: Barcelona Centre for International Affairs, Opinión 577, May, available from: https://www.cidob.org/en/publications/publication_series/opinion/seguridad_y_politica_mundial/grappling_with_dissent_how_can_think_tanks_cope_with_politicisation [accessed 14 February 2022].

Niblett, R. (2018) 'Rediscovering a sense of purpose: the challenge for Western think-tanks', *International Affairs*, 94(6): 1409–29.

Perez, M. (2014) 'Does EU policymaking allow for skilful networkers but limited knowledge managers? The think tanks' tale', *International Journal of Politics, Culture, and Society*, 27(3): 323–42.

Perthes, V. (2011) 'Scientific policy advice and foreign policymaking – Stiftung Wissenschaft und Politik (SWP), the German Institute for International and Security Affairs', in J. Lentsch and P. Weingart (eds) *The Politics of Scientific Advice: Institutional Design for Quality Assurance*, Cambridge: Cambridge University Press, pp 286–94.

R.C. (2019) 'Can think-tanks survive a post-fact world?' The Economist, 29 May, available from: https://www.economist.com/open-future/2019/05/29/can-think-tanks-survive-a-post-fact-world [accessed: 24 June 2021].

Rastrick, C.J. (2017) Thinking *A*bout Trump: American *T*hink *T*anks and *T*heir *N*ew *P*olitical *R*eality, On Think Tanks Working Paper 1, August, available from: https://onthinktanks.org/wp-content/uploads/2017/08/Rastrick_ThinkingAboutTrump.pdf [accessed: 16 May 2021].

Schenuit, F. (2017) Modelle wissenschaftlicher Politikberatung auf dem Prüfstand: Impulse für die Politikwissenschaft aus den Science and Technology Studies, SWP: Arbeitspapier FG EU/Europa, No. 03, April, available from: https://www.swp-berlin.org/publications/products/arbeitspapiere/Modelle_wissenschaftlicher_Politikberatung_FS.pdf [accessed 14 February 2022].

Shapovalova, N. (2019) 'Lobbying and interest group influence in EU foreign policy', in D. Dialer and M. Richter (eds) *Lobbying in the European Union: Strategies, Dynamics and Trends*, Cham: Springer, pp 425–34.

Thränert, O. (2020) 'Politikberatung in Corona-Zeiten', *Policy Perspectives*, 8(2), available from: https://www.research-collection.ethz.ch/handle/20.500.11850/411961 [accessed 14 February 2022].

Zuleeg, F. (2020) *An Ethical Framework for Think Tanks: Easier Drafted Than Done?* European Policy Centre Discussion Paper, 2 September, available from: https://www.epc.eu/content/PDF/2020/An_ethical_framework_for_think_tanks_v3.pdf [accessed: 16 May 2021].

18

Increasing the impact of research through policy: the role of academic publishers in bringing researchers and policymakers together

Hazel Goodes and Sarah Broadley

Introduction

Academic publishers see themselves as part of the academic community, serving the research conversation by coordinating peer review, and aggregating and disseminating scholarly work. University presses are generally recognised as departments of the institutions to which they belong, and many other publishing houses were founded by academics. It is in the interests of publishers to work hard to stay in touch with research communities and aim to be champions for the disciplines they publish in. The industry has spent around 350 years honing its role as the communication hub of the academy, facilitating the research conversation.

In recent years, universities have been working harder to demonstrate how their work serves to improve people's lives. In the UK, the first formal assessment of research quality was the Research Assessment Exercise (RAE), which started in 1986. This was replaced by the Research Excellence Framework (REF), which issued its first report in 2014. One of the ways the REF evaluated institutions was based on their impact, defined as 'an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia' (Research England, 2015). Global initiatives such as the United Nations Sustainable Development Goals (SDGs) call on universities to refocus their efforts on education, research and leadership that 'underpin and support the implementation of the SDGs by the global community' (SDSN Northern Europe, 2020). The research community has been driving action on community impact, too. The San Francisco Declaration on Research Assessment (DORA) requires that signatories move beyond inward-looking metrics, focused on restricted forms of output, such as the Journal Impact Factor. It calls on institutions to 'consider the value and impact of all research outputs (including datasets and software) in addition to research publications, and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice' (DORA, 2020).

The authors of this chapter are both employed by Emerald Publishing, an independent academic publisher based in Bingley, UK. The company was founded in 1967 by a group of academics from the School of Management at the University of Bradford, UK. Their aim for the company was that it should help to bridge the gap between research and practice in the field of management studies. The press did this by developing a portfolio of peer-reviewed journals in applied subjects, focusing on research that had clear, practical outcomes for practitioners. Emerald has evolved over the years and now publishes across a range of business and social science disciplines and in a variety of formats, from journals and ebooks to podcasts and blogs. Throughout, the company has followed its core ethos of helping to put research into practice, but, as part of the academic community, it has also adapted to meet – and *show* that it meets – the calls for greater focus on the benefits of research beyond academia, by trying also to disseminate its authors' work to non-academic audiences.

This chapter describes the efforts Emerald has made to reach one of those non-academic audiences: policymakers. The work described here was carried out by a team of volunteers from around the business, who were excited to give research a better chance of making a difference in the real world. While the content of this chapter is based only on work done by this one medium-sized, independent academic press, the team included people with many years of experience at different academic publishing houses, and the authors believe that most of the challenges described are representative of those faced by the wider industry. The chapter begins by describing Emerald's understanding of and provision for the policymaker audience in 2019 – when the company began renewed efforts for policy impact – and then goes on to describe its research methods and findings, its identification of the gap between publishers and policymakers, the solutions it considered, and the ones now being implemented. The chapter will also discuss how Emerald intends to measure the success of its attempts to reach policymakers and how it plans to evolve its strategy, working hand-in-hand with the research community and the policymakers it wants to inform.

It should be noted that other academic publishers are also working to ensure that their publications have an impact beyond the academic community. For example, Policy Press, an imprint of Bristol University Press, was founded in 1996 with the specific aim of ensuring that the research it publishes reaches beyond academia, as part of a broader conversation with communities such as policymakers and practitioners. Furthermore, the founder of SAGE Publishing, Sara Miller McCune, has stated that 'SAGE is a company that advocates for policy that will enable science, especially social and behavioral science … to make real-world impact. We believe that social and behavioral sciences can improve, and even save, lives' (McCune, nd).

Past efforts to reach policymakers

In 2019, when Emerald began its efforts to engage more closely with policymakers, the majority of the content it published was aimed at the academic researcher audience. The way in which Emerald's content is accessed contributes to this. The full archive of Emerald's content, and its current publications, can be viewed on its content platform, Emerald Insight (Emerald Publishing, 2021a). The platform is usually accessed through an institutional (usually a university) subscription, and only open access (OA) journal articles and book chapters can be freely read and downloaded.

Policy briefs

A 2019 assessment of its content established that Emerald was publishing a very small amount of content aimed specifically at, and in a format designed for, policymakers. The journal *Disaster Prevention and Management* began publishing policy briefs in 2016, and had published three policy briefs in total: 'Reducing and managing the risk of disaster in Philippine jails and prisons' (Gaillard et al, 2020), 'Cooperation towards disaster risk reduction in the Belt and Road region' (Xu et al, 2017) and 'The effectiveness of hazard risk communication – expert and community perspective on Orewa in Auckland, New Zealand' (Fakhruddin et al, 2020). Published separately to the journal's regular content, the policy briefs were free to access from the journal's homepage (Emerald Publishing, 2021b), rather than being available via Emerald Insight. Usage data for the homepage demonstrated that the policy briefs had been downloaded a total of 5,173 times.[1] This indicated that Emerald's audience included a significant number of people interested in influencing policy.

Blogs

Emerald had previously developed other formats to give its authors reach to audiences beyond the academy. These weren't developed specifically for policymakers but for a wide range of non-academics, or academics from outside an author's specialism. Emerald's Real Impact blog (Emerald Publishing, 2021c) was launched in 2019 to highlight the impact of research in the business and social sciences disciplines, together with routes to achieving impact and ways of assessing impact. Blog posts are typically written by editors and authors who have written for Emerald, with occasional contributions from Emerald staff or third parties. The blog aims be one of the leading voices in the conversation on impact in the academic community. Shaping policy is a key route to impact and a resource like the Real Impact blog would ideally have policymakers among its audience and contributors.

Many aspects of the blog are well-suited to use by policymakers: most posts are based on original research; written by academics, practitioners or policymakers; fully cited but written for a non-specialist audience; often include recommendations for action, and at 500–1,500 words can be read in just a few minutes. Examples of recent content that could support policy, as accessible summaries of actionable research findings, are given in Table 18.1.

Podcasts

Emerald has been creating research-based podcasts since 2017. As with blogs, podcasts were not developed primarily for policymakers but their features make them suitable for this audience. Initially, podcasts were published on an ad hoc basis and tied to a specific book or journal article, but they have evolved over time into the Emerald Podcast Series (Emerald Publishing, 2021d). The series was launched to establish a regular window into research supporting the SDGs. Each podcast focuses on a piece of research, explained by the researchers themselves. It is pitched at a non-specialist level and released in easy-to-digest 20–40 minute episodes. Examples of podcasts that might be accessible to and useful for a policy audience appear in Table 18.2.

Pre-existing relationships with policymakers

Emerald maintains a network of 500,000 published authors, which includes researchers and practitioners from across the globe, working in a variety of subject areas (Emerald Publishing, 2021h). In 2019, research was undertaken with the aim of assessing the connection between Emerald's network and policy. Of the disciplines Emerald covers, it was suggested that education and health were those that would be of most use to a policy audience. Emerald has a strong presence in both disciplines, with a focus on the applied aspects, making commissioning and disseminating trusted policy content feasible. Initial research focused on Emerald's network of editorial stakeholders within these subject areas – journal and book series editors, associate editors and editorial advisory board members. The results were startling – a large number of Emerald stakeholders in these subject areas alone were active in national and local governments across the world, worked with NGOs such as the World Health Organization and the United Nations, or held active links with policy centres and institutions such as think tanks.[2] Evidently, Emerald was already working closely with many researchers who would potentially have an interest in policy-related research, or perhaps even in conducting and writing up research aimed at policymakers.

Table 18.1: Selected blog posts, 2021

Blog title	Author name and expertise	Relevant policy areas/SDGs	Word count
Multifunctional field margins in agriculture landscapes: a window of opportunity for conservation agriculture and biodiversity enhancement (Amir, 2021)	Azhar Amir – a professional writer who has worked for organisations such as the United Nations Development Programme (UNDP), and the Food and Agriculture Organisation (FAO)	Food security; SDG2 – No Hunger; SDG12 – Responsible Consumption and Production; SDG15 – Life on Land	673
Slowly but surely: business schools boost sustainability (Kaplan, 2021)	Andreas Kaplan – Professor of Marketing at the ESCP Business School	Sustainable Business; SDG4 – Quality Education; SDG8 – Decent Work and Economic Growth; SDG12 – Responsible Consumption and Production	654
UN reclassification of cannabis achieves little for patients' rights (Godfrey, 2021)	Sarah Godfrey – patient-researcher and 'illegal UK cannabis patient and home-grower' (Godfrey, 2021)	Drug policy; Health; Criminal Justice; SDG3 – Good Health and Wellbeing; SDG16 – Peace, Justice and Strong Institutions	1,536
Will Australia and the EU achieve global gender equality in a post Covid-19 world? (Rodgers-Healey, 2021)	Diann Rodgers-Healey – Director of the Australian Centre for Leadership for Women (ACLW)	Gender equality; SDG5 – Gender Equality	1,253
Gender inequality in the construction industry: an open letter for change and inclusivity (Varela Alberte, 2021)	Elaine P. Varela Alberte – Universidade Federal da Bahia, Brazil	Gender equality; SDG5 – Gender Equality	787

Table 18.2: Selected podcasts, 2019–21

Podcast title	Contributor names and expertise	Relevant policy areas/SDGs	Play time (mm:ss)
Collapse of the global order on drugs (Emerald Publishing, 2019a)	Axel Klein and Blaine Stothard, co-editors of the journal *Drugs and Alcohol Today*	Drug policy; Health; SDG3 – Good Health and Wellbeing; SDG16 – Peace, Justice and Strong Institutions	30:09
How should we address the issue of sexual violence on campus? (Emerald Publishing, 2019b)	Chris Linder – Associate Professor of College Student Affairs Administration at the University of Georgia, US; Susan Marine – Researcher in access and engagement in higher education, Merrimack College, US; Niah Grimes – Counsellor in higher education, University of Georgia, US; Marvette Lacy – Director, Women's Resource Centre, University of Wisconsin Milwaukee, US	Gender equality; Higher Education; Criminal Justice; SDG4 – Quality Education; SDG5 – Gender Equality; SDG16 – Peace, Justice and Strong Institutions	25:36
Making aid agencies work (Emerald Publishing, 2020)	Terry Gibson – Operations Director, Inventing Futures, UK; Sarwar Bari – Pattan Development Organisation, Pakistan; Manu Gupta – Seeds India, India	International development; SDG1 – No Poverty; SDG8 – Decent Work and Economic Growth	26:50
Health and illness in the neoliberal era in Europe (Emerald Publishing, 2021e)	Jonathan Gabe – Emeritus Professor of Sociology at Royal Holloway, University of London, UK, and former President of the International Sociological Association RC15 Sociology of Health; Mario Cardano – Professor of Sociology of Health at the University of Turin, Italy; Angela Genova – Department of Economics, Society and Politics of the University of Urbino, Italy and past vice-chair of the European Sociological Association Sociology of Health and Illness research network	Health; Austerity; SDG3 – Good Health and Wellbeing	37:48

Table 18.2: Selected podcasts, 2019–21 (continued)

Podcast title	Contributor names and expertise	Relevant policy areas/SDGs	Play time (mm:ss)
Inequalities in mental health care for Black communities (Emerald Publishing, 2021f)	Karen Carberry – a Black British Family and Systems Psychotherapist, and Consultant Family Therapist at Orri, UK; Ted Ransaw – K-12 Outreach Specialist at the College of Education and a faculty member of African and African-American Studies at Michigan State University, US	Health; Racial equality; SDG3 – Good Health and Wellbeing; SDG10 – Reduced Inequalities	45:42
How Covid-19 has exposed inequalities in the UK food system (Emerald Publishing, 2021g)	Bob Doherty – University of York, UK, lead for N8 AgriFood. Seconded as a policy fellow into the UK Government Department for Environment, Food and Rural Affairs (DEFRA) on a new science research programme launched to inform policymaking; Madeleine Power – founder and former chair of the York Food Justice Alliance, a cross-sector partnership addressing food insecurity at the local level, and co-chair of the Independent Food Aid Network (IFAN), UK	Poverty; Food security; SDG1 – No Poverty; SDG2 – No Hunger; SDG3 – Good Health and Wellbeing	40:30

To gain further insight into the ways in which researchers work with policymakers, several of Emerald's editorial stakeholders with connections to policy were interviewed at length. The results highlighted the fact that there is often little collaboration between researchers and policymakers, and that while researchers want their research to reach policymakers, they are often unsure whether this happens in practice. The conversations confirmed that there is a need for greater cooperation between the academic and policymaker communities, and for policymakers to play a key role in identifying the challenges which research should address.[3] Researchers are then ideally placed to address these challenges using the available evidence. By bringing both communities together, the impact of research can be increased.

Understanding the needs of policymakers

The Parliamentary Office of Science and Technology (POST), UK, has been working with Emerald since 2019. POST aims to build connections between researchers and the parliamentary process, through publishing research briefings and connecting experts with policymakers at Westminster (POST, 2021). In October 2019, Emerald sponsored a POST event, which was aimed at encouraging researchers to engage with Parliament.[4] The event was attended by several Emerald authors, and feedback demonstrated that it increased the knowledge and confidence of researchers who had not previously engaged with the policymaking process.

Throughout the process of increasing Emerald's engagement with policymakers, the advice of the team at POST has been invaluable. Through working closely with POST, and meeting policymakers at the POST event, the team at Emerald gained valuable insight into the needs of policymakers. For policymakers, it is essential that research is accessible, as they are unlikely to have access to subscription research platforms, such as Emerald Insight. Research aimed at policymakers, therefore, needs to be published in an OA (open access) format. Discussions with both policymakers and POST confirmed that the references used in any research aimed at policymakers also need to be OA where possible, so that they can be easily accessed by the policymaker.

Emerald's work with POST also demonstrated the need for content aimed at policymakers to be concise. Busy policymakers do not often have the time to read long documents, and need to be able to assess information very quickly. POST advised Emerald that documents aimed at policymakers should be no more than four pages in length. Within these documents it is important that any headline statistics are highlighted, so that policymakers can grasp these quickly in order to be able to easily communicate important figures to a wider group. Finally, policymakers are usually not subject-area experts, so it is essential that research aimed directly at policymakers is structured and written differently to traditional academic research – this includes focusing on results and recommendations rather than methodology, using accessible language and avoiding jargon. At the POST event, academic researchers with experience of writing for policymakers explained that their main challenge in doing so was amending their usual writing style to ensure that their research was accessible to the policymaker audience.

Emerald's own discussions with academic researchers corroborated the findings of its work with POST. The importance of research aimed at policymakers being free to access was heavily emphasised, as was the need for research to be presented in a succinct, easily accessible manner. The results of these conversations, and Emerald's work with POST, were vital in identifying ways in which Emerald could begin to engage more effectively with policymakers. Emerald's research clearly demonstrated the eagerness of academic researchers and policymakers to find ways to work together, and the potential benefits of bringing both communities together.

Shaping content to the needs of policymakers

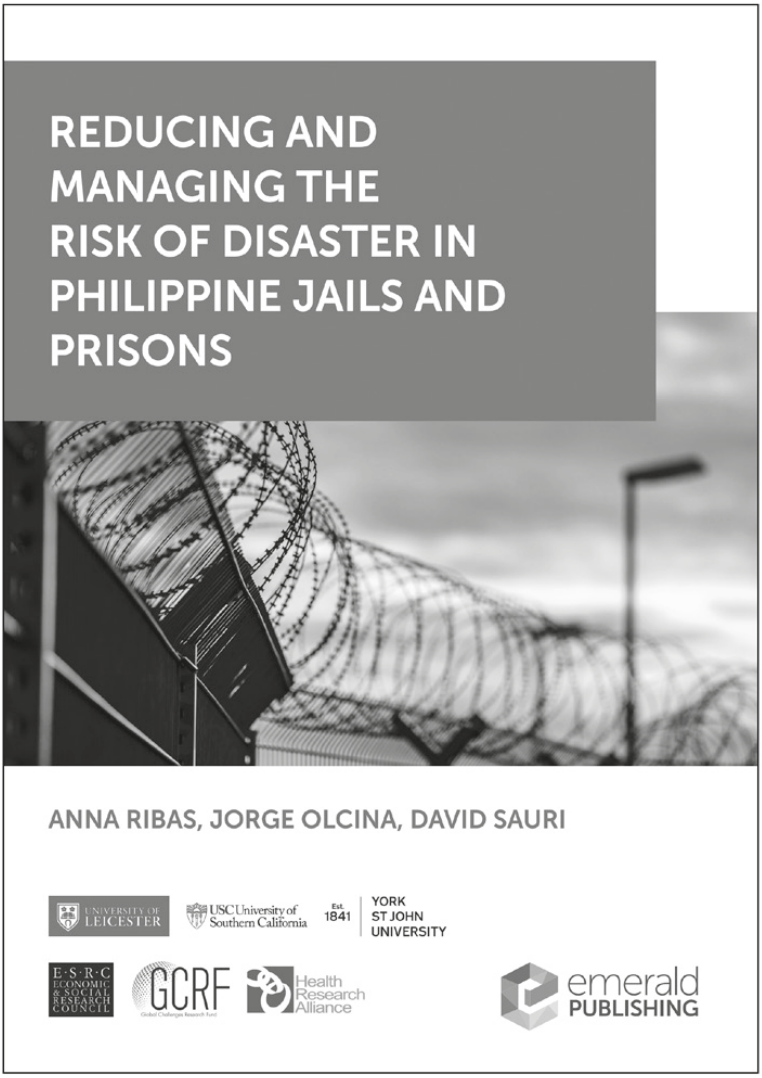
Developing policy briefs

Building on the publication of the policy briefs the journal *Disaster Prevention and Management* had published, Emerald decided to expand this format and begin publishing policy briefs as a new form of content, separate to its journals, which authors are invited to write, to help increase the reach of their research findings. The guiding principles, already discussed, were followed as Emerald worked to develop its policy brief format.

Questions about what information would be contained in policy briefs and how this should be presented were relatively easily answered. While new to Emerald, policy briefs have existed for a long time and there were plenty of examples available to draw on. Based on research of existing policy briefs and conversations with our policy contacts, a suitable structure was identified, along with a set of author guidelines – a starting point from which the documents could evolve, based on audience feedback. In addition to setting out recommended word lengths, advice on structure and helpful information about rights and permissions, these guidelines discuss the style authors should use and the type of information that should and shouldn't be included, to maximise the usability of the brief.

To further support authors and to help build recognition and trust in Emerald as a publisher of policy briefs, the team worked with in-house designers to create a standard template for policy briefs. The design team based the template on the policy briefs produced for the journal *Disaster Prevention and Management*. Text can be easily dropped into the template to create an attractive, readable, standardised document. The cover of the

Figure 18.1: Sample policy brief cover, Emerald Publishing



REDUCING AND MANAGING THE RISK OF DISASTER IN PHILIPPINE JAILS AND PRISONS

ANNA RIBAS, JORGE OLCINA, DAVID SAURI

document features a large image to help identify the broad theme the brief addresses, clearly states the brief title and author names, has spaces for the logos of the authors' institutions and funding organisations, and also features the Emerald Publishing logo (Figure 18.1).

Finding a home for policy briefs

The next step for Emerald was to identify a hosting platform for policy briefs. The question of where policy briefs should be hosted was fundamental as, without the right home online, the content would immediately fail on discoverability and accessibility goals. Emerald had identified that content for policymakers must be OA. The briefs published in connection with the journal *Disaster Prevention and Management* were free to access, as they were hosted as PDFs on the journal homepage (Emerald Publishing, 2021b). This way of hosting content limits searchability, which impacts on discoverability, and its usage is difficult to track, limiting the extent to which its success can be measured. Clearly, a new approach was needed for policy briefs.

In 2019, as part of efforts to reach new audiences beyond academia, Emerald launched Emerald Open Research (EOR) (F1000 Research, nd). EOR is a fully OA publishing platform powered by F1000, which publishes a variety of research outputs, including research articles, case studies, documents and data. All content on the platform is published under a CC-BY licence, meaning that it can be read, shared and reused by anyone, in any format. The platform offers fast publication, making it ideal for sharing practical solutions such as policy briefs, and is structured around the SDGs, freeing policy briefs from the frameworks of journals and traditional subject disciplines. Although EOR does not yet offer the same audience size as the well-established Emerald Insight platform, its flexibility, OA capability and focus on fast publication meant that it was better suited to content for a policy audience.

Getting new content to policymakers

In 2019, Emerald's policy contacts represented just a small section of the policy community. To successfully share policy-oriented content with the right audiences, Emerald needed to reach out beyond its existing networks.

Emerald has a social media presence on Twitter (more than 39,000 followers over multiple subject- and region-specific accounts), LinkedIn (more than 15,000 followers) and Facebook (more than 51,000 followers).[5] Data do not show how many of these followers are policymakers or have strong connections to policy, but Emerald believes most of its followers are academic researchers. Paid posts allow Emerald to reach beyond its followers. Depending on the platform, individuals can be targeted by characteristics such as company, job title, qualifications, interests, geographical location and characteristics shared with the followers of influential accounts in the field. Emerald is grateful to Scott Taylor from the University of Manchester for sharing his approach to using social media to bring research to people who can put it into practice (Taylor, 2021). Building a social media following within the policy community, and bringing policymakers onto its platform to engage with its content, also helps Emerald to build networks for the longer term.

Bridging the gap between research and policy

Building networks

In order for publishers to engage effectively with policymakers, it is essential that policymakers are included at every level of the publishing process. By doing so, publishers like Emerald can become more inclusive, actively listening to the evolving concerns of communities outside of academia, and ensuring that the research they publish contributes towards a wider impact. In addition to publishing policy briefs, Emerald identified other ways of reaching and engaging with the policymaker community. Staff in various publishing roles at Emerald regularly commission content from the academic and practitioner communities, including journal articles and special issues, books and book chapters, podcast episodes, blog pieces and videos. By broadening Emerald's commissioning strategy across all formats to include policymakers, the breadth of Emerald's content, as well as its potential impact, was increased.

As part of a series of workshops aimed at assisting Emerald staff in including policymakers in their commissioning strategies and increasing the representation of policymakers on their editorial teams, a representative from POST visited the company's headquarters and spoke with staff about the ways in which they could further engage with the policymaker community. Professor Peter Murphy of Nottingham Business School at Nottingham Trent University, an Emerald Editorial stakeholder and academic researcher with a wealth of policymaking experience, attended an Emerald workshop in which he spoke with Emerald staff about his experience of working with policy as both a policymaker and a researcher. The feedback from both sessions was positive – staff reported that they found them very useful in beginning conversations with policymakers.

Supporting researchers in reaching policymakers

Policy and impact cannot be separated. Through engaging with policymakers, academics can greatly increase the impact of their research. As well as trying to reach policymakers directly, Emerald has created resources that help researchers to understand the needs of policymakers. In summer 2020, the Real Impact blog featured a series of posts on the topic of closing the gap between research and policy, with practical advice for researchers. Leslie Villegas wrote on 'The do's and don'ts of brief writing for a policymaking audience' (Villegas, 2021), Maria Gintova outlined ways to make research more accessible to those in government (Gintova, 2021) and Scott Taylor explained his strategy for sharing research more effectively using social media, as already mentioned (Taylor, 2021).

The surest and most effective way to ensure research influences policy and is usable by policymakers is to include policy stakeholders at the earliest possible stage. If policymakers have been part of a project from the planning stages, they can help shape the outputs and build the networks needed for dissemination. As part of its aim to support deeper and more profound relationships between researchers and policymakers, Emerald is developing resources as part of its Impact Services initiative (Emerald Publishing, 2019–21), which aims to support academic institutions and their researchers in understanding impact, how it can be achieved through their work and how an institutional culture of impact can be cultivated. This is done through a series of resources, which have been co-created with the community. Among these is an Impact Planner, which encourages researchers to think about the change they want to influence from the outset and the importance of including relevant stakeholders at the earliest stage. There is potential for similar resources to be created around policy impact, which would aim to help researchers understand how best to interact with policymakers throughout their research journey. Emerald's research and work with POST has demonstrated that many researchers are not sure how best to work with policymakers; thus, co-created resources like these would help researchers to develop a way forward, and support greater collaboration towards their impact goals. This would greatly benefit policymakers, who would have access to the latest research in a specific field, written by authors who understand their needs.

Measuring success

As Emerald continues its efforts to engage more closely with policymakers, it is vital that progress is continually assessed to ensure that this is happening as effectively as possible. All of Emerald's content is tracked with regards to usage – the number of downloads or citations an article or document has received. This data is essential for academic researchers, who need to track which academic journals or books have cited their research. It is, however, more difficult to trace how policymakers are using research. Recently, Emerald has employed Altmetric data to provide further detail on how research outputs are being used. Altmetric data provides details of where a research output has been cited outside of journals and books, including in policy documents, in the press, in blogs and on social media. Each document receives an attention score, which indicates the total attention a research output has received (Altmetric, nd). It should be noted that this data does not correspond to the full impact of a research output. However, coupled with data showing where research has been cited by other academic research outputs, it does provide a multifaceted view of how research is being shared and disseminated online and, crucially, within policy. This can also help us to build a picture of which topics might be most relevant to policymakers and to audiences outside of academia. Emerald is working to increase the data it gathers with regards to policy, with the hope that in the future we may be able to trace the full impact of a research output on the policymaking process.

It is important that the work Emerald is doing to increase its engagement with policymakers is continually monitored. EOR is an author-led platform and, in order to evaluate its success, Emerald continuously listens to the feedback received from the communities that use it. This feedback is being used to improve the resources and research outputs Emerald publishes related to policy engagement. For example, feedback from policy brief authors has already helped to refine the publication process, and this will continue as the project develops. Indeed, Emerald is continuing to build partnerships with organisations working across the research, practitioner and policymaker communities, and will continue to work at the nexus of research and policy, broadening its understanding of how these communities can work together in order to drive impact.

Conclusion

There is a gap between researchers and policymakers that is widely recognised but not well understood. The academic and policymaker communities must be brought closer, and that is best achieved by better communication – and ongoing conversation. With a long history of supporting the conversation between researchers, academic publishers are well-placed to facilitate a dialogue between researchers and policy audiences. Those publishers whose focus has historically been on helping to put research into practice may be better positioned than most.

The increasing focus on real world impact in the higher education sector – exemplified by initiatives like the SDGs and DORA – means that reaching policymakers is no longer something that is simply ‘nice to have’ alongside publication in high impact factor journals. It is often part of a researcher’s ethical responsibility to ensure that their research findings make a difference, as this helps offset the risks, burdens and resources taken to conduct the research. This chapter has focused on the work of Emerald, but several publishers are now revisiting and renewing their efforts to reach beyond the academy, developing new forms of content, new hosting platforms and new techniques for disseminating content. Some are also working to help researchers understand what policymakers need from them, to improve the impact of academic work and to make sure policymakers have timely access to high-quality, ethical evidence to shape policy.

Academic publishers have a huge amount still to learn about policy audiences. We believe that, like universities, publishers need to be more open to the world outside the academy. This chapter has illustrated the earnest efforts of one publishing house to reach out to policymakers and begin to bridge the gap. Academic publishers want to help policymakers find the ethical evidence they need to make effective policy, and work with them to make that happen.

Acknowledgements

The authors would like thank Naomi Saint and Sarah Foxen of the Parliamentary Office of Science and Technology (POST); Peter Murphy; Scott Taylor; Ron Iphofen for kindly asking us to write this chapter; and the team at Emerald who contributed their expertise to different parts of this project: George Adams, Ciara Boardman, Kim Chadwick, Virginia Chapman, Danielle Crow, Priya Dharni, Juliet Harrison, Gemma Hemming, Claire Jackson, Emma Leverton, Louise Lister, Abi Masha, Jen McCall, Aidan Morrison, Sharon Parkinson, Tom Saxton, Anna Scaife, Jo Sharrocks, Andrew Tarrant and Sanna Zahoor.

Notes

[1] Figures accessed on 12 June 2019 at 7.40am (UK time), showed that 'Reducing and managing the risk of disaster in Philippine jails and prisons' had been downloaded 4,718 times, 'Cooperation towards disaster risk reduction in the Belt and Road region' had been downloaded 242 times, and 'The effectiveness of hazard risk communication – expert and community perspective on Orewa in Auckland, New Zealand' had been downloaded 213 times.

[2] Network mapping exercise by Emerald Publishing, 29 July 2019.

[3] Interviews with editorial stakeholders of Emerald journals and book series, across subject areas, carried out from July to September 2019.

[4] The event 'Increasing Women Researchers' Impact and Engagement at the UK Parliament', 9 October 2019, Westminster, London, UK.

[5] Aggregated data from Twitter, Facebook and LinkedIn accounts run by Emerald Publishing, accessed 24 March 2021.

References

Altmetric (nd) 'Homepage', Altmetric, available from: https://www.altmetric.com/ [accessed 19 March 2021].

Amir, A. (2021) 'Multifunctional field margins in agriculture landscapes: a window of opportunity for conservation agriculture and biodiversity enhancement', Emerald Publishing, 26 February, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/multifunctional-field-margins-agriculture-landscapes-a-window-opportunity [accessed 24 March 2021].

DORA (Declaration on Research Assessment) (2020) 'San Francisco Declaration on Research Assessment', DORA, https://sfdora.org/read/ [accessed 8 March 2021].

Emerald Publishing (2019a) 'Collapse of the global order on drugs', [podcast], available from: https://soundcloud.com/user-163454702-828217667/collapse-of-the-global-order-on-drugs [accessed 24 March 2021].

Emerald Publishing (2019b) 'How should we address the issue of sexual violence on campus?', [podcast], available from: https://soundcloud.com/user-163454702-828217667/sexual-violence-on-campus [accessed 24 March 2021].

Emerald Publishing (2019–21) 'Impact services', [video] Emerald Publishing, available from: https://impactservices.emerald.com/ [accessed 19 March 2021].

Emerald Publishing (2020) 'Making aid agencies work', [podcast], available from: https://soundcloud.com/user-163454702-828217667/making-aid-agencies-work [accessed 24 March 2021].

Emerald Publishing (2021a) 'Emerald Insight: discover journals, books & case studies', Emerald Publishing, available from: https://www.emerald.com/insight/ [accessed 19 March 2021].

Emerald Publishing (2021b) 'Disaster Prevention and Management', Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/journal/dpm [accessed 19 March 2021].

Emerald Publishing (2021c) 'Opinion & blog', Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog [accessed 24 March 2021].

Emerald Publishing (2021d) 'Emerald podcast series', Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/about/emerald-podcast-series [accessed 24 March 2021].

Emerald Publishing (2021e) 'Health and illness in the neoliberal era in Europe', [podcast] Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/health-and-illness-neoliberal-era-europe [accessed 24 March 2021].

Emerald Publishing (2021f) 'Inequalities in mental health care for black communities', [podcast] Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/inequalities-mental-health-care-black-communities [accessed 24 March 2021].

Emerald Publishing (2021g) 'How Covid-19 has exposed inequalities in the UK food system', [podcast] Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/how-covid-19-has-exposed-inequalities-uk-food-system [accessed 24 March 2021].

Emerald Publishing (2021h) 'We are a global publisher', Emerald Publishing, available from: https://www.emeraldgrouppublishing.com/ [accessed 24 March 2021].

F1000 Research (nd) 'Emerald open research', Emerald Open Research, available from: https://emeraldopenresearch.com/ [accessed 19 March 2021].

Fakhruddin, B., Kuizon, J. and Glover, C. (2020) 'The effectiveness of hazard risk communication – expert and community perspective on Orewa in Auckland, New Zealand', Emerald Open Research, available from: https://emeraldopenresearch.com/documents/2-56 [accessed 11 March 2021; originally published on the *Disaster Prevention and Management* journal homepage, 2019].

Gaillard, J.C., Casing-Baring, E.M., Sacayan, D., Balay-as, M. and Santos, M. (2020) 'Reducing and managing the risk of disaster in Philippine jails and prisons', Emerald Open Research, available from: https://emeraldopenresearch.com/documents/2-54 [accessed 11 March 2021; originally published on the *Disaster Prevention and Management* journal homepage, 2016].

Gintova, M. (2021) 'Making policy research more accessible for government practitioners', *Emerald Publishing* [blog], 12 January, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/making-policy-research-more-accessible-government-practitioners [accessed 9 March 2021].

Godfrey, S. (2021) 'UN reclassification of cannabis achieves little for patients' rights', *Emerald Publishing* [blog], 12 February, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/un-reclassification-cannabis-achieves-little-patients-rights [accessed 24 March 2021].

Kaplan, A. (2021) 'Slowly but surely: business schools boost sustainability', *Emerald Publishing* [blog], 17 November, https://www.emeraldgrouppublishing.com/opinion-and-blog/slowly-surely-business-schools-boost-sustainability [accessed 24 March 2021].

McCune, S.M. (nd) 'Social & behavioral science impact', SAGE Publishing, available from: https://group.sagepub.com/social-behavioral-science-impact [accessed 3 December 2021].

POST (Parliamentary Office of Science and Technology) (2021) 'Bridging research and policy', UK Parliament, available from: https://post.parliament.uk/ [accessed 19 March 2021].

Research England (2015) 'REF impact', UK Research and Innovation, available from: https://re.ukri.org/research/ref-impact/ [accessed 8 March 2021].

Rodgers-Healey, D. (2021) 'Will Australia and the EU achieve global gender equality in a post COVID-19 world?', *Emerald Publishing* [blog], 4 March, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/will-australia-and-eu-achieve-global-gender-equality-a-post-covid-19-world [accessed 24 March 2021].

SDSN Northern Europe (2020) 'SDGs in universities', available from: https://www.unsdsn-ne.org/our-actions/sdgs-in-universities/ [accessed 8 March 2021].

Taylor, S. (2021) 'Mobilising knowledge through social media: getting the right ideas to the right people', *Emerald Publishing* [blog], 29 April, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/mobilising-knowledge-through-social-media-getting-right-ideas-right-people [accessed 9 March 2021].

Varela Alberte, E.P. (2021) 'Gender inequality in the construction industry: an open letter for change & inclusivity', *Emerald Publishing* [blog], 12 February, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/gender-inequality-construction-industry-open-letter-change-inclusivity [accessed 24 March 2021].

Villegas, L. (2021) 'The do's and don'ts of brief writing for a policy-making audience', *Emerald Publishing* [blog], 12 January, available from: https://www.emeraldgrouppublishing.com/opinion-and-blog/dos-and-donts-brief-writing-a-policy-making-audience [accessed 9 March 2021].

Xu, X., Gaillard, J.C., Yang, S. and Velasquez, J. (2017) 'Cooperation towards disaster risk reduction in the Belt and Road region', UNDRR [United Nations Office for Disaster Risk Reduction], available from: https://www.preventionweb.net/publications/view/56056 [accessed 24 March 2021; originally published on the *Disaster Prevention and Management* journal homepage, 2017].

Index

References to figures appear in *italic* type;
those in **bold** refer to tables. References to endnotes
show both the page number and the note number (198n6).

"Wide-ranging and practical. This is a wonderful primer for the new researcher and an essential update for senior leaders in policymaking and research."
Hugo Slim, University of Oxford

"This is a superb book that should be in every policymaker's library as it contains best evidence to guide them in their decision making."
Bernadette Melnyk, The Ohio State University

EPDF and EPUB available Open Access under CC-BY-NC-ND licence.

This important book offers practical advice for using evidence and research in policymaking. The book has two aims. First, it builds a case for ethics and global values in research and knowledge exchange, and second, it examines specific policy areas and how evidence can guide practice.

The book covers important policy areas including the GM debate, the environment, Black Lives Matter and COVID-19. Each chapter assesses the ethical challenges, the status of evidence in explaining or describing the issue and possible solutions to the problem. The book will enable policymakers and their advisors to seek evidence for their decisions from research that has been conducted ethically and with integrity.

Ron Iphofen is an independent consultant based in France with international recognition for expertise on research ethics and professional standards in research.

Dónal O'Mathúna is Associate Professor at The Ohio State University, USA, and internationally recognised for his work on disaster bioethics, including the ethics of research in disasters and humanitarian settings.

Policy Press
PUBLISHING WITH A PURPOSE
@policypress PolicyPress
policy.bristoluniversitypress.co.uk

ISBN 978-1-4473-6395-8
9 781447 363958