

Refining reporting guidelines using behaviour change theory

Author:

James Harwood

Supervisors:

Prof. Gary Collins

Dr. Charlotte Albury

Dr. Jennifer de Beyer

Dr. Michael Schlüssel



ST. CROSS COLLEGE
UNIVERSITY OF OXFORD

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy
Michaelmas Term 2024*

Abstract

Refining reporting guidelines using behaviour change theory

The lasting legacy of most medical research is the written account. When writing up their work, researchers often omit information that readers — including clinicians, reviewers, patients, and other researchers — need to fully understand, appraise, replicate, or apply the research. Reporting guidelines try to solve this problem. They are community-created recommendations of information to include when writing up research so that everybody can use it. The first reporting guideline was created almost 30 years ago, and many more have been created since. Most leading medical journals ask authors to adhere to reporting guidelines, yet adherence remains low in almost all medical fields. When authors do not adhere to reporting guidelines their work is less transparent, less valuable, contributes less to patient outcomes and has potential to inadvertently harm by distorting the evidence base. The aim of my thesis is to understand why authors do not adhere to reporting guidelines when writing up medical research, and to address these reasons with the intention of increasing adherence in the future.

I began my thesis with this clear aim but without a route to get there. My first step was a qualitative synthesis (chapter 3) of research exploring authors experiences of reporting guidelines, and I discovered authors' adherence to reporting guidelines is influenced by features of the guidelines themselves, but also the system that surrounds them including the websites that disseminate them and the people and policies of the broader scholarly ecosystem. I probed this system further with a review of quantitative surveys and an evaluation of a key dissemination website (chapters 4 & 5).

Given the complexity of the system and the complexity of the behaviour in question, I realised I would need a framework to help me make sense of them. I decided to use the Behaviour Change Wheel (chapter 6), a well established and evidenced framework for designing and defining complex behaviour change interventions. I led workshops and

focus groups (chapters 7 & 8) to guide stakeholders through frameworks' steps and, ultimately, ended up identifying 32 influences that needed to change, and 128 ideas of how to change them. These ideas were applicable to guideline developers, the EQUATOR Network, journals, publishers, funders, ethics committees, and institutions.

I decided that redesigning reporting guidelines and EQUATOR's website would be an effective, affordable, acceptable, and equitable way to realise many of the generated ideas. I defined 63 intervention components and managed to incorporate 46 into a prototype website that included a redesigned reporting guideline (the Standards for Reporting Qualitative Research) and the EQUATOR Network's home page (chapter 9). I then explored authors experiences of using this prototype through in-depth interviews, observation, think aloud, and writing appraisals (chapter 10) and I identified 53 possible areas for improvement.

My key outputs include a design blueprint for user-friendly reporting guideline resources, and an online platform for creating and hosting resources following this blueprint. This platform will benefit guideline developers who typically lack the money, time, and skills to create user friendly resources. I have secured funding for a feasibility study and parallel process evaluation to explore the acceptability and usability of this platform amongst authors submitting to journals, in preparation for a future evaluation to determine whether my redesign leads to more authors adhering to reporting guidelines successfully.

Acknowledgements

Many people influenced my capability, opportunity, and motivation throughout my DPhil, and without them my journey would have been a lot harder.

I am extremely grateful to Professor Gary Collins, Dr. Charlotte Albury, Dr. Jennifer de Beyer, and Dr. Michael Schlüssel for their expert supervision and for giving me the tools I needed to explore new paths. I would like to thank my colleagues Jacqueline, Patricia, Paula, and Shona for making me feel part of the team, and I'd like to thank Caroline for welcoming me to the world of reporting guidelines and always being ready to put the world to rights. I thank my family and partner for encouraging me throughout this endeavour, and for giving me the support and space to complete it.

Finally, I thank Cancer Research UK for funding my research.

Table of contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 The problem of poor reporting in health research	1
1.2 Reporting guidelines to the rescue?	4
1.3 The EQUATOR Network unites the reporting guideline movement	10
1.4 Reporting guidelines are part of a complex behaviour intervention	10
1.5 Reporting guidelines have had little to no effect on reporting quality, and few authors adhere to them	15
1.6 Evidence gap: What more could be done to improve guideline adherence?	20
1.7 Summary	22
1.8 Aims and Objectives	23
1.9 Thesis structure	24
2 Reflections on starting my DPhil	29
I Identifying influences	35
3 What influences researchers when using reporting guidelines? Part 1: A the- matic synthesis	37
3.1 Introduction	37
3.2 Methods	39
3.3 Results	45

3.4 Discussion	77
3.5 Reflections on this chapter	80
4 What influences researchers when using reporting guidelines? Part 2: Describing the questions asked in quantitative surveys	83
4.1 Introduction	83
4.2 Methods	84
4.3 Results	95
4.4 Discussion	104
4.5 Reflections on reporting this chapter	107
5 Describing who visits EQUATOR's website for disseminating reporting guidelines, and how visitors use it	109
5.1 Introduction	109
5.2 Methods	111
5.3 Results	113
5.4 Discussion	121
5.5 Reflections on this chapter	126
II Identifying and implementing intervention changes	127
6 Selecting the Behaviour Change Wheel framework	129
6.1 Reflections on this chapter	136
7 Following the behaviour change wheel guide: Workshops with reporting guideline experts	139
7.1 Introduction	139
7.2 Methods	140
7.3 Results	143
7.4 Discussion	153
7.5 Reflections on this chapter	155
8 Generating ideas to address influences: focus groups with reporting guideline developers, advocates, and publishers	159

8.1	Introduction	159
8.2	Methods	160
8.3	Results	170
8.4	Discussion	194
8.5	Reflections on this chapter	197
9	Defining intervention components and developing a prototype	199
9.1	Introduction	199
9.2	Methods	201
9.3	Results	207
9.4	Discussion	234
9.5	Reflections on this chapter	237
III	Refining the intervention	239
10	Refining the intervention: interviewing authors to identify deficient intervention components	241
10.1	Introduction	241
10.2	Methods	242
10.3	Results	258
10.4	Discussion	285
10.5	Reflections on this chapter	292
11	Discussion	295
11.1	Chapter Overview	295
11.2	Summary of findings	295
11.3	Overview of outputs	296
11.4	Contributions and transferability to the meta-research community	298
11.5	Strengths	300
11.6	Limitations	302
11.7	Recommendations for future research	307
11.8	Implications for policy	308
11.9	Conclusions	308

References	311
Appendices	341
A PRISMA-S Checklist for chapter 3	341
B ENTREQ Checklist for chapter 3	345
C Search strategies	353
D SRQR Checklist for chapter 7	357
E BCW Step 1 – Define the problem in behavioural terms	365
F BCW Step 2 – Select the Target Behaviour	367
G BCW Step 3 – Specify the Target Behaviour	371
H BCW Step 4 – Identifying what needs to change	375
I BCW Step 4 – Consolidated list of influences needing to change for the target behaviour to occur	381
I.1 Researchers may not know what reporting guidelines are	381
I.2 Researchers may not know what reporting guidelines exist	382
I.3 Researchers may not know whether a reporting guideline applies to them	382
I.4 Researchers may not know what reporting guideline is their best fit . . .	383
I.5 Researchers may not know what resources exist for a reporting guideline	383
I.6 Researchers may not know when reporting guidelines should be used . .	384
I.7 Researchers may misunderstand	384
I.8 Researchers may not know what benefits to expect	385
I.9 Researchers may not know why items are important	385
I.10 Researchers may not know how to do an item	386
I.11 Researchers may not know how to report an item in practice	386
I.12 Researchers may not know what to write when they cannot report an item	386
I.13 Researchers have limited time	387

I.14 Researchers may not encounter reporting guidelines early enough to act on them	387
I.15 Researchers may not understand the language	388
I.16 Researchers may struggle to keep writing concise	388
I.17 Researchers may not have tools for the job at hand	389
I.18 Reporting guidelines can become outdated	389
I.19 Researchers may struggle to reconcile multiple sets of guidance	390
I.20 Researchers may be asked to remove reporting guideline content	390
I.21 Researchers may forget to use reporting guidelines at earlier research stages	390
I.22 Guidance may be difficult to find	391
I.23 Reporting guidelines may be difficult to access	391
I.24 Reporting guideline resources may not be in usable formats	391
I.25 Researchers may feel afraid to report transparently	392
I.26 Researchers may feel restricted if reporting guidelines prescribe design .	392
I.27 Researchers may feel patronized	393
I.28 Researchers may not believe stated benefits	393
I.29 Researchers may not care about the benefits of using a reporting guideline	393
I.30 Researchers may expect the costs to outweigh benefits	394
I.31 Researchers may feel that checking reporting is someone else's job.	394
I.32 Researchers may not consider writing as reporting	395
J BCW Step 5 – Identifying intervention functions	397
K BCW Step 6 – Identifying policy categories	409
L SRQR Checklist for chapter 8	419
M Focus Groups - emails	427
N Focus Groups - participant information	431
O Ideas generated from workshops and focus groups	437
Before developing a reporting guideline	437
O.1 Create reporting guidance for early stages of research	437
O.2 Avoid confusing authors with too many reporting guidelines	438

When developing a reporting guideline	439
O.3 Avoid prescribing structure	439
O.4 Keep reporting guidelines agnostic to design choices	439
O.5 Describe reporting items fully	440
O.6 Describe each reporting guideline fully	441
O.7 Keep guidance short	442
When writing guidance down and creating resources	443
O.8 Make resources ready-to-use	443
O.9 Make reporting guidelines easy to understand	443
O.10 Use persuasive language and design	443
O.11 Create additional tools	444
O.12 Make resources easy to discover and find	445
O.13 Make information digestible	446
When disseminating resources	447
O.14 Describe reporting guidelines where they are encountered	447
O.15 Make resources accessible	448
O.16 Show and encourage citations	448
O.17 Provide testimonials	449
On an ongoing basis	449
O.18 Budget for reporting	449
O.19 Create rewards	450
O.20 Create discussion spaces	450
O.21 Create ways to catch authors earlier	451
O.22 Endorse and enforce reporting guidelines	452
O.23 Evidence the benefits	452
O.24 Make reporting guidelines appear as a priority	453
O.25 Promote reporting guidelines	453
O.26 Install reporting champions	454
O.27 Provide additional teaching	454
O.28 Make updating guidelines easier	455

Q SRQR Checklist for chapter 10	501
R Author interview recruitment image	509
S Participant information sheet for author interviews	511
T Emails for author interviews	515
U Author interview schedule	519
V Deficiencies	527
V.1 Include design, features, and language to foster trust	527
V.2 Describe what reporting guidelines are where they are first encountered	527
V.3 Use tone of voice and design to communicate personal benefits; confidence and simplicity	528
V.4 Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications) . . .	528
V.5 Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for	528
V.6 For each item, provide examples of reporting in different contexts . . .	529
V.7 Define key terms	529
V.8 Cater to different kinds of user (readers vs dippers) by structuring guidance with headings, itemisation, hyperlinking to particular sections, and with optional content	531
V.9 Explain how the guidance was developed and why it can be trusted . .	533
V.10 Make guidance appear shorter by removing superfluous information, hiding optional content, splitting long guidelines, using concise language, and separating design advice	535
V.11 Create spaces for authors to discuss reporting guidelines with others .	536
V.12 Address communications to authors	537
V.13 Structure guideline items to make them quicker to digest	539
V.14 Describe the scope of a reporting guideline at the top of every resource .	539
V.15 Search function on website	540

V.16	Include testimonials from research users who benefit from complete reporting	540
V.17	Include testimonials from researchers who were nervous about being punished for reporting transparently	541
V.18	Instruct authors to cite reporting guidelines so readers may learn about them	542
V.19	Explain when reporting guidelines do not intended to prescribe structure	543
V.20	Provide links to other resources that explain how an item can be done . .	543
V.21	Links between related guidelines	544
V.22	Educate authors about writing as a process	545
V.23	For each item, explain why the <i>information</i> is important and to whom (not just what constitutes “good” design)	546
V.24	For each item, provide clear instruction of what needs to be described .	547
V.25	Gather and communicate evidence for benefits	548
V.26	Provide advice regarding how to respond if asked to remove reporting guideline content by a colleague, editor, or reviewer	549
V.27	Use plain language	550
V.28	Link all resources to each other	550
V.29	Reassure that all research has limitations to encourage explanation over perfect design	550
V.30	Tell authors when to use reporting guidelines, or that reporting guidelines are best used as early as possible	551
V.31	Use consistent terms	552
V.32	Avoid patronizing language	552
V.33	Centralised hosting	553
V.34	Provide translations	553
V.35	Tell authors how long the guidance will take to read	553
V.36	Encourage explanation even when choices are unusual or not optimal .	554
V.37	Explicitly state when no better guidance exists for a use case	555
V.38	Provide instruction as to how and where information can be reported without breaching word count limits or making articles bloated.	555
V.39	Reassure when guidelines are <i>just</i> guidelines	555

V.40 Ensure all resources and tools (e.g., checklists and templates) are in ready-to-use formats	556
V.41 Decrease fear of judgement by making reporting guidelines design agnostic	556
V.42 Explain importance of complete reporting to the scientific community .	557
V.43 Provide clear instruction of what needs to be described when an item was not done, could not be done, or does not apply	557
V.44 Remove branding and messaging that may invoke feelings of judgement, complexity, or administrative red-tape	558
V.45 Search Engine Optimization	558
V.46 Use if-then rules to direct authors to more appropriate and up-to-date guidance when available	558

List of Figures

1.1	Graphical abstract describing the outputs of research chapters and the relationships between them	27
3.1	PRISMA flow diagram	47
6.1	The COM-B system for understanding behaviour, reproduced from Michie et al. [150] used under Creative Commons Attribution License . .	131
6.2	The Behaviour Change Wheel for designing behaviour change interventions. The innermost ring is the COM-B model for understanding behaviour, the middle ring are intervention functions, and the outer ring are policy categories. Reproduced from Michie et al. [150] used under Creative Commons Attribution License	132
8.2	An example row in the Ideas Table that was co-edited by participants. The left hand column contained influences that needed to be addressed for our target behaviour to occur. Existing ideas to address each influence were initially hidden (by turning the text white), and only made visible once participants had discussed their own ideas.	164
9.1	A simplified layout of the existing EQUATOR Network website, as of the 5th of April, 2023. Users must navigate through up to 5 different web pages to reach reporting guidance. The proportion of users navigating between each step is shown in the width of the links. Links in grey are estimated proportions. “Exit” means leaving the website. For simplicity, I have not included the 7% of authors viewing pages within the “Library” subdirectory, nor the 1% visiting other content categories.	208

9.2 The layout of the new web-intervention. Users must now only need to navigate to 2 pages to access reporting guidance. The proportion of users navigating between each step is shown in the width of the links. All links in are estimated proportions, based on a realistic aim to reduce the exit rate from the home page and database page by 50%.	209
9.3 The existing EQUATOR home page, as captured on 5th of April, 2023. Labelled areas are limitations and area described in the previous paragraph.	216
9.4 Redesigned intervention home page.	217
9.5 The existing EQUATOR page for SRQR, as captured on 5th April, 2023. Labelled limitations are described in the previous paragraph.	218
9.6 The SRQR publication, captured on the 5th April, 2023. Areas A-D show the scope, checklist, link to the guidance, and its strictness.	220
9.7 Redesigned intervention reporting guideline introduction and FAQ. Labelled areas include intervention components and are described in the preceding paragraph.	222
9.8 An example item from the SRQR guideline.	225
9.9 A redesigned item from the SRQR reporting guideline. Labelled intervention changes are described in the paragraph above.	230
9.10 Intervention discussion page. Every reporting item now has its own discussion page where authors can ask and answer questions (area A), and provide feedback to guideline developers.	231
10.1 The home page before and after initial revisions based on feedback from the first five participants.	264

List of Tables

1.1	A selection of highly cited reporting guidelines and their similar dissemination strategies	6
1.2	Example journal policies regarding reporting guidelines.	11
1.3	Sources of complexity within reporting guidelines and the system drives their use.	13
1.4	Studies published between August 2020 and June 2022 assessing adherence to CONSORT, PRISMA, STROBE, or STARD and their extensions. Data from Dal Santos et al. [15]	19
3.1	Information sources and record management	40
3.2	Reporting guidelines featured on the EQUATOR Network's home page . .	44
3.3	Study characteristics	49
3.4	Codes (left, not bold), descriptive themes (bold), and analytic themes (right)	57
4.1	Studies that collected quantitative data to explore researcher's experiences of reporting guidelines	86
4.2	Codes describing the focus of questions asked and their code categories. Items in bold did not appear in the qualitative data.	96
4.3	Codes and descriptive themes identified from a qualitative evidence synthesis. Items in bold did not appear in the quantitative questions. Items in italic offer possible explanations to some quantitative findings.	98
5.1	Definitions of terms used in this chapter	111

5.2 User demographics equator-network.org for the year 2021. *There are 105 English locales, including the United States, India, Ireland etc.	114
5.3 Session information for EQUATOR-Network.org for the year 2021. A session is defined as group of user interactions that take place within a given time frame. If a user visited the site twice within a year, then that would appear as two sessions. A session ends after 30 minutes of inactivity.	116
5.4 Web pages where visitors began their session and their bounce rates between 1st of January and 1st of July 2022. Bounces are sessions where a visitor leaves without interacting with the website at all.	117
5.5 How frequently reporting guideline pages were viewed and resources accessed between the 1 st of January 2022 – 1 st of July 2022. Outbound links were mainly links to publications and files were mainly checklists, but a few guidelines also have flow diagrams which were rarely downloaded. The ten most accessed guidelines are shown.	118
5.6 Survey responses to the question “Why are you leaving?”	120
6.1 COM-B model components and examples, reproduced from Michie et al. [150] used under Creative Commons Attribution License	133
6.2 Intervention function definitions and examples, reproduced from Michie et al. [150] used under Creative Commons Attribution License	134
6.3 Policy categories, reproduced from Michie et al. [150] used under Creative Commons Attribution License	136
7.1 Techniques for establishing trustworthiness. Based on Lincoln and Guba’s Evaluative Criteria [156]	141
7.2 Favoured techniques from the Behaviour Change Technique taxonomy v1.0 [158]	151
8.1 Techniques for establishing trustworthiness. Based on Lincoln and Guba’s Evaluative Criteria [156]	167
8.2 Broad ideas and the influences they target, categorised by when they should be considered. Influences were derived and consolidated in previous chapters, and are reported fully in Appendix I)	171

9.1 Intervention components pertaining to how the website is organised, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.	210
9.2 Intervention components appearing throughout the website, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.	211
9.3 Intervention components appearing on the website's home page, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.	213
9.4 Intervention components appearing in the introductory material preceding the reporting guideline, alongside the intervention functions and behaviour change techniques (BCT) they employ. Italicized items are duplicated from previous tables. See Appendix P for further descriptions and links to barriers each component addresses.	223
9.5 Intervention components pertaining to the content and organisation of reporting guidelines, alongside the intervention functions and behaviour change techniques (BCT) they employ. A reporting guideline's content includes (but is not limited to) reporting <i>items</i> . Components specific to reporting items appear in Table 9.6. Italicized items are duplicated from previous tables. See Appendix P for further descriptions and links to barriers each component addresses.	226
9.6 Intervention components appearing in reporting guideline items, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.	228

9.7 Intervention components pertaining to the website's FAQ section, alongside the intervention functions and behaviour change techniques (BCT) they employ. <i>Italicized</i> items are duplicated from previous tables. See Appendix P for further descriptions and links to barriers each component addresses.	232
10.1 Data collection methods and when they occurred.	245
10.2 Methods used and the intervention components they explore. Intervention components are defined in chapter 9	251
10.3 Techniques used in this study for establishing trustworthiness. Based on Lincoln and Guba's Evaluative Criteria [156]	256
10.4 Recruitment and drop out of participants through different channels . .	259
10.5 Participant characteristics	260
A.1 PRISMA-S Checklist for chapter 3 [86]	341
B.1 ENTREQ Checklist [87]	345
D.1 SRQR Checklist [30]	357
E.1 Behaviours scored for impact, likelihood, spill-over, and measurement .	370
H.1 Workshop participants' ideas around what needs to change for authors to perform our target behaviour. Based on the COM-B questionnaire [154]	375
J.1 Workshop participants' considerations on the affordability, practicability, effectiveness and cost-effectiveness, acceptability, side-effects/safety, and equitability of intervention functions. Intervention functions and their definitions come from [154].	398
K.1 Workshop participants' considerations on the affordability, practicability, effectiveness and cost-effectiveness, acceptability, side-effects/safety, and equitability of policy categories. Policy categories and their definitions come from [154].	411
L.1 SRQR Checklist [30]	419

P.1 Intervention Components Table. Intervention components labelled with the behaviour change techniques and intervention functions they employ, and grouped according to the key behaviours, barriers, and behavioural drivers that they aim to target. Where possible, examples demonstrate how components were (or were not) used originally (Before) and, how they are included within the redesigned intervention (Now) . . .	458
Q.1 SRQR Checklist [30]	501

List of Abbreviations

Reporting Guidelines

AGREE	Appraisal of Guidelines, REsearch and Evaluation
ARRIVE	Animal Research: Reporting of In Vivo Experiments
CARE	guidelines for CAsE REports
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
COGS	Conference On Guideline Standardization
CONSORT	CONsolidated Standards Of Reporting Trials
COREQ	COnsolidated criteria for REporting Qualitative research
ENTREQ	ENhancing Transparency in REporting the Synthesis of Qualitative Research statement
GREET	Guideline for Reporting Evidence-based practice Educational interventions and Teaching
GRRAS	Guidelines for Reporting Reliability and Agreement Studies
MDAR	Materials Design Analysis Reporting
ORION	Outbreak Reports and Intervention Studies Of Nonsocomial infection
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA Equity	Preferred Reporting Items for Systematic Reviews and Meta-Analyses with a Focus on Health Equity
PRISMA-P	Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols

PRISMA-S	Preferred Reporting Items for Systematic Reviews and Meta-Analyses Searches
RIGHT	Reporting Items for practice Guidelines in HealThcare
SAMPL	Statistical Snalyses and Methods in the Published Literature
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
SQUIRE	Standards for QUality Improvement Reporting Excellence
SRQR	Standards for Reporting Qualitative Research
STARD	STAndards for Reporting Diagnostic accuracy studies
STREGA	STrengthening the REporting of Genetic Association Studies
STRICTA	STandardsäfor Reporting Interventions in Clinical Trials of Acupuncture
STROBE	STrengthening the Reporting of OBservational Studies in Epidemiology
TREND	Transparent Reporting of Evaluations with Nonrandomized Designs
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
Other Abbreviations	
NICE	National Institute for Health and Care Excellence
CASP	Critical Appraisal Skills Programme
EQUATOR	Enhancing the QUAlity and Transparency Of health Research
GIM	Global Index Medicus
OSF	Open Science Framework
SciELO	Scientific Electronic Library Online

Chapter 1

Introduction

In this chapter I introduce the evidence gap I have addressed and the approach I have taken to address it. I begin by describing the prevalence and consequences of poorly reported medical research. I introduce reporting guidelines and position them as part of a complex behaviour change intervention. This intervention has had a disappointing impact—few authors adhere to reporting guidelines and medical research is still poorly reported. This brings me to my evidence gap: how can we get more authors to adhere to reporting guidelines? I then outline my aims, objectives, and thesis structure.

1.1 The problem of poor reporting in health research

Facing an uncertain choice during treatment for multiple myeloma, epidemiologist Alessandro Liberati wrote “Why was I forced to make my decision knowing that information was somewhere but not available?” [1]. When I started my DPhil governments may have been asking the same question. The world was in the grip of COVID-19 and decision makers were wading through a deluge of patchy research articles missing important information [2]. In the years since, members of my family have had to make treatment decisions where the evidence is of “low certainty” because key details are missing from research articles.

A selfish silver-lining of these tumultuous years was that my family and friends finally understood the problem my thesis addresses: when medical researchers inadequately

describe what they did, why they did it, and what they found, other people cannot understand, replicate, or use their work. Research costs huge amounts of time, money, and effort, and the written account is typically its sole legacy. When details are omitted they are lost. The remaining gaps are sources of doubt; are they accidental omissions? Oversight? Cover-ups? Whatever their source, the gaps fragment the full picture, and the potential value to patients drains away.

Early concern over reporting quality came from frustrated reviewers unable to find the data they needed within research reports. For example, in 1963, Glick [3] found many reports of psychiatric therapies used ambiguous descriptions of treatment duration like “at least two months” or “from one to several months”. These descriptions were so vague they were “unsuitable for comparative purposes”. In the decades since, evidence syntheses have increasingly underpinned clinical guidance, and the Cochrane Collaboration and the National Institute for Health and Care Excellence (NICE) are two organisations conducting such syntheses. Their reviews influence clinical practice throughout the UK and globally, but are undermined when the underlying literature omits details. Dechartres et al. [4] looked at all Cochrane reviews published between 2011 and 2014. These reviews included over 20,000 randomised controlled trials. A third of these trials were so poorly reported, reviewers could not judge their potential for bias thereby limiting the confidence of conclusions. Some researchers fear that *unclear* bias may actually be *high-but-hidden* bias. For example, Savovi et al. [5] found exaggerated intervention effect estimates in randomized trials with high or unclear (versus low) risk of bias judgments. Poor reporting, therefore, not only undermines the confidence of systematic reviews but potentially allows distortions to creep into the evidence un-checked.

No similar data exists for NICE reviews, so instead I’ll draw on personal experience of consulting NICE evidence when a relative was deciding between two cancer procedures. Both options had serious side effects, similar success rates, and neither had a clear-cut advantage given my relative’s medical history. To respect my relative’s privacy I won’t cite the NICE review, but it included ten studies: seven case series, two non-randomised comparison studies, and one randomised trial. Two studies did not describe participants’ gender, one did not describe ages, four did not describe their exclusion criteria, three did not describe side-effects fully. Consequently, NICE’s recommendation was

less certain than it could have been. Did the case series include patients similar to my relative? We couldn't tell. Did participants of equivalent age experience the side effects we feared? No idea. My family echoed Liberati's frustration: the missing information was probably collected at some point and we (as tax payers) may even have paid for some of that research, but because details were not written down they would not help us in our hour of need. In typical scientific objectivity, Dechartres wrote [4] "waste related to poor reporting could be completely avoided", but it's hard to remain so dispassionate when tossing and turning in the small hours of the morning, deliberating a life-changing decision, when details matter more than ever.

Two of the studies in that NICE review barely described the procedure in question, so had the surgeon wished to repeat it for my relative he wouldn't have been able to. This reveals another consequence of poor reporting: when research is poorly described, it cannot be repeated. Doctors and service providers need clear descriptions to replicate interventions [6]. As Feinstein noted in 1974 [7], it is difficult enough for a clinician to understand the value of unfamiliar procedure, but "it is much more difficult when he is not told what that procedure was". For example, Davidson et al. [8] reviewed trials describing exercise interventions for chronic back pain and found authors often did not describe interventions sufficiently for other healthcare providers to copy them. Researchers also need clear descriptions to understand, appraise, and repeat each others' work. For example, Carp [9] described how a third of 241 brain imaging studies missed information necessary to interpret and repeat them, like the number of examinations, examination duration, and the resolution of images.

As well as being wasteful, poor reporting breeds mistrust. The COVID pandemic saw a deluge of research, yet of 251 trial reports [10], only 14% described harms adequately; fifty five trials (22%) did not report any information about harms, and most others (n=150, 60%) did not explain whether harms resulted in participants dropping out. Fear over possible side effects stopped many people from getting vaccines [11], seeking medical treatment [12], or wearing masks [13]. By not reporting side effects, trial reports did nothing to quell those fears and, in some cases, may have stoked them further. Even years later, in April 2024, a British member of parliament describing himself as "double

vaccinated and vaccine-harmed” [14] told the House of Commons that the “adverse effects of the [Covid-19] vaccine... are not currently known”, before speaking of “gaslighting” and blaming the vaccine as causing “enormous harm”. He repeatedly railed against an omission of information about harms — “crucial information was hidden from the regulator and the public”, “The public are being denied that data... yet again, data is hidden with impunity” — before jumping to un-evidenced, exaggerated, or conspiratorial claims of harm and cover-ups. The absence of evidence seemingly left a void for fantasy and fear to fill.

These are a mere handful of many studies documenting poor reporting in medical literature. A 2023 systematic review found 148 published between 2020-2022 alone [15]. All investigated reporting quality in different medical research disciplines, and *almost* all concluded reporting was sub-optimal. Hence, poor reporting is a long-standing problem, plagues all disciplines, devalues research, seeds mistrust in science, introduces avoidable doubt and patient distress, and derails the uptake of new knowledge into clinical practice.

1.2 Reporting guidelines to the rescue?

Concern over reporting quality crescendoed through the eighties and early nineties as systematic reviews became more common. Responding to calls for “strategies”, “guides”, and “lists” to help authors prepare their manuscripts, the Standards Of Reporting Trials (SORT) [16] and Asilomar [17] guidelines were published in 1994 by two independent groups of methodologists, trialists, and editors. Both advised on how to report randomised trials and had similar content. On the suggestion of JAMA’s deputy editor, the SORT and Asilomar groups met in 1996 and combined their recommendations to create the CONSolidated Standards of Reporting Trials (CONSORT) statement [18]. CONSORT is a set of recommendations detailing what information authors should include in clinical trial reports. It comprised an article describing how it was made, a checklist, flow diagram, and (after an update in 2001) and ‘Explanation and Elaboration’ publication [19–21].

CONSORT proved influential, and other groups quickly developed guidelines for different research types. Reporting guidelines are like a theme and variations, where CONSORT forged a path others have followed with varying fidelity (See Table 5.5). Most have acronym names. Most were first published as a journal article describing their development. Some, but not all, have checklists and elaboration documents. Some guideline developers publish resources as separate documents, others put them all into a single journal article. Guidelines are developed by different groups, with different compositions (possibly including methodologists, editors, clinicians etc.) and in different ways (e.g., some by Delphi consensus [22]). Although most follow CONSORT's approach of presenting recommendations focussing on reporting above conduct, guidelines differ in how forceful their recommendations are and whether they also seek to influence design.

TABLE 1.1: A selection of highly cited reporting guidelines and their similar dissemination strategies

Guideline	Definition	Study type	Year	Resources
CONSORT	CONsolidated Standards Of Reporting Trials	Randomised controlled trials	1996 [19], updated in 2001 [20], and 2010 [23]	Development article Checklist Explanatory document Flow diagram Website COBWEB writing tool[24]
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses	Systematic Reviews and Meta-Analyses	2009 [25], updated in 2021 [26]	Development article Checklist Explanatory document Flow diagram Website
PRISMA-P	Preferred Reporting Items for Systematic review and Meta-Analysis Protocols	Protocols of systematic reviews and meta-analyses	2015 [27]	Development article Checklist Explanatory document Website

Guideline	Definition	Study type	Year	Resources
ARRIVE	Animal Research: Reporting of <i>In Vivo</i> Experiments	Publications describing research involving live animals	2010 [28], updated in 2020 [29]	Development article Checklist Explanatory document Website Action plans Compliance questionnaire
SRQR	Standards for Reporting Qualitative Research	Qualitative health research	2014 [30]	Development article Checklist (not editable) Explanatory document as supplementary material
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology	Observational Epidemiology	2007 [31]	Development article Checklists (four versions: cohort, case-control, cross-sectional, and combined) Explanatory document Website
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials	Protocols of clinical trials	2013 [32]	Development article Checklist Explanatory document Website

Guideline	Definition	Study type	Year	Resources
CARE	CAse REport	Case Reports and data from the point of care	2013 [33]	Development article Checklist Explanatory document Website Training course [34] Writing application [35]
STARD	STAndards for Reporting Diagnostic accuracy	Studies of diagnostic accuracy	2003 [36] and revised in 2015 [37]	Development article Checklist Explanatory document
AGREE	Appraisal of Guidelines, Research and Evaluation	Reporting clinical practice guidelines	2016 [38]	Development article Checklist Website
SQUIRE	Standards for QUality Improvement Reporting Excellence	Quality Improvement in Health Care	2008 [39] and revised in 2016 [40]	Development article Checklist Explanatory document Website Glossary

Guideline	Definition	Study type	Year	Resources
CHEERS	CHecklist and Explanation and ElaboRation taSk force report	Economic evaluations	2013 [41] revised in 2022 [42]	Development article Checklist Explanatory document
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis	Studies developing, validating, or updating a prediction model, whether for diagnostic or prognostic purposes	2015 [43]	Development article Checklist Explanatory document Website Adherence assessment form
RIGHT	Reporting Items for practice Guidelines in HealThcare	Practice guidelines in healthcare	2017 [44]	Development article Checklist (not editable) Explanatory document as supplementary material Website
COREQ	COnsolidated criteria for REporting Qualitative research	Interviews and Focus Groups	2007 [45]	Development article Checklist (not editable)

1.3 The EQUATOR Network unites the reporting guideline movement

As reporting guidelines grew in number and the problem of poor reporting gained recognition, the late Professor Douglas Altman, a leading medical statistician and evidence synthesiser, saw the need to catalogue reporting guidelines and form a community. He united academics from around the world to form The EQUATOR Network, often simply called EQUATOR, standing for Enhancing the QUAlity and Transparency Of health Research. It was the first coordinated attempt to combat poor reporting systematically and on a global scale. One of EQUATOR's core objectives was to create a database of reporting guidelines, accessible via their website where researchers will also find training and information about developing guidelines.

There are now over 500 reporting guidelines, representing the collective work of thousands of academics. The best-known guidelines are endorsed by large numbers of medical journals and the International Committee of Medical Journal Editors, and are amongst the 1% most highly cited publications indexed by Web of Science [46].

1.4 Reporting guidelines are part of a complex behaviour intervention

The publishing community began to take note. The International Committee for Medical Journal Editors encouraged journals “to ask authors to follow [reporting] guidelines” [47]. Concerned editors sought ways to adopt reporting guidelines, and more and more journals [48,49] have since found a range of strategies to introduce reporting guidelines into their policies, outlined in Table 1.2. There is variation in the degree of enforcement (from passive recommendation through to compulsory enforcement) and variation in the guidelines subject to the policy; PRISMA and CONSORT and more commonly enforced than other reporting guidelines, and many reporting guidelines in EQUATOR’s library are not enforced nor endorsed by any journals. Instead of listing reporting guidelines by name, many journals keep their instructions vague and merely recommend

authors find an appropriate guideline on EQUATOR's website. To my knowledge, no journal explicitly advises against using a reporting guideline.

TABLE 1.2: Example journal policies regarding reporting guidelines.

Policy	Example
Enforcing adherence	An editor or peer reviewer checks the article body for reporting guideline adherence and asks the author to add missing items.
Requesting peer reviewers use reporting guidelines	Editors ask peer reviewers to consider reporting guideline adherence as part of their review. Some editors may supply the reviewer with the relevant checklist.
Enforcing checklist submission	Editorial staff require authors to submit a completed reporting checklist as part of manuscript submission. Some journals may refuse to process a submission when the checklist is missing. Some journal submission systems may include fields for authors to upload their checklists, whereas other journals may expect authors to upload checklists as a supplementary file.
Journal endorsement	The journal's instructions to authors recommends authors follow reporting guidelines. Guidelines may be specified, in which case journals may link to guideline specific websites, to the guideline publications, or to the EQUATOR Network website. Sometimes journals include a general statement but do not name guidelines, instead referring authors to the EQUATOR website with an instruction to follow "relevant guidance".

Policy	Example
Publisher endorsement	Sometimes reporting guideline policies are set at the level of the publisher, as is commonly done for editorial policies. Individual journals may point authors to their publisher policies.
No policy	Journals have no policies regarding reporting guidelines.

Other stakeholders have begun incorporating reporting guidelines into their policies. Conferences like the Peer Review Congress ask applicants to use reporting guidelines when writing their abstracts. MedRxiv, a large preprint server, asks authors to declare they have “followed all appropriate research reporting guidelines, such as any relevant EQUATOR Network research reporting checklist(s)” [50].

EQUATOR have developed training programmes based on reporting guidelines. The training covers different ways to use reporting guidelines, including drafting manuscripts, checking manuscripts you have written, and appraising the reporting of someone else’s manuscript. Researchers have developed writing software to help authors apply reporting guidelines when drafting [51,52], online applications to facilitate checklist and flow diagram completion [53,54], and resources for reviewers to check compliance [55].

Hence over the years, a system has organically grown around reporting guidelines, driven by separate groups of people doing what they felt was sensible. This system includes the guidance resources themselves (the publications, checklists, flow diagrams), the websites that host those resources (guideline websites, the EQUATOR website, publisher’s websites), organisations that promote or enforce their use (EQUATOR, publishers, ICMJE, etc.), and staff at those institutions (researchers, editors, reviewers, etc.). These components all have the same aim: to influence what information researchers include in their articles.

As the Medical Research Council notes, a system with multiple components (like those listed above) is one of many hallmarks of a complex intervention: “an intervention

might be considered complex because of properties of the intervention itself, such as the number of components involved; the range of behaviours targeted; expertise and skills required by those delivering and receiving the intervention; the number of groups, settings, or levels targeted; or the permitted level of flexibility of the intervention or its components.”. The reporting guideline system exhibits these sources of complexity, as described in Table 1.3.

TABLE 1.3: Sources of complexity within reporting guidelines and the system drives their use.

Source of complexity	Example
Number of components involved	Reporting guidelines often consist of guidance documents, checklists, and flow diagrams, and other tools. These are disseminated through websites, publishing platforms, submission systems, and they are endorsed and enforced by stakeholders including publishers, the EQUATOR Network, conference organisers, and pre-print platforms.
Range of behaviours targeted	Guidelines comprise “reporting items”. Some items are relatively simple, like asking the author to specify their study design in the title. Others are harder, perhaps because they require time, expertise, or prerequisite tasks. For instance, some items may require authors to have conducted their study or analysis in a certain way, or to have collected particular information.
Expertise and skills required by those delivering and receiving the intervention	Academics from a particular field write reporting guidelines for their peers (as opposed to a lay audience), and so authors, editors, and reviewers must have sufficient expertise to use them.

Source of complexity	Example
The number of groups, settings, or levels targeted	<p>Groups: Users of reporting guidelines differ in their field of expertise, their experience, place of work.</p> <p>Settings: Although mostly written with authoring in mind, most guideline developers may also hope their resources are used by editors or peer reviewers for checking or appraising research articles.</p> <p>Levels: Guidelines are written with individuals in mind, but their efficacy is generally measured at group level (e.g., articles from a particular field published in a period time).</p>
Flexibility of the intervention or its components	There is variation between guideline content, resources, and the implementation strategies that development groups, publishers, and other stakeholders employ.

Viewing reporting guidelines as part of a complex behaviour change intervention may seem novel. Researchers often call reporting guidelines “tools” or “strategies” (e.g., [56,57]). In their history of the EQUATOR Network [58], co-founders Douglas Altman and Iveta Simera refer to reporting guidelines as “resources” that “influence” reporting, but do not call them interventions. However, I believe my perspective is not radical and I will now outline studies exploring the efficacy of reporting guidelines and argue they too take a systems perspective although they seldom acknowledge it explicitly.

1.5 Reporting guidelines have had little to no effect on reporting quality, and few authors adhere to them

Most studies exploring reporting guidelines' impact have used a non-randomised design, often comparing reporting quality before and after reporting guidelines were published and/or journals began asking authors to use them. For example, in 2023 Kilicoglu et al. [59] analysed 176,620 randomized trial reports published between 1966 and 2018. They used software to identify sentences pertaining to CONSORT methodology items. They found reporting quality in clinical trials has improved over time but remains sub-optimal: articles published before CONSORT was published reported 24% of CONSORT items, whereas articles published afterwards reported 48%. The fastest improvement occurred in the years immediately after CONSORT was published.

This result mirrors similar, manual, efforts to monitor reporting quality over time. For example, Dechartres et al. [4] looked at CONSORT items related to bias in 20,920 trial reports. They found information on sequence generation was missing from 69% of articles published before CONSORT (1986-1990), but only 31% of articles published after CONSORT (2011-2014). The proportion of articles missing information on allocation concealment fell from 70% to 45% over the same periods.

These two examples focus on clinical trials, but other research types have seen modest improvements too. De Jong et al. [60] reviewed 284 qualitative evidence syntheses. These syntheses appraised whether their primary studies had adhered to COREQ (a reporting guideline for qualitative research). Studies published before COREQ was developed reported 16 of COREQ's 32 items, whereas studies published after reported 18. Similarly, when comparing the reporting quality of genetic association studies, Nedovic et al. [61] found reporting quality was better in generic association articles published after STREGA but only in journals that endorsed STREGA (63% vs 49% showing full adherence).

Not all studies have found relationships like these. For example, Howell et al. found no

improvement in the reporting of quality improvement studies after the SQUIRE guidelines were published [62]. Similarly, Pouwels et al. found no improvement in observational epidemiology following STROBE's publication [56] and another study [63] found that reporting quality improved *before* STROBE was published, but not *afterwards*.

Other studies have explored the effect of different journal policies. For example, Hopewell et al. [64] assessed the reporting of clinical trial abstracts before and after the publication of CONSORT for Abstracts and compared journals that a) had no reporting guideline policy b) endorsed the guideline and c) actively enforced guideline adherence. They only found an effect in the active enforcement group, where 5.41 items (out of 9) were reported, which was 50% higher than expected. No improvement was seen in journals that endorsed the guideline without enforcing it.

Some observational studies have focussed on a single time period. In 2018, before my DPhil, I collaborated with the UK EQUATOR Centre and BMJ Open to compare manuscript versions before and after authors completed a reporting checklist [54]. Authors made few edits to their work. Of 20 included authors, three added information for a single reporting item, one added information for two items, and one added information for six items. The remaining 15 authors made no changes. On average, manuscripts described 57% of necessary reporting items before the author completed the checklist and 60% afterwards.

In 2018 the senior managing editor of the Journal of the National Cancer Institute asked 2000 submitting authors whether they had used a reporting guideline [65]. She then asked peer reviewers to rate manuscripts for their clarity and adherence to reporting guidelines. Declared guideline use was associated with better adherence to guidelines, but not associated with improved clarity nor acceptance rates.

Some experimental studies have tried to isolate the effect of journal policies by randomising authors. In 2015 PLOS One randomly allocated 1689 incoming submissions reporting *in vivo* animal research manuscripts to either a) request completion the ARRIVE reporting checklist or b) usual practice [66]. No article achieved full compliance with ARRIVE, and only one sub item (details of animal husbandry) showed improvement between groups.

In another study [24], 197 authors submitting to 46 participating journals were randomly allocated to receive either a) access to an online tool (WebCONSORT) to generate customised reporting checklists and flow diagrams based on CONSORT and its extensions or b) a standard CONSORT flow diagram generator without a reporting checklist. There was no difference in reporting quality between groups: authors only reported half of required items.

In two earlier studies, Cobo et al. explored the role of peer review. Simply providing peer reviewers with reporting guidelines had no effect [67]. Adding a reviewer whose task was to check for guideline adherence did lead to improved reporting quality [68], but “the observed effect was smaller than hypothesised and not definitively demonstrated”.

It would be unsurprising to find that the stricter the enforcement, the better the adherence. Pandis et al. [69] describe an enforcement strategy used in a small Dentistry journal where the associate editor checked manuscripts for adherence to CONSORT. The associate editor would complete a CONSORT checklist for each manuscript, making note of unclear or unreported items and suggesting ways to improve the manuscript. This would be sent back to the author. Resubmitted manuscripts were subject to the same process, and the manuscript would only be sent out for peer review once the reporting was deemed satisfactory. Over two years, 23 manuscripts were handled in this way. The policy was effective. All studies reported at least 33 of 37 CONSORT items (compared to 15 items before the policy was introduced). However, even with this heavy-handed approach, “four items were still unreported in all trials: changes to methods (3b), changes to outcomes (6b), interim analysis (7b), and trial stopping (14b).”.

To summarise, reporting standards may have improved over the last two decades. There is some evidence that reporting guidelines may have contributed to this change, but if they have, their effect has only been modest, and the bottom line is that most research still does not include the details these guidelines call for. There is huge room for improvement. In a systematic review of 124 studies assessing adherence to one of eight reporting guidelines, Jin et al [70] found 88% of studies reported suboptimal adherence to reporting guidelines. More recently, Dal Santo et al [15] performed a systematic review of studies assessing adherence to CONSORT, PRISMA, STARD, and STROBE. Of 148

studies published between August 2020 and June 2022, Dal Santo et al. only 6 described adherence as acceptable Table 1.4.

TABLE 1.4: Studies published between August 2020 and June 2022 assessing adherence to CONSORT, PRISMA, STROBE, or STARD and their extensions. Data from Dal Santos et al. [15]

Subgroup	No. of studies	Adequate	Inadequate	Mixed	Vague	No mention
All	148	6 (4%)	84 (57%)	10 (7%)	22 (15%)	26 (18%)
CONSORT	61	3 (5%)	41 (67%)	3 (5%)	9 (15%)	5 (8%)
PRISMA	59	1 (2%)	31 (53%)	7 (12%)	11 (19%)	9 (15%)
STARD	10	0 (0%)	6 (60%)	0 (0%)	2 (20%)	2 (20%)
STROBE	18	2 (11%)	6 (33%)	0 (0%)	0 (0%)	10 (56%)

1.5.1 Studies exploring the efficacy of reporting guidelines are actually exploring the efficacy of the reporting guideline system

Because reporting guidelines do not exist in a vacuum, it is difficult to separate the guidelines themselves from the policies, people, websites, and tools involved in their implementation. For example, many before-and-after studies use the publication of the guideline as their defining time point. However, these observational studies must disentangle the effect of guidelines coming into existence with the effect of subsequent journal policies and editorial practices. Similarly, in experimental studies comparing the effect of asking authors to complete a checklist or use a resource, the intervention groups included changes to editorial workflows. These changes were external to the resource being tested, but could be equally important to its success. For example, in the WebCONSORT study, editors' inability to identify randomised trial reports was an important source of failure external to the tool being tested. Hence, in describing reporting guidelines as being part of a complex behaviour change intervention, I believe I am explicitly articulating a systems perspective that previous studies have hinted at, and I am exploring that system in more granularity.

1.6 Evidence gap: What more could be done to improve guideline adherence?

Some of the articles I have cited end with rallying cries like “major improvements need active enforcement” [71]. It would be tempting to look at Pandis’ results as support for heavy editorial enforcement being the best option, but this approach may not generalise to other journals and other guidelines. The dentistry journal in this study was small. Only 23 manuscripts underwent this treatment over 2 years, and despite giving “30 to 60 minutes” of editorial attention to each manuscript, not all completely adhered to CONSORT. The study authors admit the benefits should be “considered in the light of the additional time requirement and need for greater editorial input during the peer review process”.

Other articles have called for lighter forms of enforcement; “We need to promote more active implementation, such as submission of the checklist with the manuscript” wrote

Dechartres [4], and “it is not sufficient for journals to simply recommend the use of STREGA to authors in the authors’ instructions; instead, journals should require submission of the STREGA checklist together with the manuscript” wrote Nedovic [61]. But the PLOS One [66] and BMJ Open studies [54] found little effect of checklist completion on reporting quality. Additionally, the PLOS One study found that enforcing checklists, although less burdensome than the editorial enforcement described by Pandis, still came with costs to both editors and authors and significantly prolonged publication times. Peer reviews focussing on reporting might help [68] but is relatively unexplored as an option and may be unacceptable to reviewers.

These studies all focussed on different methods of enforcement. Some include incidental findings hinting at areas-for-improvement unfixable by enforcement alone. For example, because reviewers assessing adherence in the PLOS One study [66] did not always agree or fully understand the guidance, the study authors suggested refining the guideline’s “content” and “perceived clarity”. In the WebCONSORT study, [24] Hopewell et al. made some guesses for why their intervention failed. They had to exclude 39% of manuscripts because editors had incorrectly identified them as randomised trials, and a quarter of authors selected inappropriate extensions. As a solution, they suggest a tool to help authors and editors identify study types. The study authors also raised other hypotheses to explain why their intervention failed: perhaps the custom combined checklists were too long, unclear, or perhaps giving feedback during manuscript revision was too late.

Neither the PLOS One study [66] nor the WebCONSORT study [24] explored these incidental findings. Both may have benefited from a qualitative component to understand why the interventions were not working. In our BMJ Open study [54] we surveyed authors after they completed checklists. Many reported finding the checklist too long, confusing, or irrelevant. However, because we used a multiple choice question with a (small) box for a free text answer, and because we did not survey authors if they *did not* complete a checklist, our results were fairly superficial.

Together, these incidental findings suggest authors may face challenges when trying to use reporting guidelines that require solutions beyond enforcement, and some groups

have tried to address these challenges. Noting the outcome assessors' confusion in the PLOS One study, ARRIVE's developers took steps to refine its clarity when they revised it [72]. They also decided to prioritise items to make the guidance quicker to apply. Hopewell et al. [24] created WebCONSORT because they worried that combining CONSORT with its extensions may be "cumbersome and difficult" without providing evidence for this claim. In 2018 I worked with the UK EQUATOR Centre as a freelance developer to create GoodReports.org, a website where authors could answer a questionnaire to find the right guideline, and then complete the checklist online (previously, some reporting guidelines came with non-editable PDF checklists).

These innovation efforts shared limitations. None took steps to identify barriers thoroughly, and by focussing on a few barriers they may have neglected others or introduced new ones. For example, in trying to make combining checklists easier, WebCONSORT may inadvertently made checklists longer, and increased the risk of authors selecting inappropriate guidance. Secondly, these studies did not systematically consider options to solve those barriers. For example, ARRIVE's development team decided to prioritize items as a way to make the guidance quicker to apply, but this is not the only solution. They could also have considered making guidance more concise, providing suggested wording or creating tools to speed up writing. Thirdly, although the studies describe their innovations, they do not always describe changes beyond the tool in question (e.g., changes to editorial practice) or how changes are expected to alter behaviour.

1.7 Summary

In summary, medical research is often poorly reported, and this makes it difficult for other researchers to understand, appraise, synthesize, or replicate studies. This, in turn, makes research less useful to patients and potentially harmful if it distorts the evidence base.

I've introduced reporting guidelines as recommendations, created by the research community with the aim to improve reporting quality. I've described how a system of tools, websites, people, and policies has organically grown around reporting guidelines, and I

have argued this system forms a complex behaviour change intervention with the goal of altering what authors write.

I've discussed how this system has had only a modest effect on reporting quality, at best. I've described how studies exploring modifications to this system are limited because they did not thoroughly explore barriers or facilitators that influence reporting guideline adherence, and lacked a systematic method to identify options to address those influences.

1.8 Aims and Objectives

My aim was to identify and address the influences affecting whether authors adhere to reporting guidelines. I wanted to explore the entire reporting guideline system, and I wanted to be thorough: I wanted to identify as many influences as possible, and as many solutions as possible, before deciding which to implement.

My objectives were:

- To identify influences that may affect reporting guideline impact by synthesising existing research and by evaluating the EQUATOR Network website (addressed in chapters 3 - 5)
- To work with key stakeholders to identify intervention changes to address these influences (addressed in chapters 7 - 8)
- To implement these changes (described in 9)
- To refine the new intervention in response to feedback from authors (addressed in chapter 10)

My thesis bares hallmarks of pragmatism. I used both qualitative and quantitative methods. Constraints (like time and access to participants) influenced my decisions. I balanced participants' views with my own; I sought to remove my views as much as possible in all chapters except for the workshops I conducted with EQUATOR (chapter 7) and when designing the intervention (chapter 9). I balanced inductive and deductive reasoning; my early chapters were exploratory and inductive, and my later chapters became increasingly deductive as my focus narrowed and I relied more heavily on

frameworks.

1.9 Thesis structure

I have structured my thesis in three parts. Part 1 (chapters 3 - 5) addresses my first objective, part 2 (chapters 6 - 9) addresses my second and third, and part 3 (chapter 10) addresses my fourth objective. I'll now give a short summary of each chapter.

Chapter 2 - Reflections on starting my DPhil

I reflect on my background, my prior held opinions, and those of my supervision team, and how these may have influenced the direction of this thesis.

Chapter 3 - What influences researchers when using reporting guidelines? Part 1: A thematic synthesis

The next three chapters pertain to my first objective - to identify possible influences affecting whether authors adhere to reporting guidelines. Chapter 3 describes a thematic synthesis of qualitative studies exploring authors' experiences of using reporting guidelines.

Chapter 4 - What influences researchers when using reporting guidelines? Part 2: Describing the questions asked in quantitative surveys

This chapter builds on the previous one by identifying additional possible influences from the content of quantitative survey questions.

Chapter 5 - Describing who visits EQUATOR's website for disseminating reporting guidelines, and how visitors use it

This chapter describes the characteristics and behaviour of visitors to the EQUATOR Network website. This website, although an important piece reporting guideline infrastructure, was rarely explored by studies reviewed in the previous two chapters. This evaluation adds evidence and context to influences identified in the previous chapters, and goes on to inform discussions and decisions in later chapters.

Chapter 6 - Selecting the Behaviour Change Wheel framework

The next 3 chapters pertain to my second objective — identifying intervention changes. Chapter 6 introduces the Behaviour Change Wheel, which is a framework for designing behaviour change interventions. I explain how my thesis gained form at this point in time; my view of reporting guidelines as a system crystallised, and Charlotte Albury joined my supervision team as my plans took an unexpected qualitative turn. I wanted my thesis to accurately reflect the twists of my DPhil, and so instead of introducing my chosen framework at start of my thesis (which may be customary), I introduce it in the middle to more accurately reflect my journey.

Chapter 7 - Following the behaviour change wheel guide: Workshops with reporting guideline experts

This chapter describes how I led workshops with reporting guideline experts from the UK EQUATOR Centre to identify intervention options using the Behaviour Change Wheel framework.

Chapter 8 - Generating ideas to address influences: focus groups with reporting guideline developers, advocates, and publishers

This chapter reports focus groups where I collected ideas on how intervention options could be realised.

Chapter 9 - Defining intervention components and developing a prototype

In this chapter, I bring together the outputs of the previous two chapters to create a table of intervention components. I then describe how I implemented these components by redesigning a reporting guideline (SRQR), the EQUATOR Network website's home page, and creating a web platform for reporting guideline developers to create and disseminate resources.

Chapter 10 - Refining the intervention: interviewing authors to identify deficient intervention components

In this chapter I address my final objective by refining the redesigned reporting guideline and home page in response to feedback from authors. I describe a qualitative study where I used observation, think aloud, structured interviews, and a writing evaluation, to gather feedback from an international sample of authors.

Chapter 11 - Discussion

After summarising my findings and outputs, I reflect on the strengths and limitations of my approach, next steps, and the impact I hope my work will have on scholarly publishing, the meta-research community, and implications for policy.

Figure 1.1 summarises the outputs of each research chapter and shows how chapters fed into each other.

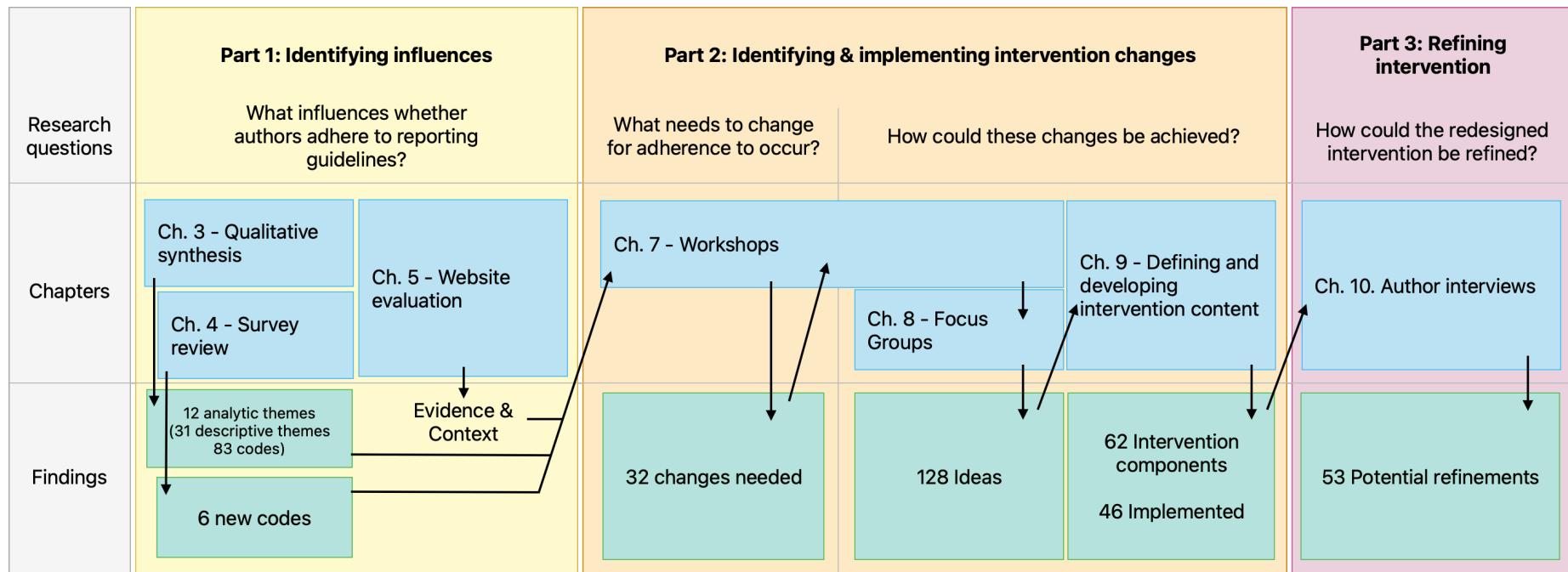


FIGURE 1.1: Graphical abstract describing the outputs of research chapters and the relationships between them

Chapter 2

Reflections on starting my DPhil

I use qualitative methods throughout my thesis. Reflexivity is important in qualitative research as “A researcher’s background and position will affect what they choose to investigate, the angle of investigation, the methods judged most adequate for this purpose, the findings considered most appropriate, and the framing and communication of conclusions” [73]. More succinctly, “Despite the sterility of the instruments, we never come innocent to a research task” [74]. In this chapter, I examine my own lack of innocence.

Reflexivity requires self reflection to acknowledge one’s own beliefs, bias, and judgement systems before, during, and after the research process (Shaw 2010). Researchers engage with self reflection to different degrees, and Woolgar [75] describes a continuum from positivists’ benign (often retrospective) introspection to constructivists’ radical constitutive reflexivity. My reflexivity falls in the middle and shifted over time. Reflexivity was a new concept to me when I started my DPhil and in my early chapters (3–4) I was mainly concerned with representing participants’ views accurately, thereby placing me at the positivist’s end of Woolgar’s continuum. In this chapter, I reflect on my experiences with reporting guidelines before starting my DPhil, and the experiences of my supervisors. This account documents my initial stance and helps explain where this work grew from. Wilkinson [76] describes this as personal reflexivity – how a researchers’ identity and experience may influence what they study.

Wilkinson [76] differentiates personal reflexivity from functional reflexivity (how researchers undertake research and how research shapes them in return) and disciplinary reflexivity (how prevailing paradigms, attitudes and power structures shape research). As my DPhil progressed my thoughts strayed further into functional reflexivity. I became increasingly aware of my inescapable role in knowledge generation in later chapters, especially in the workshops I led with reporting guideline experts (chapter 7), and when collecting feedback on design decisions (chapters 10). I kept a research diary throughout my DPhil and the remaining chapters end with shorter reflexivity sections where I consider how my beliefs may have shaped my research and how doing the research had, in turn, shaped my beliefs. This chapter, however, focuses solely on personal reflexivity.

This thesis marks the ten-year anniversary of my attempts to improve research reporting. My early interest came from reading about psychology's reproducibility crisis [77], the cancer reproducibility project [78], Ben Goldacre's book *Bad Pharma* [79], and Ioannidis' essay "Why Most Published Research Findings Are False" [80]. I was working in a psychology laboratory at the time, having recently finished a MSc in Neuroscience. The compelling Lancet series on research waste [81] drilled home a magnitude of inefficiencies in medical research. I was most stuck by the article on incomplete or unusable research reports [71], where a key takeaway was that "although reporting guidelines are important, major improvements need active enforcement" by editors or reviewers. I took Paul Glasziou's words as a call to action, and I set about creating tools to help editors enforce good reporting.

I began by writing software to check whether a manuscript cited a reporting guideline. Next I wrote code to check the reporting of statistical analyses. These were both pet projects. My first "proper" software was a manuscript checker that evaluates whether a manuscript adheres to journal guidelines [82]. I began working with academic journals, and came to appreciate that reporting guidelines are one of many priorities for editors, and often fall behind more pressing issues like ethics, consent, image duplication, and paper mills.

I later worked with the UK EQUATOR Centre to create GoodReports.org [83], a website

where authors can find and complete reporting checklists. EQUATOR and I collaborated with BMJ Open in 2018 to see whether authors improved their manuscripts after completing a checklist as part of manuscript submission [54]. The results were disappointing. Few authors made any changes.

I was awfully demoralized. It was tempting to blame the failure on lazy authors paying lip service to the checklists because they know that equally-lazy editors will not check. Indeed, this is a refrain I have heard bandied around by frustrated guideline developers. However, a tenet of software development is that when it comes to how a user experiences a product (or service), the user is never wrong. A software developer that blames poor design on “lazy users” will quickly find themselves out of a job. If users do not use your product, or do not use it how you would expect them to, then your product may lack usability (users cannot understand or use it), product market fit (users do not need or want it), or awareness (users do not know about it). It is the creator’s responsibility to make their product known, foolproof, and useful.

Understanding users’ experiences and needs is therefore central to making a successful product. It is also very difficult. Like many developers, I am happier behind a screen and talking to users never came naturally to me. Instead, I often latched on to an idea before understanding the people I was trying to help, or the problem I was trying to solve. This was a good strategy for building things nobody wanted.

I almost fell into the same trap when starting my DPhil. In my interview I adamantly pitched a tool to create personalised checklists by combining reporting guidelines with journal, funder, and institutional requirements. I envisioned something akin to Web-CONSORT on steroids. The old me would have happily spent months falling down that rabbit hole and building something (probably) useless; the custom checklists would have been incredibly long and probably confusing. The new me, the me writing this paragraph, is glad I paused and spent a year trying to better understand authors’ experiences and needs.

On starting my DPhil I had never used a reporting guideline myself, despite recommending others to do so. Yet I held them in high regard, as did my three initial supervisors. Jen, Michael, and Gary are affiliated with the EQUATOR Network. Gary is

Director of the UK EQUATOR Centre and an author of many reporting guidelines and a working-group member for many others. Michael studies reporting completeness, the robustness of reporting guideline development methods, and the consolidation of different reporting guidelines for the reporting of studies of nutritional interventions. Jen is involved in many studies investigating reporting and reporting guidelines, and runs training on writing and reporting guidelines.

Charlotte joined my supervision team in my second year, once we realised my thesis was fixed on a qualitative course. Qualitative behaviour change research was new territory for Gary, Jen, Michael, and I, and we needed an expert. Charlotte came to the rescue. She has a lot of experience in qualitative research and behaviour change theory, both from her own research and from leading the Oxford course on qualitative methods. However, she had never studied reporting guidelines before, nor any other meta-research phenomenon. Charlotte has published a suggested amendment to a reporting guideline she uses in her own work [84]. Whereas Jen, Michael, and Gary can be described as reporting guideline advocates, Charlotte was a little more cool-headed. Although she considered reporting guidelines useful for writing up quantitative work, she found guidelines for qualitative research frustrating because they were developed from a positivist perspective, and so did not fit all qualitative research.

In summary, I came to this DPhil with an existing passion for improving research, a deep-but-naïve respect for reporting guidelines and EQUATOR, and with determination to make something helpful after rebounding from the disappointing BMJ Open study. Sitting atop some rusty, decade-old research experience, I had a software developer's vocabulary and mindset. Core to this mindset was a belief that if people do not use what you have made in the way you want them to, it is not their responsibility to change their own behaviour. It is *your* responsibility to change what you have made. The challenge is figuring out what changes you need to make.

Having set the scene, I'll begin my thesis where I began my research: by trying to understand why authors do not adhere to reporting guidelines. In the previous chapter I described the evidence that reporting guidelines have had little effect on reporting quality. Next I wanted to know *why*. Chapters 3 - 5 describe my mixed methods approach

to finding possible answers to this question, beginning with a qualitative evidence synthesis.

Part I

Identifying influences

Chapter 3

What influences researchers when using reporting guidelines? Part 1: A thematic synthesis

3.1 Introduction

In chapter 1 I discussed evidence that reporting guidelines have only had a modest impact on reporting quality. Although a few of these studies put forward some possible explanations for their disappointing findings, these were often based on untested assumptions and few included any qualitative exploration of why reporting guidelines were (or were not) working.

I also described how reporting guidelines are disseminated in similar ways; mostly as checklists and an accompanying “Explanation and Elaboration” (E&E) document published in academic journals, often separately. I described how these resources sit within a broader system that includes the guidance itself, tools (e.g., checklists), websites (e.g., journal instructions, the EQUATOR Network website, guideline-specific websites), and the behaviour of others (e.g., editors, peer reviewers, co-authors, colleagues). In this chapter, my objective was to identify influences originating from any part of this system that may affect whether an author adheres to reporting guidelines. I did this by

synthesising qualitative studies exploring authors' experiences of reporting guidelines.

I will briefly describe what I mean by *influence*. Health behaviour researchers often talk of "barriers and enablers", or "barriers and facilitators". Following this tradition, I also intended to look for barriers and facilitators when first designing this study. I expected to end up with two tidy lists; a list of things that stopped authors from adhering to guidelines, and a list of things that helped. However, whilst coding I found the data did not fall into two binary categories so neatly. What hindered one person sometimes helped another, or what helped in one scenario did not help in another. I subsequently found that other researchers struggle to fit their data into these binary categories. Haynes and Loblay [85] address this struggle and argue against the use of "barriers and facilitators", which they describe as an "untheorised framing device that often rests on unexamined assumptions yet, because it is familiar, is chosen as a 'safe' approach by inexperienced qualitative researchers and those who seek ways of making qualitative research appear more palatable to reviewers". Haynes and Loblay argue that a barrier and enabler approach is a false friend; "If it can claim any theoretical basis", they argue, "it would be that the approach is loosely underpinned by a linear-rationalist behavioural paradigm", but this simplified linearity can lead to superficial findings that do not consider context nor encourage reflexivity.

After reading Haynes and Loblay I felt justified letting go of my barrier and facilitator plan. My next challenge was one of vocabulary. If I was no longer identifying barriers and facilitators, what noun could I use in their place? Haynes and Loblay didn't present me with any solutions here. Although they argue against using barriers and enablers, they use the terms exclusively. I tried "factors that influence reporting guideline adherence", but an editor complained that the word "factors" had a precise and protected definition in epidemiology, and if I were to use it readers would expect a quantitative cohort, cross-section, or case-control study. I expected this review to deepen my understanding of authors' experience of reporting guidelines and the surrounding system, and to reveal parts of that experience that may influence adherence. Ultimately I settled on the word "influence" instead of factors, barriers, or facilitators. The Oxford dictionary defines an influence as something that has "capacity to have an effect on the character, development, or behaviour of someone or something". Like the term "facilitators and

enablers”, It is untheorized, but unapologetically so. Its definition is sufficiently vague to accommodate context, nuance, and uncertainty, and it fits comfortably within my chapter’s research question: what might influence whether authors adhere to reporting guidelines?

3.2 Methods

In order to identify influences that may affect whether authors adhere to reporting guidelines, I performed a thematic synthesis of qualitative research that explored researchers’ experiences of reporting guidelines. When writing this chapter, I used the PRISMA-S guidelines for reporting systematic searches (see appendix A) and the ENTREQ guidelines for reporting qualitative evidence syntheses (see appendix B) [86,87].

3.2.1 Researcher roles and characteristics

In the previous chapter I introduced my supervisors Gary Collins, Charlotte Albury, Jennifer de Beyer and Michael Schlüssel, and our relationships with reporting guidelines. In this chapter I also worked with Shona Kirtley, a search specialist at the UK EQUATOR Centre and an author of the PRISMA-S guidelines for reporting systematic review literature searches, and I sought help from Yuting Duan and Lingyun Zhao at the Chinese EQUATOR Centre, run by Professor Zhao-Xiang Bian. As with all work in this thesis, I conceived, coordinated, and led all research.

3.2.2 Approach to searching and data sources

My search strategy sought all research publications reporting qualitative data exploring researchers’ experiences of using reporting guidelines. I wanted to capture the experiences of researchers from around the world, so I included international databases in my information sources. In addition to searching databases and websites, I emailed guideline developers and performed forwards and backwards citation searches. All data sources are listed in Table 3.1.

TABLE 3.1: Information sources and record management

Source	Search platform	Date searched	Record management
Chinese Biomedical Literature Database [88]	Chinese Biomedical Literature Service System [89]	25/10/2021	Yuting Duan used Zotero to manage, deduplicate and screen records.
China National Knowledge Infrastructure [90]	https://www.cnki.net/	"	"
Wanfang Data [91]	http://www.wanfangdata.com/	"	"
VIP Chinese Medical Journal Database [92]	http://www.cqvip.com/	"	"
Medline	Ovid	08/12/2021	I used Zotero to manage and deduplicate records. Michael Schlüssel used Rayyan to do a second deduplication and screen records.
Embase	"	"	"
Allied Complementary Medicine Database (AMED)	"	"	"

Source	Search platform	Date searched	Record management
PsycInfo	"	"	"
Latin American and Caribbean Health Sciences Literature [93]	WHO Global Index Medicus (GIM) [94]	08/12/2021	"
African Index Medicus [95]	"	"	"
Western Pacific Region Index Medicus [96]	"	"	"
Index Medicus for South-East Asia region [97]	"	"	"
Index Medicus for the Eastern Mediterranean Region [98]	"	"	"
Scientific Electronic Library Online [99]	https://scielo.org/en/	08/12/2021	"
Open Science Framework (OSF)	https://osf.io/	15/12/2021	"
Methods in Research on Research website [100]	http://miror-ejd.eu/publications/	14/12/2021	"
Emailing developers of guidelines listed in Table 2	n/a	08/01/2022	"

Source	Search platform	Date searched	Record management
Forward and backward citation searching	n/a	08/01/2022	"

3.2.3 Inclusion and exclusion criteria

I included published research articles reporting researchers' experiences of using reporting guidelines derived through qualitative methods. I excluded articles published before 1996, the year the CONSORT statement was first published. With the assistance of Yuting Duan, Lingyun Zhai, and Michael Schlüssel, I was able to screen studies written in Chinese, Spanish, and Portuguese as well as English, and so I excluded articles written in any other language during screening.

I included articles reporting feedback from researchers as part of guideline development. I decided not to include reporting guideline development studies where this feedback came exclusively from the development group members, as I considered this context to be too different to how ordinary researchers experience reporting guidelines.

Many research articles I found used a mix of quantitative and qualitative survey questions. I did not consider categorical survey questions with a free text option for "other" to be qualitative, but I did include findings from free text questions inviting participants to provide context to a previous (not qualitative) question. I describe the quantitative studies and the questions they asked in my next chapter.

3.2.4 Electronic search strategy

The UK EQUATOR Centre's search specialist (Shona Kirtley) helped develop comprehensive search strategies, which had a component for reporting guidelines and another for qualitative methods. I constructed my reporting guidelines component from the acronyms of frequently accessed guidelines (Table 5.5) with generic terms for reporting guidelines to capture guidelines not named explicitly. My qualitative component came from a review of search filters [101], which recommended a sensitive qualitative filter for systematic reviews [102]. I extended the filter to include descriptive methods because I knew some of my target records were mixed method surveys. I conducted scoping searches, but my search strategies were not peer reviewed before execution and I did not set up article alerts. My search strategies are reported fully in Appendix C.

TABLE 3.2: Reporting guidelines featured on the EQUATOR Network's home page

Name	Full name
CONSORT	Consolidated Standards of Reporting Trials
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P	PRISMA for systematic review protocols
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
STARD	STAndards for Reporting Diagnostic accuracy studies
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
CARE	guidelines for CAse REports
AGREE	Appraisal of Guidelines, REsearch and Evaluation
RIGHT	Reporting Items for practice Guidelines in HealThcare
SRQR	Standards for Reporting Qualitative Research
COREQ	COnsolidated criteria for REporting Qualitative research
ARRIVE	Animal Research: Reporting of In Vivo Experiments
SQUIRE	Standards for QUality Improvement Reporting Excellence
CHEERS	Consolidated Health Economic Evaluation Reporting Standards

3.2.5 Screening

I screened titles and abstracts to identify articles exploring researchers' experiences and then screened full texts to identify whether those articles used qualitative methods. A colleague from the UK EQUATOR Centre (Michael Schlüssel) double-screened a random 10% sample and we resolved differences through discussion. Yuting Duan from the Chinese EQUATOR Centre screened Chinese records in the same way and asked her colleague Lingyun Zhao for second opinions when necessary.

3.2.6 Describing and appraising records

I extracted study characteristics and used the Critical Appraisal Skills Programme Qualitative (CASP-Qual) checklist [103] to critically appraise included studies. I expected this appraisal to help me consider strengths and weaknesses of each study when synthesising them.

3.2.7 Synthesis methodology

I used thematic synthesis as defined by Thomas and Harden [104] because it can handle studies with “thin” descriptions, because it allowed me to infer influences from research that may not have addressed my concern directly, and because I expected its output, grouped by themes, to be useful to guideline developers.

I imported files into NVivo 12.0 for Mac and coded all sentences from the results section and relevant supplementary materials that reported qualitative findings. I assigned each sentence one or more descriptive codes that sought to distil the essence of what was written, creating new codes when necessary and without using a framework. I then used mind-mapping software [105] to visualise similarities and differences between codes and aggregate them inductively into descriptive themes. These themes were descriptive in that they captured the meaning of the codes they contained. I then used my research question to infer influences from these descriptive themes, thereby producing analytic themes. I discussed all steps with Jennifer de Beyer, and we resolved conflicts through discussion.

3.3 Results

3.3.1 Search

My search yielded 18 articles (see Figure 3.1 for full search results). I and Michael Schlüssel double-screened 10% of the non-Chinese titles and abstracts and agreed on 98.3% of them (170/173). We resolved the remaining three through discussion and consensus. Citations searches did not reveal any additional articles, nor did my emails with

guideline developers. All eligible records were in English; there were no eligible articles written in Chinese, Spanish, or Portuguese. Eligible articles included surveys, semi-structured interviews, focus groups, and writing tasks.

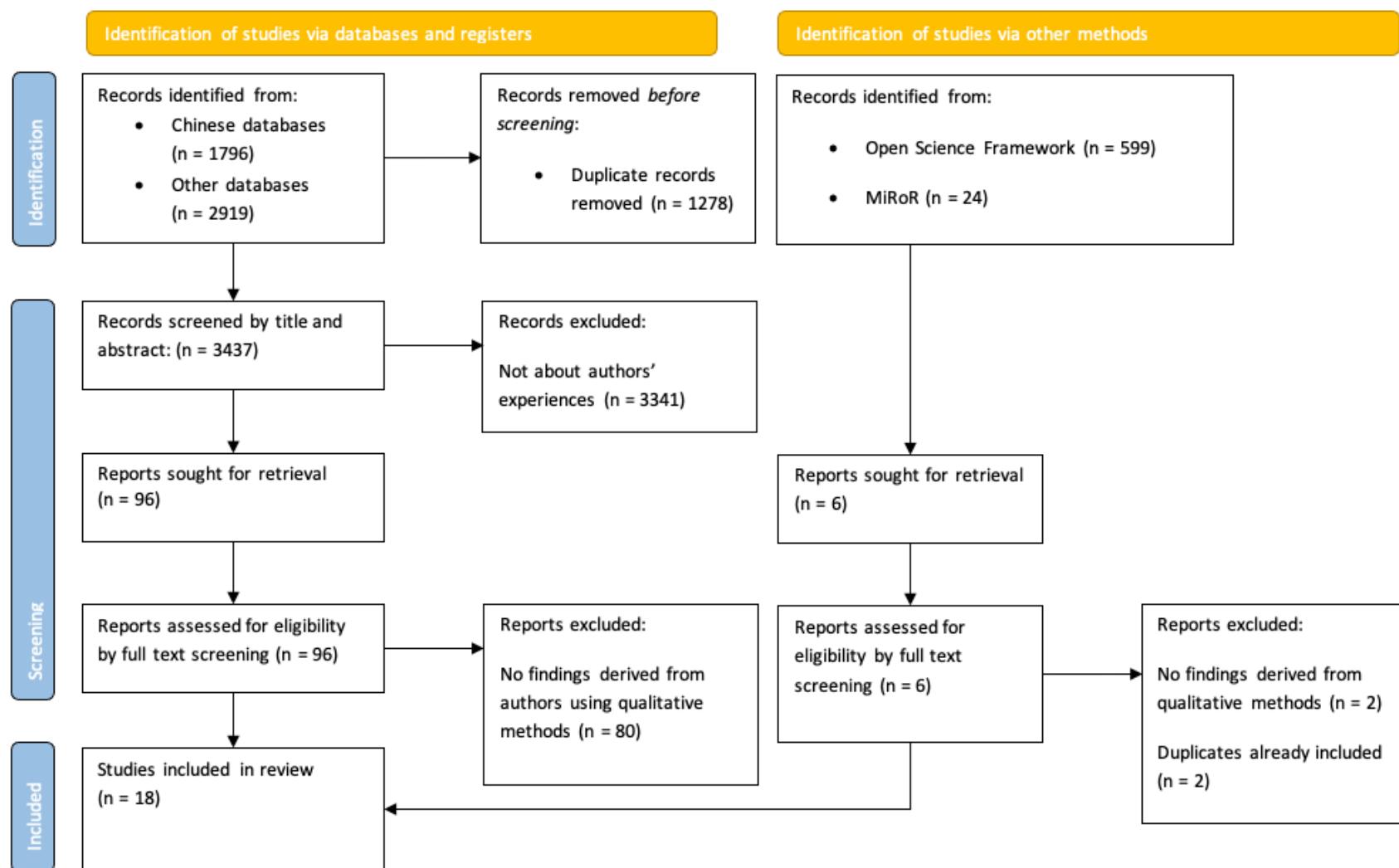


FIGURE 3.1: PRISMA flow diagram

Only 7 of the 18 records reported where participants came from. Three mixed method survey studies [57,106,107] included participants from a wide range of countries but it was not possible to tell which participants completed the optional qualitative questions. Three interview studies [108–110] included participants exclusively from North America, Europe, and Australia, whilst the fourth [29] included a participant from Brazil.

Critical appraisal of the studies using CASP-Qual rated the studies ranging from valuable to not very valuable; the less valuable studies had few qualitative components or minimal reporting of qualitative analysis or findings. I did not exclude any studies based on critical appraisal assessment. Study characteristics are reported in Table 3.3.

TABLE 3.3: Study characteristics

Guideline(s)						
Authors	Participants	Participant's country	considered	Methods	Phenomena of interest	CASP rating
Burford et al., 2013 [111]	151 systematic review authors	Not reported	PRISMA Equity	Mixed methods survey	Perceived utility, facilitators, and barriers	Fairly valuable
Davies et al., 2015 [108]	18 experts and 29 end users	USA, Canada, Sweden, the UK, and Norway	SQUIRE	Focus groups and interviews	Experiences with and impressions of the SQUIRE Guidelines	Valuable
Davies et al., 2016 [112]	44 graduates faculty, and directors of healthcare	Not reported	SQUIRE	Mixed methods survey and written exercise	“whether SQUIRE 1.6 was understood and implemented as intended by the developers”	Valuable

Authors	Participants	Participant's country	Guideline(s)			
			considered	Methods	Phenomena of interest	CASP rating
De Vries et al. 2015 [113]	7 researchers	Not reported	Systematic review protocols of animal intervention studies	Participants were asked to give feedback, but it is unclear whether this feedback was given in an interview or in writing.	Feedback on usability, missing items, possibilities for improvement and clarity	Not very valuable

Guideline(s)						
Authors	Participants	Participant's country	considered	Methods	Phenomena of interest	CASP rating
Dewey et al., 2019 [106]	74 out of 831 survey	The full survey was answered by respondents that provided (optional) free text comments	CONSORT, STROBE, PRISMA, STARD	Mixed methods survey	“(1) When and how are reporting guidelines and checklists used by authors and reviewers? (2) What is their impact on the content of final manuscript drafts according to authors? and (3) How do authors and reviewers perceive the value of reporting guidelines and related checklists?”	Fairly valuable
Eysenbach 2013 [114]	61 authors	Not reported	CONSORT Ehealth	Mixed methods survey	Views on completing the checklist as part of submission	Fairly valuable

Guideline(s)						
Authors	Participants	Participant's country	considered	Methods	Phenomena of interest	CASP rating
Fuller et al., 2015 [109]	5 authors	USA and Australia	TREND and Reporting Guidelines in general	Semi structured interviews	Factors that affected authors' use of TREND and other reporting guidelines	Valuable
Korevaar et al., 2016 [110]	4 radiology residents, 8 laboratory medicine experts	Radiology residents were from the Netherlands. No geographical details provided for experts	STARD	Interview (residents) and mixed methods survey (experts)	To identify items that were vague, ambiguous, difficult to interpret, or missing	Fairly valuable
Macleod et al., 2021 [107]	211 authors, but only some answered the free text question	The full survey was answered by participants in the USA, China, Japan, Germany, EU, and “Other” areas. It is unclear who answered the free text question.	Materials Design Analysis Reporting framework	Mixed methods survey	Whether the checklist was clear and useful	Fairly valuable

Authors	Participants	Participant's country	Guideline(s)			
			considered	Methods	Phenomena of interest	CASP rating
Page et al., 2021 [115]	110 systematic review authors and experts	Not reported	PRISMA	Mixed methods survey	Opinions on PRISMA and proposed changes	Valuable
Prady et al., 2007 [116]	40 authors	Not reported	Standards for Reporting Interven- tions in Controlled Trials of Acupunc- ture	Mixed methods survey	Experiences using PRISMA	Fairly valuable
Prager et al., 2021 [117]	5 of 18 survey respondents that answered the free text question	Not reported	STARD	Mixed methods survey	Barriers to STARD 2015 adherence	Fairly valuable

Guideline(s)						
Authors	Participants	Participant's country	considered	Methods	Phenomena of interest	CASP rating
Rader et al., 2014 [118]	263 systematic reviewers	Not reported	PRISMA	Mixed methods survey	Barriers or difficulties in meeting more detailed reporting standards in PRISMA	Fairly valuable
du Sert et al., 2020 [29]	11 authors	UK, USA, Belgium, Brazil	ARRIVE	Interview and writing task	Authors' opinions, interpretation, and experiences of updated ARRIVE guidelines	Fairly valuable
Sharp et al., 2020 [57]	203 of 1015 researchers that answered free text questions	The full survey was answered by participants in Africa, Asia, Europe, North and South America, Middle East, and Pacific Region. It is unclear who answered the free text question.	STROBE	Mixed methods survey	Experiences and attitudes towards STROBE	Valuable

Authors	Participants	Participant's country	Guideline(s)				CASP rating
			considered	Methods	Phenomena of interest		
Struthers et al., 2021 [54]	623 authors, 274 of whom answered	Not reported	Reporting guidelines in general	Mixed methods survey	The question asked, “What could I do to improve the guideline?”	Fairly valuable	
Svensøy et al., 2021 [119]	10 authors	Not reported	Not specified	Semi structured interviews	Experiences using guidelines or templates	Valuable	
Tam et al., 2019 [120]	230 authors, 62 of whom answered the open-ended questions	Not reported	PRISMA	Mixed methods survey	Opinions on PRISMA	Fairly valuable	

3.3.2 Synthesis findings

The relationships between my codes, descriptive themes, and analytic themes are reported in Table 3.4. I identified the following analytic themes: 1) Researchers may not understand guidance as intended or what reporting guidelines are, even if they think they do; 2) Researchers report a variety of reasons for using reporting guidelines, and that some are more important than others; 3) Researchers report using reporting guidelines for different tasks and wanting guidance to be delivered in ways that better fit their needs; 4) Using reporting guidelines has costs which researchers may feel outweigh benefits; 5) Reporting guidelines may need to be revised and updated; 6) Researchers may not be able to report all items, which can leave them feeling uncertain or worried; 7) Awareness and accessibility may limit reporting guideline usage; 8) Reporting guidelines may be more useful to less experienced researchers, but these researchers may find them harder to use; 9) Researchers want or need design advice, but reporting guidelines may not be the right place to find it; 10) Reporting guidelines can be harder to use if their scope is too broad, too narrow, or poorly defined; 11) Researchers may have to use multiple sets of reporting guidelines, multiplying complexity and costs; 12) Researchers may use checklists but never read the full guidance.

As mentioned in the introduction of this chapter, I identified that barriers and facilitators were not consistent experiences. What may be a barrier for one person might be an enabler to others or when occurring in different context, and so I refrained from labelling my analytic themes as one or the other and described them as *influences* instead.

TABLE 3.4: Codes (left, not bold), descriptive themes (bold), and analytic themes (right)

DESCRIPTIVE THEMES	
CODES	ANALYTIC THEMES
What does this mean?	Researchers may not understand the guidance as intended, or what reporting guidelines are, even if they think they do
What does this term mean? [54,108,110,112,115]	
What does this item mean? [54,108,110,112,115,116]	
How are these items different? [108,114–116]	
Have I understood this as intended? [108,112]	
Examples help me understand items [29,113,115]	
Why is this item important?	
Why is this item important? [108,110,115,120]	
Who is this item important to? [57,108,115]	
Does this apply to me?	
Have I understood the guideline's scope as intended?	
[54,115]	
Does this item apply to me? [54,108,114–116]	
Is this item optional? [108,116]	
I do not understand what reporting guidelines are	
What are reporting guidelines? [57,117]	
How should I use a reporting guideline? [109]	

DESCRIPTIVE THEMES**CODES****Reporting guidelines benefit me**

I find guidelines useful in general [54,106]

Guidelines make me feel confident [57]

Guidelines help me develop as a researcher [57,111]

Guidelines may help me improve my manuscript
[57,106,108,111,114]

I believe guidelines may help me publish more easily
[119]

I use guidelines because of other people

I may use guidelines because journals and editors tell
me to [57,109,111,119]

I may use guidelines because other researchers expect
it [109,119]

Guidelines benefit others

Standardized reporting benefits the community
[26,57,119]

Some benefits are more important than others

Immediate benefits are more important than
hypothetical ones [57,119]

Personal benefits are more important than benefits to
others [119]

ANALYTIC THEMES

Researchers report a variety
of reasons for using
reporting guidelines, and
that some are more
important than others

DESCRIPTIVE THEMES**CODES**

Researchers use reporting guidelines for different tasks

I use reporting guidelines for planning research [57,108]

I use reporting guidelines for designing research [57,106,116,117]

I use reporting guidelines for writing [57,106,108,116]

I use reporting guidelines for checking my own or other people's writing [57,117]

I use reporting guidelines to appraise the quality of other people's reporting [110]

I use reporting guidelines for peer reviewing [57]

I want guidance presented in formats that are better suited to the task I am doing

I want items presented in the order in which I must do them [26,113]

I want design or methods advice [57,108,115]

I want templates for writing [106]

I want checklists that are easy to fill in [54,107]

I want checklists embedded into journal submission workflows [106]

I want items embedded into data collection tools [111]

ANALYTIC THEMES

Researchers report using reporting guidelines for different tasks and wanting guidance to be delivered in ways that better fit their needs

DESCRIPTIVE THEMES**CODES****Guidelines take time**

Guidelines take time to read, understand and apply [109,111,119]

Some items require extra work which takes time and effort [108,112,118]

I want an indication of which items to prioritize [108,116]

Perceived complexity [106–108,119]

Long guidelines are off-putting [54,57,111,114]

Itemization may decrease costs

Itemization helps me navigate guidance [115]

Itemization summarizes the guidance [106]

Itemization may increase perceived costs

Itemization makes guidance appear longer [115]

Itemization blocks the bigger picture [108]

I think guidelines make my manuscripts long and bloated

Following reporting guidance can result in long, bloated articles [108,111,114,116]

Long, bloated articles may exceed journal word limits [107,109,114,116]

I want options for where to report this item [57,108,109,112,114,115]

The benefits of using a reporting guideline may not outweigh the costs

The benefits of using a reporting guideline may not outweigh the costs [57,109,114]

The balance of benefits vs costs may be more favourable when guidelines are used early

Guidelines are more valuable when used early [54,57,106,108]

ANALYTIC THEMES

Using reporting guidelines has costs, and researchers may not feel that benefits outweigh the costs

DESCRIPTIVE THEMES

CODES	ANALYTIC THEMES
I think the guidance could be improved	Reporting guidelines may need to be revised and updated for different reasons
I would clarify this item [115,116]	
I would move this item [108,112]	
I would split this item into two [108,113,115]	
I would add or remove items from this guideline [108,110,115,116]	
I would add or remove requirements from this item [29,57,115,116,120]	
Guidelines need to be kept updated	
Guidelines can become out of date [108]	
Guidelines need to be updated [115]	
I feel unable to report this	Researchers may not be able to report all items which can leave them feeling uncertain or worried
I cannot report this because I didn't do it [108,114–116]	
I cannot report this because of intellectual property issues [114]	
I cannot report this because it clashes with journal guidelines [115]	
I cannot report this because data was missing from my primary studies [111]	
Editors, reviewers or co-authors asked me to remove this item [116,118]	
I feel nervous or uncertain if I am unable to report an item	
I feel uncertain because I don't know how to say that I didn't do it [115]	
I feel worried that I will be judged for transparently reporting something I didn't do [57,115]	

DESCRIPTIVE THEMES

CODES

I can only use what I know about and have

I may not know that reporting guidelines exist, or what guidance exists [54,106,109,110,119]

I may not be able to easily access guidance [54,119]

Reporting guidelines are more valuable to inexperienced researchers

Reporting guidelines may be less valuable to experienced researchers [57,106,114]

Experienced researchers feel that they already know how to report [57,106,108]

Experienced researchers find guidance patronizing and feel untrusted [107,109,114,115]

Reporting guidelines can be hard to use at first but get easier with experience

Reporting guidelines can be hard to use at first but get easier with experience [108,109,119]

I want or need design advice

I want design or methodological advice [57,107,115]

I don't know how to do this item [108,115,116]

I think this guidance prescribes how research should be designed

Guidelines are procedural straightjackets [57]

This guideline is too prescriptive [57,115,120]

ANALYTIC THEMES

Awareness and accessibility may limit reporting guideline usage

Reporting guidelines may be more useful to less experienced researchers, but less experienced researchers may find them harder to use

Researchers want or need design advice, but reporting guidelines may not be the right place

DESCRIPTIVE THEMES**CODES****A guideline's scope can be unclear**

The guideline's applicability criteria are not clear
[54,106,110]

A guideline can be too narrow

This guideline isn't a perfect fit for me [54]
This guideline doesn't generalise [57,106,107,115,120]

A guideline's scope can be too broad

I don't want to see optional items that only apply to other types of study [54,116]

I often need to adhere to multiple sets of guidance

I need to adhere to journal guidelines or other research guidelines [106,109,115,116]

I might need to use multiple reporting guidelines [57]

I want guidelines to harmonize

I want reporting guidelines to be linked or embedded [110,115]

I want reporting guidelines to use similar structure [115]

I want reporting guidelines to use similar terms [115]

I experience reporting guidelines primarily as, or through, checklists

I don't like checklists [54,57,106,114]

I may use the checklist instead of the full guidance [29]

ANALYTIC THEMES

Reporting guidelines can be harder to use if their scope is too broad, too narrow, or poorly defined

Researchers may have to use multiple sets of reporting guidelines, multiplying complexity and costs

Researchers may use checklists but never read the full guidance

1) Researchers may not understand guidance as intended or what reporting guidelines are, even if they think they do

Researchers commonly stated that they need more information to fully understand the intention of the guideline developer. When asked about the clarity of guidance, researchers across many studies reported difficulty in understanding certain terms, concepts, or checklist items:

“Does [the word *outcome*] mean the domain, or does it mean the domain measure, metric, method of aggregation, and time? [115].”

“Primary and secondary improvement related question is confusing, what does that mean?... I had a hard time with the [difference between the] improvement question and the study question.” [108]

A few researchers reported ignoring an item if they could not understand it:

“Only one item was identified as hard to understand by more than one respondent: ‘methods employed to ensure completeness of data’, which two participants said they left out because of difficulty in comprehending the item” [112]

Some researchers reported feeling that reporting guidelines were “simply not comprehensible” [108]. Others reported that they had understood, but further investigation revealed that their interpretations could be “different from that intended by the developers” [112]. For example, Davies et al. [112] found that one SQUIRE item “was reported as used fully [in the quantitative survey], but the qualitative analysis revealed that its usage was frequently inconsistent with the intention of the developers”. One reason for this may be because different researchers may interpret the guidance in different ways depending on their prior experience, the research context, or if the guidance is ambiguous. For example, the SQUIRE developers found that the word ‘theory’ “meant different things to different people. For some, the word ‘theory’ meant ‘mechanism by which an intervention was expected to work’, for others it meant ‘lean or six sigma for example’, and for still others it meant ‘logic model’ ” [108].

Even when researchers reported understanding what an item meant, they may not have understood why it is important or who it is important to, leading them to remark that an item “seems unnecessary” [29]. Few researchers referenced the needs of evidence synthesizers or patients as consumers of research, but more reported considering whether an item would be useful to other researchers, editors, and reviewers:

“the information provided does not matter as the reviewers do not know what to do with it’ ’[57]

In addition to not understanding guidance or who it is important to, many researchers expressed difficulty understanding whether an item was applicable to their work. Some reporting guidelines specify that not all items are compulsory or that some items may only apply to a subset of research articles. Researchers highlighted that this may not be obvious, especially if this nuance is buried in a long elaboration document. Some researchers therefore reported uncertainty over which items applied to them:

“Authors asked for clarification of which items were always required and which were nonessential” [116]

“Not always clear what was relevant to their study” [29]

“He had realised with experience and re-reading the Guidelines that SQUIRE did not require him to include every item in the manuscript”.[108]

This uncertainty may extend to the entire reporting guideline if researchers don’t know when to use one over another. One researcher declared that “PRISMA guidelines can also be used rather than the MOOSE” [54], when the two are primarily for reviews of intervention studies and observational studies respectively. Sometimes there may not be a perfect reporting guideline for a given study, as one researcher commented after using ARRIVE (which focusses on experimental research involving laboratory animals):

“Our report was an animal based cadaveric study looking at accuracy of drill guides. I were unsure which category it should fall under.” [29]

Even if a researcher understands the guidance, why it is important, and why it applies to them, they may not understand how to report it or “how much detail to report”

[29]. Some researchers “used examples [included in the guidance] to understand what should be reported” because they “demonstrate what is meant in practice” [29].

At a more fundamental level, researchers varied in their understanding of what reporting guidelines are. Often researchers would talk about reporting guidelines as if they were design guidelines, e.g., describing STROBE as “woefully decent in encouraging...use of appropriate data analytic approaches” [57]. This suggested that the researcher had not noticed the stipulation that “these recommendations are not prescriptions for designing or conducting studies” included in STROBE’s explanation and elaboration document [31]. Other researchers wrote about STARD as is if the guidance was to be used when collecting imaging data:

“Two comments suggested that reporting quality may be impacted by the physical environment in which [...] data are collected. These comments may indicate an incomplete understanding of reporting guidelines which pertain to reporting results during manuscript writing, not the process of imaging acquisition itself.” [117]

2) Researchers report a variety of reasons for using reporting guidelines, and that some are more important than others

Some researchers listed personal benefits to using reporting guidelines. Some described reporting guidelines as a “training tool” [111] for personal development, noting that guidance helps “develop a strong foundation and habits” [57]. Some talked about how guidance made them feel: “As a junior scientist it gives me confidence to request the reporting of a certain piece of information” [57]. Others said that reporting guidelines are “a helpful reminder” [111] and that going through the checklist “improved their manuscripts” [114]. Some saw value in fostering a “transparent reporting process” [106] and for making sure your “project is written up [...] rigorously” [108].

A couple of researchers noted altruistic benefits, reporting that widespread reporting guideline adherence “helps in standardizing how research is reported” [57] and calling “for more scientific reports to be published, preferably using a template or guideline to make them comparable” [119].

In the absence of anticipated benefits, some researchers said that they use reporting guidelines simply because “it was what was implicitly expected of them to do” [109], and that these expectations came from journals and their peers. Some used “tools promoted by journals, which often promised to ease the publishing process” [119] but others wrote that they found this to be an empty promise:

“I have never had (nor have I heard of) an editor or reviewer pushing back on a claim that all STROBE criteria were met. Therefore, when a STROBE checklist is required for manuscript submission, it seems to turn into a[n] exercise in additional administrative busywork without really improving the research.’’[57]

A few researchers reported being more likely to comply with journal requirements if they thought the journal was likely to enforce them: “Does the journal only suggest or actually require submission of a reporting guideline checklist?” [109]. Some said they were more likely to comply if “it was a high impact factor journal and I thought that I would only get one crack at it” [109].

A few researchers compared different motivations for using reporting guidance, noting that personal, guaranteed, and immediate benefits were more motivating than hypothetical benefits or benefits to others:

“I suppose you are looking for short-term gain, short-term benefits as a writer of a report” [119]

“it can be difficult to put the energy into using STROBE (or any other) one a priori since ultimately, it depends on the journal submitted to and accepted to” [57]

“All the researchers wanted more homogenous reporting but emphasized that:”As an individual reporter, one is prone to choose the easiest and most accessible one.” ” [119]

3) Researchers report using reporting guidelines for different tasks and wanting guidance to be delivered in ways that better fit their needs

Although reporting guidelines were designed to help researchers draft and check their manuscripts, many researchers mentioned using reporting guidelines for other tasks including designing research, planning, peer-reviewing, and educating. A few researchers suggested different ways that the guidance could be delivered to make their task easier. For example, some “thought [the] order of items should reflect [the] order considered when designing the experiment” [29]. Others wanted “a manuscript template” [106] to make writing easier. Some suggested that “online form[s]” [107] or software to “mark in the text what corresponds to each item in the list” [54] would make it easier to complete a reporting checklist as part of journal submission.

4) Using reporting guidelines has costs which researchers may feel outweigh benefits

Researchers noted that some items require extra work, either to collect the necessary information or just to think about and report, and that sometimes this workload felt overly burdensome:

“If [reporting guideline developers] put the onus on everybody out there who’s trying to improve care to deal with that sophisticated question [...], I just think [they] are putting a barrier in place that is going to be a mountain”
[112]

This work requires time, and “the length of time it would take to consider the items” [111] was cited by many researchers as a cost, with some asking themselves whether “sufcient time [was] available to comply with [the] reporting guideline” [109].

Researchers noted that a reporting guideline’s “length and content is a key factor influencing the time needed to complete it.” [57]. Some found checklists to be “very complete, but to follow every single point is overwhelming” [114]. As a solution, many wanted to “simplify” or “shorten the checklists” [54]. A few researchers wanted a “hierarchy” to know which items were most “important to include” [112]. Another suggested

that checklists presented as online forms could include “logic for irrelevant [items]” [107] so that the users are presented with only items that apply to them.

Complexity was sometimes mentioned alongside time: “As the research often was performed out of work hours, the required time and complexity of the guidelines or templates may have played a crucial role [in deciding whether to use a reporting guideline]” [119]. Researchers had conflicting opinions about whether itemization reduces perceived complexity. Proponents noted that the “Checklist is a very helpful summary of sometimes confusing guidelines” [106] and that itemization made guidance “easier to follow” and “more approachable” [115]. But a few said that presenting guidance in small pieces made it difficult to “get the whole picture of what you are supposed to be doing” [108] and that itemization makes “the checklist appear more daunting for users” because it adds vertical length. Consequently, “If you make the checklist too long people will see it as too complicated and then won’t use it” [115].

Another concern cited by many researchers was that following a reporting guideline can result in long reports:

“I use SQUIRE a lot for planning—I complete the sections up through the methods at the time I design the study... [but] SQUIRE creates sort of long reports if followed exactly.” [108]

“the document you create if you use SQUIRE exactly as written is unintelligible” [108]

“this [item] would require another paper” [114]

This problem was exacerbated by journal word limits:

“I believe it is a useful instrument but it is unrealistic to assume that every single suggestion can be detailed in a 6000-words manuscript.” [114]

“two remarked that word limitations has necessitated removal of many items” [116]

Although a handful of researchers noted that “the relaxation of word limits” [109] would help, many researchers objected to long articles because they were bloated, harder to read, or simply “unintelligible” [108] and requested strategies to “enhance readability” regardless of journal policies. Some wondered where they could place this information outside of the article body, such as “in an appendix”, an “online supplement or repository”, or a figure [115]. Some researchers preferred to report information in the checklist instead of the article body because of “space restrictions, because [it was] a minor component of the study, because they considered the information to be obvious, or because they were unsure of how to incorporate it in the manuscript.” [29]. Some used this strategy to report items that had “not been used or observed during the study, for example that no inclusion or exclusion criteria had been set, no data had been excluded, randomisation and blinding had not been used...” [29] although it was not clear whether this was motivated by a desire for a concise article or a concern about highlighting potential weaknesses.

Faced with the costs of time, work and article length, some researchers explicitly weighed perceived benefits against costs and disagreed about the balance:

“The manuscript has improved. However, I felt that the amount of effort was considerably greater than the degree of improvement.” [114]

“it also adds to the time required to put together a manuscript, and I am not sure how much it improves the chances of a manuscript being published”
[57]

“it does increase the quality of the articles, it is clearly worth the time” [57]

The balance of costs versus benefits may be most favourable when guidance is used early in the research workflow. Researchers who used reporting guidelines earlier in their workflow (e.g., for planning research or drafting) used language that implied it was something they did regularly (e.g., “I use SQUIRE a lot for planning” [108]). Some reported that they had come to this habit by their own initiative and that reporting guideline developers should “encourage people to use the criteria early in the writing process (I have, which probably is why I only changed one thing [at the point of submission])”

[54]. One researcher suggested that “policy that focuses on a front end approach would be helpful” [57], noting that “To fully apply the criteria, I would need to systematically apply the STROBE criteria on the front end design of a project, grant, etc. rather than at the time of writing a project” [57].

Conversely, many authors who completed a checklist during manuscript submission, very late in their workflow, emphasised the costs, using words like “arduous” [57] and expressing negative opinions of this process (see *Researchers may use the checklist but never read the full guidance*). This may be because researchers lack the motivation, time, or ability to edit their manuscripts at this point.

5) Reporting guidelines may need to be revised and updated for different reasons

Researchers in most studies had opinions on how guidance could be improved through clarifying, reorganising, splitting, merging, adding, or deleting items, and sometimes these views fed into the revision of reporting guidelines [112,115]. This feedback may be useful for reporting guideline developers. Even if a reporting guideline was considered perfect at one point in time, researchers noted that guidance must be kept up to date in response to changes in the field and broader scientific ecosystem:

“The evolution of the healthcare improvement scholarly literature in the intervening years since the publication of the SQUIRE Guidelines has led to the development of concepts that were not fully anticipated at the time of initial release” [108].

Updates to one reporting guideline may necessitate the update of another. For instance, as PRISMA was being updated, a few researchers “supported referring to PRISMA for Abstracts, but suggested it also needs updating” to reflect updates being made to PRISMA [115].

6) Researchers may not be able to report all items, which can leave them feeling uncertain or worried

Some researchers described being unable to report items because of external factors, including intellectual property or data rules, disagreement between co-authors, or because “peer reviewers or editors had suggested editing out much of their [reporting guideline]-specific text” [116]. Others reported feeling unable to report an item because they did not do it, whether on purpose, due to an oversight, or because requirements had changed since the study began:

“[This item was] not part of the study objectives” [114]

“This [item] is a good idea, but I did not do this.” [114]

“The RCT was initiated before trial registration became customary in Norway, and therefore does not have a Trial ID number.” [114]

This left some researchers fearing that “an incomplete checklist [gave] the impression that their study is less than perfect.”[57]. Some expressed concern that strict wording that assumed something was done may “force people to lie/mislead by asking a question they cannot answer” [115] and suggested that guidance should instead use more agnostic language and specify what to do if an item were not addressed, such as “If no publicly accessible protocol is available, please state this” [115].

7) Awareness and accessibility may limit reporting guideline usage

Researchers may not know what guidance exists and may be more likely to use whatever is most accessible and discoverable:

“Several of the researchers did not have extensive knowledge about the different reporting tools, so the accessibility of the guideline or template was often a decisive factor.” [119]

One researcher wrote that “poor dissemination strategy by authors of reporting guidelines had inhibited uptake” [109], and others recognised that reporting guidelines could be “better highlighted” [106] by journals or advertised on “social media platforms” [54].

8) Reporting guidelines may be more useful to less experienced researchers, but these researchers may find them harder to use

Some researchers reported that they didn't need the guidance as they were experienced enough to know what they were doing:

“One of the most prevalent themes was the expression of self-assuredness. ‘[I] follow the STROBE guidelines in my reporting reasonably well without actually referring to them or using a checklist’ (group 3, ID1) and ‘[I] already apply the STROBE recommendations despite not having heard of it until today’” [57]

Sometimes this was accompanied by an acknowledgement that reporting guidance may be more beneficial to less experienced researchers:

“Despite experienced researchers generally not seeing a benefit to personally using STROBE, there were strong feelings that it is valuable to early-career researchers” [57]

“Helpful at beginning of career, but not at later stage” [106]

“this exercise might be good for college students but is insulting for professionals” [114]

However, less experienced researchers often reported finding “reporting guidelines being difficult to use initially” [109], or that a reporting guideline became easier after repeated use, after using other reporting guidelines, or with experience in medical writing in general. For instance, “Participants with less experience in scholarly medical writing found the SQUIRE Guidelines harder” [108].

9) Researchers want or need design advice, but reporting guidelines may not be the right place

Many researchers reported wanting advice on design choices but disagreed on where that design guidance should go. Some researchers suggested referring researchers to other design resources through hyperlinks or citations. Others explicitly wanted design

guidance to be written into reporting guidelines so that others would read it. Some went as far as calling for reporting guidelines to express an opinion and encourage one technique over another. One researcher objected to a “neutral tone” [115] in a reporting guideline that may give the impression that a design choice (that they disapproved of) was reasonable practice.

However, other researchers objected to reporting guidelines that were opinionated about design choices. One user described STROBE as a “procedural straightjacket” [57], suggesting that it dictates how studies should be conducted. Users who encounter the guidance late in writing may be unable to act on any design recommendations and consequently may feel fearful of reporting transparently if their design choices deviate from what the guideline recommends as best practice (see *Researchers may not be able to report all items, which can leave them feeling uncertain or worried*).

Perhaps with these concerns in mind, one wrote about the “need to make sure that the language around this elaboration gives [researchers] some flexibility” [115], with another noting they were “OK with the idea of emphasizing the value of [this design choice], but [they would not] mandate it” [115].

10) Reporting guidelines can be harder to use if their scope is too broad, too narrow, or poorly defined

Reporting guideline developers may narrow the scope of their guidance by limiting it to certain design choices or research contexts. This frustrated some researchers, who noted that narrow “checklists cannot fit all types of research” [106] and “cautioned that ‘balance between freedom and structure is important to consider’ [...] and that it is ‘important to recognise that each study/analysis is unique and doesn’t always fit with the recommendations’ ”.[57]

Scope may not always be clearly communicated. One PRISMA user opined that “the assessment of risk of bias, statement of risk ratio and explaining additional analyses depend on the study design ... [For] a systematic review of cross-sectional surveys or a meta-synthesis I do not need this information” [120], suggesting they were unaware of PRISMA’s focus on interventional studies or that MOOSE and ENTREQ would be more

appropriate for these kinds of studies (see previous themes for further discussion of awareness and understanding the applicability of reporting guidelines).

Researchers noted that scope could be made broader by removing items or, more commonly, by extending items with more options and examples:

“omit”(benefits or harms)” from the checklist item to be more inclusive of reviews that do not examine effects of interventions” [115]

“If the new PRISMA will more explicitly embrace topics other than interventions (which I think it should), then some additional examples could be added to the parenthesis (e.g., sensitivity and specificity, disease prevalence, regression coefficient)” [115]

However, extending guidance with options can make the guidance appear longer and means researchers must work out which parts apply to them.

11) Researchers may have to use multiple sets of reporting guidelines, multiplying complexity and costs

There are now over 500 reporting guidelines indexed on the EQUATOR Network website, with more added each year as reporting guideline developers seek to cover more and more use cases. Researchers may be expected to use a second or third reporting guideline in addition to the original one. Reporting guidelines that are intended to be used in addition to another are called extensions. Some researchers “pointed out that these extensions have created needless complexity and additional confusion in reporting of observational studies [...] and that the number of extensions has become excessive, especially given that multiple extensions may apply to a single study” [57].

One researcher wrote: “it would be good to have better connection between different checklists (perhaps using digital linking, decision-trees, etc.)” [115]. Some showed concern that hyperlinks to extensions will go unused and so developers should “incorporate all relevant details in the [...] checklist and elaboration (in case authors don’t read the extension)” [115]. When writing about PRISMA, one researcher noted that “it would be wise to limit the number of additional documents to look up. This is only item 7, and

I have already been referred to PRISMA for Abstracts and PRISMA for Searches. As a systematic review author, reviewer, or editor, I would be unlikely to go to several sources for reporting guidance” [115].

A few researchers wrote that related reporting guidelines should be mutually updated to keep in sync with each other before linking or embedding them. Researchers wanted the instruction, terminology, and structure of different sets of reporting guidelines to be coherent, suggesting, for example, that the updated PRISMA should be structured to be “in line with PRISMA-P” [115].

In addition to reconciling multiple reporting guidelines, researchers must also comply with journal, funder, and other scientific guidelines and expressed frustration when instructions contradicted each other. For example, some reporting guidelines specify subheadings for abstracts and one researcher pointed out that a “major issue is that journals wildly differ in requirements/what is allowed in abstracts” [115].

12) Researchers may use checklists but never read the full guidance

Reporting guidelines typically consist of the guidance itself and a checklist that serves as a summary of the guidance and a tool to demonstrate compliance. Sometimes the document containing the full guidance is called the Explanation and Elaboration (or E&E for short). When talking about a reporting guideline, it was often unclear whether the researcher was talking about the checklist or the E&E.

Some researchers implied that their only experience with reporting guidelines was completing a checklist as part of submission. I noticed that many negative statements were directed specifically at this process, describing checklists as “painful” [106], “pedantic”, “annoying” [57], or a “stupid exercise” [114].

One study explored researchers’ use of checklists and E&E documents, noting that “Participants used the guidelines and the E&E in different ways. Some did not read the E&E and used only the checklist, others read the E&E first and then used the checklist and a further group used the checklist and referred to the E&E for help with specific items.” [29]. One researcher even went as far as to say that the “E&E appeared to be redundant” [29].

If some researchers only use checklists, which typically lack any nuance included in the E&E, this may explain why some described reporting guidance as inflexible and prescriptive, warning that “Blind checklists are not relevant to most work” [54] or that “Authors may fear the ‘Checklist Manifesto’ becoming a rigid bureaucracy, and also becoming contrived” [57].

3.4 Discussion

Researchers face many challenges when trying to use reporting guidelines and have many questions, opinions, and suggestions that could be useful for reporting guideline developers. Researchers also report personal benefits, especially personal benefits when using a reporting guideline early. These benefits are at odds with how reporting guidelines are typically disseminated (as pre-submission checklists) and presented (as a benefit to others). The polarity and severity of these influences differ according to context (e.g., when or how a reporting guideline used), and personal characteristics (e.g., experience in academic writing). These findings could help increase the impact of reporting guidelines if taken into consideration during their development, dissemination, and implementation phases.

The reporting guideline development community has typically relied on journals to promote their resources and have called on editors to better enforce reporting guideline adherence [121]. However, the results presented here suggest that focusing solely on enforcement may be short sighted and that guideline developers have the power to address many influences, at least in part. Doing so may in turn make it easier for journals and funders to enforce reporting guidelines. For example, it is difficult to enforce a reporting guideline that is complicated to understand or if the guideline’s applicability criteria are unclear.

Qualitative methods uncovered issues that may be masked by quantitative surveys. For example, Davies et al. [112] found that one SQUIRE item “was reported as used fully [in the quantitative survey], but the qualitative analysis revealed that its usage was frequently inconsistent with the intention of the developers”. Most studies used mixed method surveys and often the qualitative component was small, perhaps limited to

a single question like “Please add your comments and suggestions in the free text below” [106] or “Any other feedback?” [107], and consequently resulted in thin data. Future studies seeking richer qualitative data should consider using interviews and focus groups above surveys.

Despite there being hundreds of reporting guidelines, my search found only 18 studies that collected qualitative data, covering 12 reporting guidelines in total, and only six of the 15 reporting guidelines listed on the EQUATOR Network’s homepage. Of notable absence were the COREQ and SRQR guidelines for qualitative research. Given that qualitative research differs from quantitative research in terms of its ontologies, epistemologies, and how it considers replicability and best practice, it would be interesting to know whether qualitative researchers face additional influences not covered here.

Many of the themes centred on understanding and usability. For example, not understanding what reporting guidelines are, their scope, their content, or how to use them. These issues could be resolved through user-friendly design and refinement in response to feedback, but reporting guideline developers may lack the funds, time, motivation, or expertise to design user-friendly resources or effectively test them. The EQUATOR Network guidance for reporting guideline developers [122], published in 2010, covers steps from inception to dissemination but largely neglects user experience or user testing. Only two short sections, totalling eight sentences of an eight-page document, address the importance of gathering user feedback, but the guidance offers no instruction on how to do this. Nor does the guidance advise on usability best-practices. Reporting guideline developers may benefit from advice on how to design and refine their resources, and assistance in conducting user testing.

3.4.1 Limitations

I was limited by the availability of literature and the relative thinness of some studies’ qualitative analysis. Most studies relied on participants recalling what they had done or thought in the past, and so may be subject to recall bias. Future studies could consider using ‘in the moment’ methods like think aloud tasks.

Surveys could be subject to question order bias. Often the qualitative question appeared at the end of a survey, and so proceeding quantitative questions could have influenced participants' responses. In the next chapter I describe how most concepts explored by the quantitative survey questions also appeared in the qualitative data and some of these quantitative questions were leadingly phrased. However, the qualitative data contained many additional themes that did not appear in the quantitative questions and so I believe the results in this chapter are not merely the results of question order bias.

I tried to capture the experience of a diverse range of researchers, but most participants of the reviewed studies were from western countries. My Chinese database searches yielded no relevant studies. Likewise, I found no studies on this topic published in Spanish or Portuguese. Around a quarter of visitors to the EQUATOR Network's website have their browser set to a language other than English (I talk about this more in chapter 5), and non native-English speaking researchers may well face additional challenges not covered here. The studies synthesised in this chapter were all conducted in English. This may explain why language barriers did not appear as a theme, despite being identified as a potential issue in quantitative surveys [117] as discussed in the next chapter.

I did not distinguish between different guidelines and expect that the themes I found may apply to reporting guidelines to different degrees. I also expected code frequency to be biased by the questions asked in each study. I therefore decided not to prioritise themes by importance or frequency.

I considered including grey literature, commentaries, and opinion pieces. These may have contributed themes to my analysis but finding these pieces (many of which may have been on private blog posts not indexed by search tools), and the extra work of synthesising primary and secondary order constructs was not feasible. I also considered synthesising quantitative survey data. This data was mostly categorical or ordinal. I decided against it because the surveys differed in content and measure and because numbers would not give me reasons. Instead, in the next chapter I describe how I looked at

the questions themselves and compared them with the themes identified in this chapter.

3.4.2 Conclusions

Researchers encounter many influences when using reporting guidelines. Overall, I found few reporting guidelines have been evaluated qualitatively and the few that have may suffer from thin description, recall bias, question order bias, and lack diverse sampling. In chapter 10 I describe how I avoided these limitations in my own qualitative study where I used interviews, observation, think-aloud tasks, and writing tasks to explore a diverse group of authors' experiences of a redesigned reporting guideline. Reporting guideline developers should be encouraged, supported, and funded to evaluate their resources using in-depth qualitative methods like these. Most of the studies included in this review were surveys, and during my search I found other (quantitative) surveys. In the next chapter, I explain why and how I explored the questions contained in these surveys.

3.5 Reflections on this chapter

As with all qualitative research it is important for me to reflect on how my stance may influence my data collection and interpretation. Because I'm studying reporting guidelines I felt it was particularly important to keep a record of my own experience using them for the first time. This was the first chapter I wrote. Before Charlotte Albury joined my supervisory team none of us had ever done a qualitative synthesis before. We did not even realise our planned study *was* a qualitative evidence synthesis and consequently did not know which reporting guideline I should use to write my protocol. One supervisor suggested the Non-Intervention, Reproducible, and Open Systematic Reviews framework [123] before I found the Journal Article Reporting Standards for Qualitative meta-analyses (JARS-Qual)[124]. Developed by the American Psychological Association, it felt like a good fit as my phenomena (experiences of using a reporting guideline) fell within the realm of social sciences. However, when I tried to publish my protocol the editor (who came from the medical meta-research world) insisted I use ENTREQ [87] which was developed for reporting syntheses of *health* research, even though

my phenomena was not health-related. Although I disliked being strong-armed to use a reporting guideline I felt was less relevant to my work, there was sufficient overlap to switch without much complaint. Neither ENTREQ nor JARS-Qual have ready-to-use checklists, so when I wanted to complete a reporting checklist I had to make my own. I found parts of ENTREQ confusing. The article states “the ENTREQ statement consists of 21 items grouped into five main domains” and continues to discuss each domain. But the checklist (presented as a table) does not mention domains. This makes it difficult to work out where each item is discussed in the guidance text. If I want to read more about item 15 (“software”) should I look under Methods and Methodology, Literature Search and Selection, Appraisal, or Synthesis of Findings? This was my first attempt at using a reporting guideline and already I was encountering issues with selecting the right one, understanding it, and applying it to my writing. Some of the influences my synthesis had identified were beginning to ring true, and going forward I attempted to avoid giving precedence to barriers I experienced myself.

Chapter 4

What influences researchers when using reporting guidelines? Part 2: Describing the questions asked in quantitative surveys

4.1 Introduction

In the previous chapter I describe my systematic search and thematic synthesis of qualitative research exploring authors' experiences of using reporting guidelines. I identified influences that may affect whether an author adheres to reporting guidelines. Many of the studies I included used mixed methods surveys and my search also identified purely quantitative survey studies. Despite not including these quantitative data in my qualitative synthesis, I decided the quantitative *questions* were important to investigate for three reasons.

My primary reason was to continue pursuing the objective of my previous chapter: to identify influences that may affect whether authors adhere to reporting guidelines. As many authors of these studies were themselves users or developers of reporting guidelines the questions may reflect real influences they have experienced, witnessed, or are

trying to avoid or achieve.

Secondly, I wanted to know how many reporting guidelines had undergone any user testing, qualitative or quantitative. My qualitative synthesis found few reporting guidelines had undergone qualitative evaluation. Quantitative surveys might be easier for guideline developers to run (as most come from quantitative disciplines), and so I wanted to know which guidelines had been evaluated quantitatively and how this compared with those from my previous chapter.

Thirdly, I expected this analysis would provide additional context for my qualitative synthesis. In mixed-method surveys, quantitative questions may bias responses to subsequent qualitative questions and, therefore, the findings of my thematic synthesis. For instance, qualitative questions like “Anything else?” or “Please elaborate” may lead respondents to neglect or repeat topics covered by the previous quantitative questions.

My objectives, therefore, were to:

1. Describe the landscape of quantitative survey studies exploring authors’ experiences of using reporting guidelines.
2. Identify additional possible influences that were absent from my qualitative evidence synthesis.
3. Compare themes arising from the quantitative questions with those from my qualitative synthesis.

4.2 Methods

My qualitative synthesis included a systematic search that sought to capture all survey studies investigating reporting guidelines (see chapter 3 for full search details). I found 22 studies using quantitative survey questions, 14 of which also included one or more qualitative questions (see Table 4.1). Although Chinese databases did not yield qualitative studies to include in my previous chapter, I did find two Chinese quantitative surveys [125,126]. Yuting Duan from the Chinese EQUATOR Centre translated these two

studies into English. I imported files into NVivo, including the full surveys where available, labelled all questions with descriptive codes, creating new codes when necessary, and then inductively grouped related codes into broad categories.

Describing the questions asked in quantitative surveys

TABLE 4.1: Studies that collected quantitative data to explore researcher's experiences of reporting guidelines

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Brouwers et al. 2016 [38]	The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines	AGREE Reporting Checklist	Not reported	15	Quantitative
Burford, Welch, Waters et al., 2013 [111]	Testing the PRISMA-Equity 2012 reporting guideline: the perspectives of systematic review authors	PRISMA-Equity Checklist items embedded into survey	Not reported	151	Mixed methods
Davies, Donnelly, Goodman, Ogrinc 2016 [112]	Findings from a novel approach to publication guideline revision: user road testing of a draft version of SQUIRE 2.0	SQUIRE Guidelines, which are presented as a checklist	Not reported but invited participants were from USA, UK Lebanon, Sweden.	44	Mixed methods

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Dewey, Levine, Bossuyt et al., 2019 [106]	Impact and perceived value of journal reporting guidelines among Radiology authors and reviewers	CONSORT, STROBE, PRISMA, STARD checklists	USA, Canada, China, South Korea, Japan, Germany, France , Italy, UK, Other European countries, Middle East, Latin America and 'Other'.	831	Mixed methods
Eysenbach, 2013 [114]	CONSORT-EHEALTH: Implementation of a Checklist for Authors and editors to improve reporting of web-based and mobile randomized controlled trials	CONSORT-Ehealth checklist	Not reported	61	Mixed methods

Describing the questions asked in quantitative surveys

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Fang, Xi, Liu et al. 2016 [125]	A survey on awareness of the ARRIVE Guideline and GSPC in researchers field in animal experiments field in Lanzhou City	ARRIVE Guidelines and Gold Standard Publication Checklist	China	287	Quantitative
Fuller, Pearson, Peters, Anderson, 2015 [109]	What affects authors' and editors' use of reporting guidelines? Findings from an online survey and qualitative interviews	TREND and reporting guidelines in general	Predominantly North America	56	Mixed methods
Giray et al. 2020 [127]	Assessment of the knowledge and awareness of a sample of young researcher physicians on reporting guidelines and the EQUATOR network: A single center cross-sectional study	CONSORT, PRISMA, CARE, GRASS, STARD, STROBE, ARRIVE, SAMPL guidelines	Turkey	100	Quantitative

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Guo, Qi, Yang et al., 2018 [128]	Recognition status of quality assessment and standards for reporting randomized controlled trials of traditional Chinese medicine researchers	CONSORT Statement, STRICTA guidelines and CONSORT extension for Traditional Chinese Medicine	China	180	Quantitative
Korevaar, Cohen, Reitsma, et al, 2016 [110]	Updating standards for reporting diagnostic accuracy: the development of STARD 2015	STARD checklist	Not reported for quantitative survey	12	Mixed methods
Ma et al. 2017 [126]	Survey of basic medical researchers on the awareness of animal experimental designs and reporting standards in China	ARRIVE guidelines and Gold Standard Publication Checklist	China	266	Quantitative

Describing the questions asked in quantitative surveys

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Macleod, Collings, Graf et al. 2021 [107]	The MDAR (Materials Design Analysis Reporting) Framework for transparent reporting in the life sciences	MDAR checklist	USA, China, Japan, Germany, Other EU, 'Other'	211	Mixed methods
McDonough et al. 2011 [129]	Familiarity of non-industry authors with good publication practice and clinical data reporting guidelines	CONSORT guidelines	USA, UK, Canada, South Africa, Israel, China	23	Quantitative

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Öncel et al. 2018 [130]	Knowledge and awareness of optimal use of reporting guidelines in paediatricians: A cross-sectional study	CONSORT guidelines, STROBE, PRISMA, CARE, SRQR, STARD, SQUIRE, CHEERS, SPIRIT, ARRIVE, TREND, STREGA, the Conference on Guideline Standardization (COGS), Outbreak Reports and Intervention Studies Of Nosocomial infection (ORION)	Turkey	244	Quantitative
Page, McKenzie, Bossuyt et al. 2021 [115]	Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement	PRISMA statement	Not reported	110	Mixed methods

Describing the questions asked in quantitative surveys

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Prady & MacPherson 2007 [116]	Assessing the Utility of the Standards for Reporting Trials of Acupuncture (STRICTA): A survey of authors	STRICTA	Not reported	28	Mixed methods
Prager, Gannon, Bowdridge et al. 2021 [117]	Barriers to reporting guideline adherence in point-of care ultrasound research: a cross-sectional survey of authors and journal editors	STARD	Not reported	18	Mixed methods
Phillips et al 2015 [131]	Pilot testing of the Guideline for Reporting of Evidence-Based Practice Educational Interventions and Teaching (GREET)	GREET checklist and E&E	Not reported	31	Quantitative

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Rader, Mann, Stransfield et al., 2014 [118]	Methods for documenting systematic review searches: a discussion of common issues	PRISMA statement	Not reported	263	Mixed methods
Sharp, Glonti, Hren, 2020 [132]	Using the STROBE statement: survey findings emphasized the role of journals in enforcing reporting guidelines	STROBE statement	The full survey was answered by participants in Africa, Asia, Europe, North and South America, Middle East, and Pacific Region. It is unclear who answered the free text question.	1015	Mixed methods
Struthers, Harwood, de Beyer et al., 2021 [54]	GoodReports: developing a website to help health researchers find and use reporting guidelines	Reporting guidelines in general	Not reported	274	Mixed methods

Describing the questions asked in quantitative surveys

Citation	Title	Guidelines studied	Sample geographics	Sample size	Quantitative or mixed methods
Tam, Tang, Woo, Goh 2019 [120]	Perception of the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) statement of authors publishing reviews in nursing journals: a cross-sectional online survey	PRISMA statement	Not reported	230	Mixed methods

4.3 Results

4.3.1 What reporting guidelines were studied?

Between the 22 studies 25 reporting guidelines were mentioned, most frequently PRISMA (n=6), STARD (n=6), CONSORT (n=6) and ARRIVE (n=5) (see my [List of Abbreviations](#) for the full titles of each reporting guideline). Thus, only a small proportion of the reporting guidelines indexed in EQUATOR's database [133] have been evaluated with quantitative questions. Fourteen studies focussed on a single guideline (n=14), with others asking questions about multiple guidelines (e.g., "which reporting guidelines [participants] had known" [127]) or guidelines in general (e.g., "whether they had used reporting guidelines in their publications" [130]). Most studies included participants from the USA, Europe, and Canada and only a few studies were conducted elsewhere (e.g., China and Turkey) (see Table 4.1).

In comparison, my thematic synthesis (chapter 3) identified 18 studies that collected qualitative data. These studies covered only 12 reporting guidelines and were all conducted in western countries, hence were slightly less diverse than the quantitative survey studies.

4.3.2 The focus of quantitative questions

Survey studies asked participants:

- whether they were aware or familiar with certain reporting guidelines,
- how often they used them and what for,
- whether reporting guidelines had influenced their behaviour,
- whether guidance was usable and useful,
- their opinions on guidance content,
- their reasons for using a reporting guideline,
- their opinions on reporting quality in the literature,
- whether reporting guidelines were easy to find and access,
- whose role it was to check for compliance,
- whether the aim of the guidance was clear,

- opinions on things explicitly named as a barrier including the length of the guidance, the language it is written in, and the time needed to use it, and
- opinions on things explicitly named as a facilitator or motivator including endorsements, evidence, explanatory information, training, the behaviour of peers, and the development process of the guidance.

4.3.3 Comparing the focus of quantitative questions with themes derived from qualitative data

The quantitative questions included some novel influences not contained in the qualitative data (shown in bold in Table 4.2), such as training as a possible facilitator [109], whether authors had heard of the EQUATOR Network [109,127], and whether transparency in guideline development is important [109]. One study asked whether language may be a barrier to using reporting guidelines for some [117]. Both quantitative questions and the qualitative data mentioned journals enforcing reporting guidelines, but only quantitative questions asked whether funders and employers should also enforce them.

TABLE 4.2: Codes describing the focus of questions asked and their code categories. Items in bold did not appear in the qualitative data.

CODE	CATEGORY
Participant's experience [106,109–112,117,125–130,132]	Demographics
Participant's speciality [109–112,117,118,120,125–127,132]	
Participant's age [117,120,125–128,130,132]	
Participant's gender [117,120,125,127,128,130,132]	
Participant's geography [106,109,129,132]	
Participant's stage of current research project [111]	
Awareness of a particular guideline [109,111,116,117,120,125–127,129,130,132]	Awareness
Awareness of EQUATOR [109,127]	
How did they first hear about guidelines or EQUATOR? [109,130,132]	
When did they first learn about a guideline? [109]	

CODE	CATEGORY
How frequently do they use guidelines? [106,109,111,112,117,120,127,130,132]	Usage
When should guidelines be used? [106,109,117,127,130,132]	
Would they use a guideline, hypothetically [38,111,132]	
Did the guidance impact subsequent behaviour? [38,54,106,111,112,114,116]	Impact on behaviour
Is the guidance usable? [117,131,132]	Usability
Is the guidance easy to understand? [54,110,112,117,128,132]	
Is the guidance useful? [38,54,106,107,109,126,132]	Usefulness
Is the guidance important? [109,112,114,120,126,128]	Importance
Are time and length barriers? [54,109,114,117,132]	Barriers
Is language of guidance a barrier? [117]	
Are guidelines lacking for study type? [109]	
Is the layout OK? [38,110,132]	Opinions on content
Should the content be modified? [38,110,115]	
Is the guidance relevant? [54]	
Are guidelines prescriptive? [109]	
Will using a guideline benefit the manuscript? [38,111,117,132]	Reasons for using a guideline
Productivity benefits of using guidelines [132]	
Using guidelines because of journal requirements [109,132]	
Using guidelines because of funder requirements [109]	
Using guidelines because of employment requirements [109]	
Using guidelines because of other researchers expecting it [132]	
Opinions on reporting quality of the literature [109,117,125,126]	Opinions on reporting quality
Are guidelines easy to find and access [109,117,126]	Accessibility
Who should complete the checklist? [109,117]	Roles

CODE	CATEGORY
Are endorsements a facilitator? [109]	Facilitators and motivators
Is evidence of increased chance of publication a facilitator? [109]	
Is evidence of improved reporting quality a facilitator? [109]	
Is explanatory information a facilitator? [109]	
Is training a facilitator? [109]	
Is the behaviour of peers a facilitator or motivator? [109]	
Is the evidence base underlying a reporting guideline a motivator? [109]	
Is transparency in guideline development a motivator? [109]	
Is the aim of the guidance clear? [110]	Aim of guidance

TABLE 4.3: Codes and descriptive themes identified from a qualitative evidence synthesis. Items in bold did not appear in the quantitative questions. Items in italic offer possible explanations to some quantitative findings.

CODES	DESCRIPTIVE THEMES
<i>What does this term mean? [112; 108; 110; 115; 54]</i>	What does this
<i>What does this item mean? [54,108,110,112,115,116]</i>	mean?
<i>How are these items different? [108,114–116]</i>	
<i>Have I understood this as intended? [108,112]</i>	
<i>Examples help me understand items [29,113,115]</i>	
<i>Why is this item important? [108,110,115,120]</i>	Why is this item
<i>Who is this item important to? [57,108,115]</i>	important?
<i>Have I understood the guideline's scope as intended? [54,115]</i>	Does this apply to
<i>Does this item apply to me? [54,108,114–116]</i>	me?
<i>Is this item optional? [108,116]</i>	
What are reporting guidelines? [57,117]	I don't understand
How should I use a reporting guideline? [109]	what reporting
	guidelines are

CODES	DESCRIPTIVE
	THEMES
I find guidelines useful in general [54,106]	Guidelines benefit
Guidelines make me feel confident [57]	me
Guidelines help me develop as a researcher [57,111]	
Guidelines may help me improve my manuscript [57,106,108,111,114]	
I believe guidelines may help me publish more easily [119]	
I may use guidelines because journals and editors tell me to [57,109,111,119]	I use guidelines because of other people
I may use guidelines because other researchers expect it [109,119]	
Standardized reporting benefits the community [26,57,119]	Guidelines benefit others
Immediate benefits are more important than hypothetical ones [57,119]	Some benefits are more important
Personal benefits are more important than benefits to others [119]	than others
I use reporting guidelines for planning research [57,108]	Researchers use
I use reporting guidelines for designing research [57,106,116,117]	reporting guidelines for different tasks
I use reporting guidelines for writing [57,106,108,116]	
I use reporting guidelines for checking my own or other people's writing [57,117]	
I use reporting guidelines to appraise the quality of other people's reporting [110]	
I use reporting guidelines for peer reviewing [57]	

CODES	DESCRIPTIVE	THEMES
I want items presented in the order in which I must do them [26,29,113]		I want guidance presented in
I want design or methods advice [57,108,115]		formats that are
I want templates for writing [106]		better suited to the
I want checklists that are easy to fill in [54,107]		task I am doing
I want checklists embedded into journal submission workflows [106]		
I want items embedded into data collection tools [111]		
Guidelines take time to read, understand and apply [109,111,119]		Guidelines take time
Some items require extra work which takes time and effort [108,112,118]		
I want an indication of which items to prioritize [108,116]		
<i>Perceived complexity [106–108,119]</i>		
<i>Long guidelines are off-putting [54,57,111,114]</i>		
Itemization helps me navigate guidance [115]		Itemization may
Itemization summarizes the guidance [106]		decrease costs
Itemization makes guidance appear longer [115]		Itemization may
Itemization blocks the bigger picture [108]		increase perceived
		costs
Following reporting guidance can result in long, bloated articles [108,111,114,116]		I think guidelines make my
Long, bloated articles may exceed journal word limits [107,109,114,116]		manuscripts long and bloated
I want options for where to report this item		
[57,108,109,112,114,115]		

CODES	DESCRIPTIVE THEMES
<i>The benefits of using a reporting guideline may not outweigh the costs [57,109,114]</i>	The benefits of using a reporting guideline may not outweigh the costs
Guidelines are more valuable when used early [54,57,106,108]	The balance of benefits vs costs may be more favourable when guidelines are used early
I would clarify this item [115,116]	I think the guidance could be improved
I would move this item [108,112]	
I would split this item into two [108,113,115]	
I would add or remove items from this guideline [108,110,115,116]	
I would add or remove requirements from this item [29,57,115,116,120]	
Guidelines can become out of date [108]	Guidelines need to be kept updated
Guidelines need to be updated [115]	
I cannot report this because I didn't do it [116]	I feel unable to report this
I cannot report this because of intellectual property issues [114]	
I cannot report this because it clashes with journal guidelines [115]	
I cannot report this because data was missing from my primary studies [111]	
Editors, reviewers or co-authors asked me to remove this item [116,118]	

CODES	DESCRIPTIVE THEMES
I feel uncertain because I don't know how to say that I didn't do it [115]	I feel nervous or uncertain if I am unable to report an item
I feel worried that I will be judged for transparently reporting something I didn't do [57,115]	unable to report an item
I may not know that reporting guidelines exist [54,106,109,110,119]	I can only use what I know about and have
I may not be able to easily access guidance [54,119]	
Reporting guidelines may be less valuable to experienced researchers [57,106,114]	Reporting guidelines are more valuable to inexperienced researchers
Experienced researchers feel that they already know how to report [57,106,108]	
Experienced researchers find guidance patronizing and feel untrusted [107,109,114,115]	
Reporting guidelines can be hard to use at first but get easier with experience [108,109,119]	Reporting guidelines can be hard to use at first but get easier with experience
I want design or methodological advice [57,107,115]	I want or need design advice
I don't know how to do this item [108,115,116]	
Guidelines are procedural straightjackets [57]	I think this guidance prescribes how research should be designed
This guideline is too prescriptive [57,115,120]	
<i>The guideline's applicability criteria are not clear [54,106,110]</i>	A guideline's scope can be unclear
<i>This guideline isn't a perfect fit for me [54]</i>	A guideline can be too narrow
This guideline doesn't generalise [57,106,107,115,120]	
This guideline is too prescriptive [57,115,120]	

CODES	THEMES
I don't want to see optional items that only apply to other types of study [54,116]	A guideline's scope can be too broad
I need to adhere to journal guidelines or other research guidelines [106,109,115,116]	Authors often need to adhere to multiple sets of guidance
I might need to use multiple reporting guidelines [57]	
I want reporting guidelines to be linked or embedded [110,115]	I want guidelines to harmonize
I want reporting guidelines to use similar structure [115]	
I want reporting guidelines to use similar terms [115]	
I don't like checklists [54,57,106,114]	I experience reporting guidelines primarily as, or through, checklists
I may use the checklist instead of the full guidance [29]	
I may use the checklist before I read the full guidance [29]	

Most ideas captured in the quantitative questions also appeared in the qualitative data. This may indicate that the quantitative questions asked were pertinent, or perhaps that they influenced participants' responses to subsequent, qualitative questions.

Overall, although the quantitative questions contained some novel themes, the qualitative data contained many more ideas that were not addressed by the quantitative questions (see bold items in Table 4.3). These included what authors understand reporting guidelines to be, the pros and cons of itemization, ideas of how guidance could be improved, negative feelings when an item cannot be reported as desired, the pros and cons of including design advice in reporting guidance, whether optional items were understood as being optional, and frustration when the scope of a reporting guideline is too broad, narrow, or unclear.

The qualitative data sometimes provided context to or explanation for quantitative answers (see italicised items in Table 4.3). For example, many of the quantitative surveys

asked participants whether they could understand the guidance. However, a quantitative answer to this question does not reveal *what* the participant understands, *how* they understand it, or whether they understand it *as intended*. The qualitative data contained reports of people failing to understand the wording of an item, how to report that item in practice, whether an item applies to them, whether a reporting guideline applies to them, what the intended scope of a reporting guideline is, or even what a reporting guideline is at all. One study found that although authors reported understanding an item, their writing showed that they had interpreted it differently to how the reporting guideline developers had intended [112].

4.4 Discussion

The aim of this chapter was to reveal which reporting guidelines had been evaluated using quantitative survey questions, what influences those questions explored, and to compare those influences with the results of my previous chapter.

Few reporting guidelines have been evaluated using quantitative or mixed surveys. Given that even fewer have been tested qualitatively, this means that hardly any of the hundreds of reporting guidelines in the EQUATOR Network's database have undergone meaningful user testing. To encourage and support future guideline developers, advice on user testing could be included in an update of guidance for guideline developers, the current version of which contains little advice on how to evaluate reporting guidelines [122].

The quantitative survey questions offered few new influences and covered few of the themes I identified from qualitative studies. Quantitative surveys often asked about awareness, usage, usability, usefulness, importance, barriers, facilitators, content, and whether reporting guidelines had led to a change in behaviour. However, many of the themes I identified in my qualitative synthesis did not appear in the quantitative surveys. This suggests that the quantitative questions did not bias or limit the qualitative

data, and also suggests that quantitative surveys may often miss themes. Often the qualitative data provided context and explanation for the quantitative data. Because quantitative surveys can miss important findings or fail to explain them, developers seeking actionable feedback should collect qualitative data when assessing how researchers understand or feel about reporting guidelines, or what could be done to improve the guidance.

4.4.1 Limitations of included studies and advice for future research

All included survey studies were subject to recall bias as participants were describing past behaviour or opinions. Future studies could consider methods that allow researchers to document experiences in real time, like observation or think aloud tasks.

The included studies' participants lacked diversity. Studies should ensure participants represent expected users in terms of academic writing experience, discipline, profession, experience (or naivety) with reporting guidelines, and language, or even focus on differential experiences of specific target groups. For example, the EQUATOR Network website gets similar levels of traffic from Asia and Europe, yet very little research into usability or barriers of reporting guidelines has included authors from Asian countries (see chapter 5). The website also sees many new visitors who abandon the site quickly, without accessing any reporting guidance. These visitors may be authors who are naïve to reporting guidelines and decide not to use one. Understanding why people choose not to use a guideline is equally important as understanding the experiences of those that do. Many of the included studies used snowball sampling via social networks connected to the researchers, and hence may have attracted participants who knew about reporting guidelines and already held opinions on them. Many studies required authors to read the guidance as part of the study itself, thereby forcing participants to engage with it. Consequently these studies do not capture perspectives of less-engaged authors, or explain why authors choose not to use a guideline.

Some included surveys contained leading questions. For example, the Likert rated statements “The STARD 2015 guidelines are easy to follow” [117] and “The time required to adhere to the STARD 2015 guidelines is a barrier to using the guidelines” [117] are both

subject to acquiescence bias; the tendency for participants to agree with research statements [134]. Future studies should consider using neutral questions.

Studies used lots of different words to describe reporting guidelines, including *guidelines*, *standards*, *requirements*, *checklist*, *explanation and elaboration*, or just an acronym, e.g., CONSORT. This became a problem in studies where participants were not supplied with guidance documents as part of the study, as it was not always clear which document a participant was considering. For instance, asking participants whether PRISMA is easy to understand will not tell you whether they are talking about the PRISMA checklist, statement, or explanation and elaboration document. Future studies should be specific when asking questions and reporting results.

4.4.2 Conclusions

Very few reporting guidelines have been evaluated using either quantitative or qualitative methods. Reviewing the content of quantitative surveys revealed few novel influences which were absent from the qualitative data synthesised in chapter 3. However, the qualitative data contained many more themes than the quantitative questions. This suggests that the results of my qualitative evidence synthesis were not limited by the content of quantitative survey questions within the mixed methods surveys. Because the qualitative data revealed more themes and provided explanation and context to findings, reporting guideline developers who want to make sure their resources are easy to use should consider using qualitative methods, which may produce richer, actionable insights.

Two studies asked participants whether they had heard of the EQUATOR Network, noting that it is a “valuable resource for users and potential users of reporting guidelines” that 44% (19/43) of editors [109] and 38% of authors (38/100) [127] are aware of (published in 2015 and 2020 respectively). Although these studies asked participants whether they were familiar with EQUATOR, authors’ experiences of using EQUATOR’s website has never been explored. In the next chapter I describe EQUATOR’s website and key characteristics of its web traffic, before discussing how well it is helping authors find reporting guidelines.

4.5 Reflections on reporting this chapter

This chapter could have been combined with the one before. They came from the same starting idea, and their findings are intertwined. I separated them because I felt like my first chapter was already getting too long, and I wanted to keep its structure simple, familiar, and in line with ENTREQ. Would I have made this decision if ENTREQ did not exist? Possibly not. Separating the chapters draws a distinction between their objectives and methods, but in drawing this line I am diverging from the true origin story of this chapter. The two were not designed as separate studies, so in presenting them as such am I twisting the truth so it falls in line with a reporting guideline? This might not matter for these exploratory studies, but it made me question whether other researchers may feel pressure to package a messy research journey into a neat-and-tidy reporting guideline format.

Chapter 5

Describing who visits EQUATOR’s website for disseminating reporting guidelines, and how visitors use it

5.1 Introduction

When trying to book their COVID vaccine booster, a relative gave up because they could not understand the website booking process. When interventions rely on a website their success relies, in part, on whether people can easily understand and use that website. The same is true for reporting guidelines which are disseminated through a number of webpages like journal author guidelines, guideline-specific websites, and the focus of this chapter; the EQUATOR Network’s website.

EQUATOR’s organisational aims include maintaining a “comprehensive collection of online resources providing up-to-date information, tools and other materials related to health research reporting” and to assist in the “dissemination and implementation of robust reporting guidelines”. The website is the face of this “comprehensive collection” and is key to dissemination. The website’s annual traffic has increased from 100,000 users to almost 1 million over the last 10 years, attracting visitors from all around the

world. For context, the International Association of Scientific, Technical, and Medical Publishers estimated that 11 million people worked in research and development in 2018 [135]. Whilst these numbers will have grown since, they suggest that EQUATOR’s website traffic numbers are within a single order of magnitude of its potential audience size. This is impressive, especially considering how the website was built on a shoestring budget and without any in-house expertise in software, design, or user experience.

Despite the website’s success and rising importance within the medical publishing ecosystem, there has never been a formal investigation into how successfully the website contributes towards EQUATOR’s aim of disseminating reporting guidelines. In chapter 4 I found two studies exploring authors’ awareness of EQUATOR and its website [109,127], but no studies asked authors whether they found it easy to use.

I wanted to explore how well EQUATOR’s website was serving the organisation’s aim of disseminating reporting guidelines, because I expected my findings to extend or add context to the lists of influences I identified in chapters 3 and 4. My objectives were to answer the following questions:

1. How many people visit the website each year?
2. What countries are visitors from, and how does this compare with the global distribution of researchers?
3. How often do visitors come back?
4. How do visitors get to the website?
5. Do visitors engage with the website?
6. What content do visitors view?
7. What reporting guideline database records do visitors view?
8. How many visitors continue to access guidance on a third party site?
9. How many visitors access publications vs. checklists?
10. Do authors manage to achieve or find what they wanted?

5.2 Methods

Assessing how effectively a website contributes towards an organisation's objectives is generally referred to as a website service evaluation, although there is no standard definition nor approach [136]. Some evaluations use in-depth qualitative methods like interviews, card-sorting, and observed browsing or task completion. In chapter 10 I describe how I use some of these techniques but at this stage I wanted to perform a broad descriptive exploration of user activity examining the types of users, where they come from, and how they interacted with the website, and so I decided to use a suite of commonly used software tools: Google Analytics, Google Tag Manager, and a survey tool called Popup Smart. These are described in Table 5.1 along with some of the core vocabulary they use.

TABLE 5.1: Definitions of terms used in this chapter

Term	Definition
Reporting guideline database record	EQUATOR maintains a database of reporting guidelines. Each database entry is viewable as a webpage that displays the name of the guideline, bibliographic data, meta data, and links to the associated publications and files. The database page does not include the guidance itself.
Google Analytics	A web analytics service that tracks and reports on how many people visit a website and how they interact with it.
Google Tag Manager	A system that allows you to “inject” code into a website to facilitate custom analytics.
Google Search Console	A service that allows website owners to check how Google’s search engine indexes a website and to optimize its visibility.
Popup Smart	A service for adding surveys to websites and to control when they appear.
User	A person visiting the website (or, more specifically, a unique browser cookie).
Session	A single visit where one or more pages are viewed.

Term	Definition
Bounce	A session where a user visits a single page and performs no actions. In our setup, we considered outbound link clicks to guideline articles to be actions (because we want authors to click through to guidance). So if a user visited a reporting guideline database record and then immediately clicked an outbound link to a guidance publication, this would <i>not</i> be counted as a bounce.
Bounce Rate	The ratio of bounce sessions to total sessions within a given time duration.

Google Analytics [137] is a web analytics service that helps website owners track and understand users' activity. Used by 85% of websites globally, it is the most popular web analytics service by far [138]. EQUATOR has used Google Analytics to collect data since creating their website. Mostly they use the data to report high level impact metrics to funders (such as number of visitors) but they have never used it to evaluate their website in depth.

There are a huge number of metrics you can consider when evaluating a website; Google Analytics collects over 50 by default [139,140], and Google Tag Manager [141] allows you to collect additional custom metrics. My first seven research questions could be explored using Google Analytics default data collection settings. Questions eight and nine could not: Google Analytics recorded which reporting guideline database records visitors looked at, but not how many people went on to view guidance on third party websites, or whether visitors went on to view checklists or full guidance. Therefore I used Google Tag Manager to create two custom metrics: one to count when visitors downloaded a reporting checklist file, and another to count when visitors accessed a third party website.

The tenth question required exploring whether website users were able to access the *right* resource. Although Google Analytics could tell me *what* pages visitors access, it

could not tell me what they *needed*, or *why*. To explore this, I decided to use PopupSmart [142] to build an exit survey (an online questionnaire that pops up when a user appears to be leaving the website). I did not want to annoy users or block them from using the website, and I wanted to keep the survey short in the hope of maximizing response rate, and so I decided to limit the survey to a single question. I decided to use an open ended question in the hope of receiving richer responses.

When deciding what this question should be I consulted reporting guideline experts from the UK EQUATOR Centre who are responsible for maintaining the website. We decided against asking “What were you looking for today?” as that would not tell us whether users had found what they needed. We considered “How easy was it to find what you were looking for?” but decided it was too closed, the word “easy” was too subjective, and it assumed too much about the users’ intent. Instead, given that the questionnaire would only appear as users prepared to leave the site, we decided to ask “Why are you leaving?”. We intended to add narrower follow up questions in the future, depending on the responses to this initial, broad question. We hoped the open question would cater to all users, whether or not they found what they needed, or whether they were leaving for any other reason. For example, one EQUATOR member joked “We might get answers like ‘because it was not the right website I was looking for, I was looking for geography websites’...! haha” (spoiler alert: this was not far off the mark!).

5.3 Results

5.3.1 How many people visit the website?

830 134 users visited equator-network.org in 2021 (see Table 5.2), 150,000 more than the previous year, thus continuing growth seen since attracting 20,000 visitors in 2008.

5.3.2 What countries are visitors from, and how does this compare with the global distribution of researchers?

See table Table 5.2 for users by country. A third of users were from the United States or United Kingdom. The geographic distribution of users didn’t always align with global publishing trends. For example, Brazil accounted for the same proportion of users as

the UK (7%) despite producing far fewer citable, medical documents [143]. Conversely, China produces twice as many citable documents as the UK but accounts for only 5% of users. Two fifths of users had their browsers set to a language other than English (Table 4), most frequently Portuguese, Chinese, and Spanish.

5.3.3 How often do visitors come back?

Google Analytics classified almost all users as new (see Table 5.2), which means they had not visited the site within the previous 2 years (the default expiry limit for Google Analytics cookies). Most users only visited the site once within 2021.

TABLE 5.2: User demographics equator-network.org for the year 2021.

*There are 105 English locales, including the United States, India, Ireland etc.

	Number	% total users
Total users	830,134	
Users who had not visited before	823,087	99%
Users by country (top 10)		
United States	195,217	24%
United Kingdom	61,292	7%
Brazil	55,894	7%
China	41,628	5%
India	41,196	5%
Australia	30,647	4%
Canada	29,620	4%
Netherlands	25,426	3%
Germany	24,561	3%
Spain	23,020	3%
Users by browser language setting (top 10)		
English (all varieties*)	498,286	60%
English (United States)	371,331	45%
English (Great Britain)	86,580	10%
Portuguese (Brazil)	50,172	6%

	Number	% total users
Chinese	42,951	5%
Spanish	36,204	4%
French	14,708	2%
Dutch	14,331	2%
Spanish (Latin America and Caribbean)	14,259	2%
Japanese	13,381	2%
German	12,460	2%

Users by the number of sessions they made

within 2021

1 or more sessions	818,512	99%
2 or more sessions	176,570	21%
> 3 sessions	72,886	9%

5.3.4 How do visitors get to the website?

The source of half of the traffic was labelled as “unknown”, meaning that Google Analytics had no way of knowing where that traffic came from (see Table 5.3). Although this could be because authors type the URL into their browser, or because they are clicking hyperlinks within offline documents, I believe it is more likely a result of journals incorrectly linking to EQUATOR’s website (see discussion). A third of users arrive at the site via a search engine. An eighth of traffic was explicitly labelled as referrals from other websites, most commonly Wiley and Elsevier journals and Manuscript Central.

A third of sessions begin with users arriving on the home page (see Table 5.4). Most other sessions begin with authors arriving directly on a reporting guideline database record page.

5.3.5 Do visitors engage with the website?

Over half of sessions (53%) ended without the user interacting with the site at all (see Table 5.3). Google Analytics calls this behaviour bouncing. Bounce behaviour was similar regardless of which webpage user arrive on: 45% for the home page, >50% for most

reporting guideline database record pages. Two thirds of sessions lasted less than 10 seconds, suggesting little interaction.

TABLE 5.3: Session information for EQUATOR-Network.org for the year 2021. A session is defined as group of user interactions that take place within a given time frame. If a user visited the site twice within a year, then that would appear as two sessions. A session ends after 30 minutes of inactivity.

	Number	Percent
Number of sessions	1,209,420	
Mean number of sessions per user	1.5	
Number of sessions originating from...		
a referral from another website	142,158	12%
a search engine	417,671	35%
unknown sources	590,659	49%
Number of sessions lasting...		
< 10s	760,967	63%
< 1 min	908,516	75%
Number of pages viewed within a session		
0	1,451	0%
1	676,670	56%
2	223,864	19%
3	81,878	7%
4 or more	225,557	19%
Number of sessions ending without interaction	648,370	54%
Number of sessions including views of...		
reporting guidelines	605,570	50%
landing page	411,429	34%
library	83,238	7%
toolkits	15,574	1%
study design info	15,268	1%
(All other categories of content were viewed in less than 1% of sessions)		

TABLE 5.4: Web pages where visitors began their session and their bounce rates between 1st of January and 1st of July 2022. Bounces are sessions where a visitor leaves without interacting with the website at all.

Landing Page	Sessions starting on page	Percentage of total sessions	Bounce rate
All pages	602,921	100%	53%
Home page	218,260	36%	45%
STROBE	68,274	11%	56%
PRISMA	29,398	5%	60%
Reporting guidelines search page	27,122	4%	40%
CONSORT	24,615	4%	55%
COREQ	19,962	3%	73%
STARD	19,316	3%	57%
SRQR	12,039	2%	65%
CARE	10,759	2%	56%
TRIPOD	10,342	2%	58%

5.3.6 What content do visitors view?

Reporting guideline database record pages were viewed in half of all sessions (see Table 5.3). The home page was viewed in a third of sessions. The rest of the website is organised within the menu categories Library, Toolkits, Courses and Events, News, Blog, Librarian Network, and Content. Hardly any of these categories received meaningful traffic. Only 7% of sessions included a view of any page listed within the Library category. All other content categories were viewed in 1% of sessions or fewer.

5.3.7 What guideline database records do visitors view?

There are over 500 reporting guidelines indexed in EQUATOR’s database, but few were viewed regularly. STROBE was the most viewed reporting guideline, viewed in 110k out of 600k sessions in the first 6 months of 2022. These numbers dropped rapidly: SQUIRE, the tenth most viewed reporting guideline record, was viewed ten times less frequently

than STROBE (see Table 5.5). Only 13 guideline database records were viewed in more than 1% of sessions (more than 6,000 sessions in six months), all of which appear in the list of “reporting guidelines for main study types” featured on EQUATOR’s home page and in a side bar on all reporting guideline sub-pages. Only 65 reporting guideline pages were viewed in more than 0.1% of sessions (more than 600 sessions in six months). The remainder were viewed in hardly any sessions, or were only viewed by bots (software that crawls web pages, often for search engine indexing purposes).

TABLE 5.5: How frequently reporting guideline pages were viewed and resources accessed between the 1st of January 2022 – 1st of July 2022. Outbound links were mainly links to publications and files were mainly checklists, but a few guidelines also have flow diagrams which were rarely downloaded. The ten most accessed guidelines are shown.

Reporting Guideline	Sessions where database page was viewed	Users that accessed an outbound link	Users that accessed a checklist
STROBE	110,910	2,502	38,780
PRISMA	50,146	17,240	15,522
CONSORT	45,832	2,465	13,862
COREQ	30,487	11,152	-
STARD	27,301	987	7,803
SRQR	26,132	10,035	-
CARE	21,892	8,761	8,350
TRIPOD	15,240	4,678	3,960
SPIRIT	11,108	3,985	3,244
SQUIRE	11,061	1,022	3,061

5.3.8 How many visitors continue to access guidance on a third party site?

Many users leave the reporting guideline database page without accessing any resources. For example, the STROBE database record was viewed 110,000 times, but the checklist only downloaded 30,000 times (~1 in 3) and the link to the full guidance was only clicked 2,500 times (~1 in 44). Overall, only around 1 in 4 users end up clicking a link that will direct them to a checklist or guideline publication (see Table 5.5).

5.3.9 How many visitors access publications vs. checklists?

Of the ten most viewed reporting guideline database pages, all except COREQ and SRQR offer downloadable checklists. For STROBE, STARD, CONSORT, and SQUIRE, the checklist links are clicked more often than links to the full guidance (15:1, 8:1, 6:1, and 3:1 respectively). For PRISMA, CARE, TRIPOD, SPIRIT, link clicks to each resource are more similar (around 1:1).

5.3.10 Survey data suggest some visitors may not understand what the website is about, and may not find it useful

Our survey received 33 responses in 2 weeks (see Table 5.6), after being viewed by 25,660 people, (a response rate of 0.1%).

A few responses indicated clearly that the user had found what they were looking for; “got what I was looking for!” wrote one user, “found what I needed!” wrote another. Other users had not been so successful, writing “Could not find what I needed”, “I cannot find the guidance that I seek”, “I don’t see what I want”, and “could not find any reporting guidelines for reporting guidelines (specifically, abstracts for reporting guidelines)”. Some visitors voiced frustration with the website (“i did not understand anything”, “The site is very complex”, “Too big a mess”).

Some users hinted at why they were on the website; e.g., “to get reporting guidelines”, “a tool called standard for reporting qualitative research”. One user wrote “word format would be more easy to fulfil”, suggesting they had been looking for a checklist. Others clearly didn’t understand what the EQUATOR website was about. Two authors seemed to be looking for requirements for specific journals: one wrote “format for paper submission to Hindawi”, and another wrote “awful site. I just want to know the requirements in terms of number of words and format for a submission and cannot seem to find this anywhere”. Another author was looking for a “quality of life questionnaire”, another for “scientific research”. More cryptically, one visitor simply wrote “ALGERIA”, another “Germany” (perhaps EQUATOR staff were correct to worry that their website could be mistaken as a geography resource).

Some users voiced frustration with the survey itself (“do something about your annoying pop up!”). Wary of annoying users and given the poor response rate I decided to take the survey down instead of modifying it.

TABLE 5.6: Survey responses to the question “Why are you leaving?”

Survey responses

qualitive

word format would be more easy to fulfil

I've just arrived, actually. As I started to look at the page content, this message was thrust in my face - why is it that websites so often ask immediately for feedback on material the viewer hasn't yet had a chance to explore?

I am Research

You are perfect Thanks

format for paper submission in Hindawi

I cannot find the guidance that I seek. The site is very complex.

I'm not, so you might want to do something about your annoying pop up!

I could not find any reporting guidelines for reporting guidelines (specifically, abstracts for reporting guidelines)

awful site. I just want to know the requirements in terms of number of words and format for a submission and cannot seem to find this anywhere

i did not understand any thing

ALGERIA

to get reporting guidelines

scientific research

I just needed the URL to recommend the site to someone else.

Found what I needed!

because I am in the wrong place

I am not eligible

No desired

I got what I was looking for! Thank you :)

Could not find what I needed

Survey responses

relevant titlke is not found

Diseño y validación de un nuevo registro clínico de enfermería, para la continuidad de los cuidados y seguridad del paciente en hemodiálisis

I just got here?

Following a link in the website to a different page on the guidelines

dont ask that

Great!

germany

Too big a mess

link doesn't work

not relevant

quality of life questionnaire

there is another tool called standard for reporting qualitative research. I want that tool because it is development is better than this.

5.4 Discussion

A core objective of the EQUATOR Network website is to help the global research community to find and access reporting guidelines. My aim in this chapter was to explore how well the website serves that objective. Over 800,000 people visited the website in 2021, thus EQUATOR have been successful in obtaining visitors. However, once they arrive, few visitors engage with the website meaningfully. Half leave within seconds without engaging at all, and even those that stay longer rarely return. The survey responses suggest some users do not understand what the website is about or how to find what they need. Other than the reporting guideline pages, most of the content on EQUATOR's website is rarely accessed. Only 1 in 4 visitors click a link to a checklist or to a guideline's PubMed page. Fewer users would end up reaching the full guideline from this PubMed page, as users would have to access the published article (which may be paywalled), and then find the guidance within the publication itself which can be hidden in a table or supplement. Together, these data suggest the website has huge room

for improvement.

These findings add support to influences I identified in my qualitative evidence synthesis in chapter 3 and, conversely, my synthesis offers possible explanation for some findings in this chapter. My synthesis found that authors may primarily use checklists and rarely (if ever) read the full guidance. The data in this study support this theory. Many authors do appear to access checklists over and above full guidance. This could be their intention, or it could be because EQUATOR places links to checklists at the top of reporting guideline database record pages, making them more prominent than links to full guidance. Either way, guideline developers should be aware that checklists may be the primary way that authors interact with the guidance, and so checklists should link to the full guidance, and should include key information about the guideline like its aim, scope, and how it is intended to be used.

My thematic synthesis suggested some authors may struggle to find the most appropriate guideline for their research. Some of the survey responses echoed this influence. Besides these survey responses, I had no way of knowing which reporting guideline would have been most appropriate for each visitor, but indirect evidence comes from the fact that very few guidelines receive any meaningful traffic. Only 13 guideline database records were viewed in more than 1% of visits. The remaining guidelines form a long tail, with some guidelines receiving no traffic at all. Of course, some guidelines will naturally be accessed more often (most people would expect general guidance for systematic reviews to be accessed more than, say, guidance for eye-tracking studies in dentistry), and guidelines endorsed by journals will attract more visits than un-endorsed guidelines. But the severity of the skew suggests that authors may not be discovering guidance that is most appropriate for them. Alternatively, a cynic may interpret this skew as evidence that some reporting guidelines simply aren't useful.

My evidence synthesis also found that authors may not know what reporting guidelines are, why they are useful, or may choose not to use them if the costs outweigh the benefits. These could explain some of the poor engagement seen in this study. If visitors do not know what reporting guidelines are then they may also not understand the purpose of a website that exists to disseminate those guidelines. If the website looks confusing

(as suggested by some survey responses), then perhaps visitors expect the costs of using it (time and effort) outweigh the benefits they might get, if any. The data from this study are not sufficient to draw any conclusions here: although they show engagement is poor, the data do not explain why or how it could be improved. This brings me to the first limitation of this study.

5.4.1 Limitations

Ultimately, numbers can only tell you so much. There are many possible reasons for poor engagement. Perhaps the website's content or structure is too complex, perhaps the design puts people off. These reasons could be explored qualitatively using interviews or think aloud.

One way to infer why visitors come to the website and what they might be trying to achieve is by looking at which website they were viewing previously. For example, if lots of visitors arrive on EQUATOR's website after clicking a link within a journal's submission system this would suggest they are in the midst of submitting an article and might be looking for a checklist. I had been hopeful that Google Analytics would tell us this information. Unfortunately, half of the traffic comes from "unknown" sources. Google Analytics labels traffic as "unknown" when it has no information about where that traffic comes from. Although it's possible that some of this traffic comes from links within offline documents (e.g., PDF publications), it is likely that a most of this traffic represents referral traffic from websites that are linking to EQUATOR using links that start with *http* instead of *https*. When a secure website (with an address that starts with *https*) links to a less secure address (starting with *http*), no referral data gets sent. EQUATOR upgraded its website to use *https* years ago, but journals have continued to link to EQUATOR using the old, less secure, *http* address. Hence a lot of traffic coming from journals and submissions is being labelled as "unknown" and EQUATOR has no way of knowing which journal or stage of submission those referrals are coming from.

Another potential limitation is that Google Analytics counts cookies, not people. If a visitor clears their cookies between visits or uses multiple devices or browsers, the visitor will appear as multiple visitors. Cookies expire after 2 years by default. The proportion of new vs. returning users is thus an overestimation but, nevertheless, still high.

A final limitation is that the data presented here cannot be used to draw comparisons with other websites. For example, I was tempted to compare EQUATOR’s bounce rate of 53% with eCommerce benchmarks (around 40%, from an analysis of 1068 European eCommerce websites [144]). But this comparison isn’t useful. EQUATOR’s website is a free learning resource, not a shop, and its users are different in terms of who they are, what they are trying to do, why they are doing it, and how they get to the website. Consequently it’s not clear whether EQUATOR’s bounce rate should ideally be lower than, higher than, or equivalent to eCommerce websites. Perhaps it could be useful to compare EQUATOR’s website with similar learning resources but, as yet, I’ve not found any with publicly available analytics.

5.4.2 Implications for future work

Whilst writing this chapter, potential ways to improve the website naturally occurred to me and I will list some of those here. In my personal opinion, the website violates some common usability heuristics: the design is unattractive, pages are crammed with text and do not follow standard design principles (for example, the search button is situated in a bizarre location). Any redesign should consider adhering to standard heuristics and best practices. Qualitative research would help understand who is visiting the website, their reasons for doing so, and their opinions and experiences of using it. This may shed light on why so many visitors leave without engaging. Such research should include the perspectives of non-native English speakers as a large proportion of visitors had their browsers set to a language other than English.

Search engine optimisation could help EQUATOR reach more authors. To rank higher in search engines, EQUATOR could add metadata, optimize their website for mobile phones, and take advantage of Google Search’s featured snippets and description features. EQUATOR could consider optimising for keywords that may help catch authors at earlier stages of writing. Google Search Console (a Google product that allows website owners to view how their site performs in Google searches) shows that when users search for “STROBE guidelines,” EQUATOR’s site appears at the top of search results and 36% click this result. However, if a user searches for “how to write an epidemiological report,” EQUATOR drops to 29th place with a click rate of 0%. EQUATOR could

optimize its website to attract naïve authors at an early stage of writing who may not know guideline acronyms.

Optimizing pages for searchability might also help visitors discover content on the site. Currently only 13 reporting guidelines receive meaningful traffic. The rest of EQUATOR's massive website goes largely unvisited. EQUATOR could consider reorganising content to make it easier to find, or pruning dead-weight to simplify the website.

EQUATOR could also ask journals and submission systems to update how they link to EQUATOR's website. This would result in more correct referral data. This data would tell EQUATOR which journals are successfully recommending reporting guidelines and it would allow EQUATOR to infer visitors' intentions. For example, traffic from submission systems may signify authors who are in the very late stages of writing, and may be seeking a checklist.

These ideas – usability heuristics evaluations, search optimization, and updating links – came to mind as I wrote this chapter. However, in my introduction I declared that I wanted to avoid this kind of ad-hoc ideation. These first three chapters had helped me systematically explore and understand authors' experience of reporting guidelines and some of the influences affecting adherence. These influences were a jumble. Some were characteristics of the guidelines themselves (e.g., whether they are easy to understand), the websites disseminating them (e.g., whether they are easy to find and navigate), the authors using guidelines (e.g., their experience in writing). Others pertained to time and context (e.g., discovering guidelines at the point of publication vs drafting), to believed or perceived costs and benefits, and the expectations and policies of others. At this point in my DPhil I needed to make sense of this jumble and begin to systematically identify ways to address influences. I describe this journey in the next 4 chapters, which begin with how I chose a framework to help me.

5.4.3 Conclusions

This chapter presents a first step in evaluating the EQUATOR Network's website. These data will be a useful baseline against which to compare future changes. The findings add support to some of the influences I identified in chapters 3 and 4, and demonstrate

how some of those influences apply to EQUATOR’s website. This context will be useful when considering possibilities to address influences, which is the subject of later chapters.

5.5 Reflections on this chapter

At first my supervisors could not understand why I wanted to write this chapter, or why I wanted to look at the EQUATOR website all. I was warned that a website “is not an intervention”. I do not think they realised how significant their website was (in attracting many visitors) or how much it could be improved. Besides reporting visitor numbers to funders, EQUATOR made little use of their Google Analytics data and had no idea where visitors were coming from, what content they accessed, or how long they stayed. As a software developer these questions were marketing 1-0-1 to me, but I had to avoid jargon like “conversion rates” or “acquisition funnels” as these projected an image of corporate analysis and not research. It took a while for me to learn how to position this chapter and to convince my supervisors it belonged in a thesis. Eventually I believe I succeeded, and I smiled when a supervisor later commented with an echo of my own words: “is it worth reminding the reader that the EQUATOR website and its contents are an intervention to improve transparency and research integrity?”. I’m glad I succeeded, as the findings from this chapter went on to influence discussions in my workshops with EQUATOR staff in chapter 7, my design choices in chapter 9, and were echoed by participants I interviewed in chapter 10.

Part II

Identifying and implementing intervention changes

Chapter 6

Selecting the Behaviour Change Wheel framework

In my introduction chapter I argued that reporting guidelines are part of a system. In chapters 3 - 5 I explored authors' experiences of reporting guidelines and what might influence adherence. Some influences were inherent to the guidelines to themselves or the authors using them, whereas others originated from other parts of the system through which reporting guidelines are disseminated, such as other people, policies, or websites. Until now, I had not used a framework to guide my inquiry, and I had grouped themes in ways I felt would be useful to guideline developers, but these groupings became less useful once it came to unpicking and addressing influences. For example, in chapter 3 my analytic theme "using reporting guidelines has costs, and researchers may not feel that benefits outweigh the costs" felt like a useful and interesting way to summarize 14 codes, including the costs themselves (time, lengthy manuscripts), the perception of those costs, and how they compare to perceived benefits. But in reality these codes are different problems requiring different solutions. In this chapter, I explain how I chose a framework to help me unpick my findings from chapters 3 - 5, understand what drives them, and to identify ways to address them.

I considered the MRC guidance for developing and evaluating complex interventions [145] and the Person Based Approach (PBA) [146]. Both emphasise the importance

of understanding context, evidencing needs, and iterative evaluation and refinement. Both recommend involving stakeholders and the intervention recipients. I have adhered to these principles in my thesis: in chapters 3 - 5 I collected evidence for needs and context, in chapters 7, 8, and 10 I explain how I involved stakeholders and users, and in chapters 10 and 11 I explain how I refined the intervention in response to author feedback and possible next steps for iterative development.

However, at this stage in my DPhil my aim was to address the influences I had identified and described in chapters 3 - 5. Neither MRC guidelines nor PBA advise you how to do this directly. Instead, they encourage intervention designers to theorise how the intervention is working. To do this, designers need to select a theoretical framework. First I considered the Theoretical Domains Framework (TDF) [147,148]. The TDF was developed to help implementation researchers identify what influences the behaviour of health professionals and patients. It was made by synthesising 33 theories of behaviour and behaviour change into 14 domains. These domains cover the cognitive, affective, social, and environmental influences on behaviour. The TDF and its accompanying guide [149] are useful, but they require considerable understanding of psychological constructs that might not be familiar to non-behavioural experts, and although the TDF helps designers understand what influences behaviour, there is little guidance on how to then go about *changing* behaviour.

The Behaviour Change Wheel (BCW) [150] addresses these limitations. It is designed for “policy makers, intervention designers, researchers and practitioners... from a wide range of disciplines with varying levels of expertise”. It was made by synthesizing 19 existing frameworks. In doing so, Michie et al. [150] ensured the BCW is comprehensive (it covers all intervention options), conceptually coherent (category members are of the same type), and linked to an overarching model of behaviour. This was one of their core objectives, as prevailing frameworks at the time lacked these qualities. For example, they criticise MINDSPACE [151] (once popular amongst UK policy makers) for not including all intervention types and for conflating psychological constructs with mechanisms of action and modes or delivery.

The BCW’s core is the COM-B model of behaviour which is simpler and more intuitive

than the TDF and stands for Capability, Opportunity, Motivation and Behaviour. These are defined in Table 6.1 but, briefly; capability includes having the required knowledge, skills, stamina, and physical strength; opportunity includes the necessary resources such as time, materials, or reminders; and motivation means having the required intentions, beliefs, wants and needs to perform the behaviour. When making COM-B, Michie et al. began with motivation, defined as: “brain processes that energize and direct behaviour” [152]. Their next step was “to consider the minimum number of additional factors needed to account for whether change in the behavioural target would occur, given sufficient motivation” [150]. They used a US consensus meeting of behavioural theorists [153] and a centuries-old principle of criminal law: that to prove guilt one must show means or capability, opportunity, and motive. Michie et al. argued that the “common conclusion from these two separate strands of thought lends confidence to this model of behaviour”[150], and they added further confidence by mapping COM-B to the TDF. When describing COM-B, Michie et al. explain how the components capability, opportunity, and motivation interact with each other to generate behaviour which, in turn, may influence these components in a circular way. (See Figure 6.1)

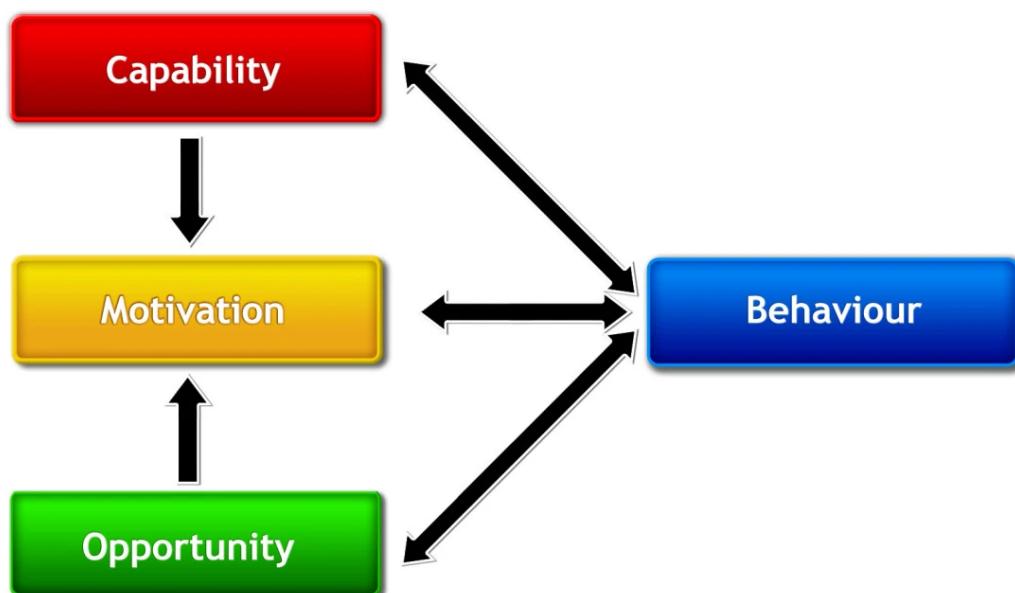


FIGURE 6.1: The COM-B system for understanding behaviour, reproduced from Michie et al. [150] used under Creative Commons Attribution License

Changing behaviour requires altering one or more of these components and the BCW goes beyond the TDF by helping intervention designers select options to achieve this.



FIGURE 6.2: The Behaviour Change Wheel for designing behaviour change interventions. The innermost ring is the COM-B model for understanding behaviour, the middle ring are intervention functions, and the outer ring are policy categories. Reproduced from Michie et al. [150] used under Creative Commons Attribution License

The wheel's middle ring contains intervention functions, defined as "broad categories of means by which an intervention can change behaviour" [154]. An intervention may have multiple functions. For example, a television advert about drink driving may educate its audience with information about alcohol limits and persuade using emotive imagery. In this example, education is being used as an intervention function to target psychological capability (teaching people how much they can safely drink before driving) and persuasion is being used to target motivation (targeting people's emotions to make them less likely to want to engage in drink driving). The functions you choose depend on the behavioural drivers you are trying to target. Intervention functions are described in Table 6.2.

Having decided what functions your intervention needs, the guide helps you decide how to implement them with policy categories, defined as "types of decisions made by authorities that help to support and enact the interventions" [154]. Categories include communication/marketing, guidelines, regulation, and service provision, as described in Table 6.3.

Recognizing that intervention designers come from various disciplines and places of work, and that not all will have experience in applying behaviour change theory, Michie et al. published a book to help designers apply the BCW [154]. This book contains suggested steps, exercises, and worksheets. In the next chapter, I describe how I used this approach to guide reporting guideline experts from the UK EQUATOR Centre through applying the BCW.

TABLE 6.1: COM-B model components and examples, reproduced from Michie et al. [150] used under Creative Commons Attribution License

Component	Definition	Example
Physical capability	Physical skill, strength, or stamina	Having the skill to take a blood sample
Psychological capability	Knowledge or psychological skills, strength, or stamina to engage in the necessary mental processes	Understanding the impact of carbon dioxide on the environment

Component	Definition	Example
Physical opportunity	Opportunity offered by the environment involving time, resources, locations, cues, physical ‘affordance’	Being able to go running because one owns appropriate shoes
Social opportunity	Opportunity afforded by interpersonal influences, social cues and cultural norms that influence the way that we think about things, e.g., the words and concepts that make up our language	Being able to smoke in the house of someone who smokes but not in the middle of a board meeting
Reflective motivation	Reflective processes involving plans (self-conscious intentions) and evaluations (beliefs about what is good and bad)	Intending to stop smoking
Automatic motivation	Automatic process involving emotional reactions, desires (wants and needs), impulses, inhibitions, drive states and reflex responses	Feeling anticipated pleasure at the prospect of eating a piece of chocolate cake

TABLE 6.2: Intervention function definitions and examples, reproduced from Michie et al. [150] used under Creative Commons Attribution License

Intervention		
Function	Definition	Example
Education	Increasing knowledge or understanding	Providing information to promote healthy eating
Persuasion	Using information to induce positive or negative feelings or stimulate action	Using imagery to motivate increases in physical activity

Intervention

Function	Definition	Example
Incentivisation	Creating an expectation of reward	Using prize draws to induce attempts to stop smoking
Coercion	Creating an expectation of punishment or cost	Raising the financial cost to reduce excessive alcohol consumption
Training	Imparting skills	Advanced driver training to increase safe driving
Restriction	Using rules to reduce the opportunity to engage in the target behaviour (or to increase the target behaviour by reducing the opportunity to engage in competing behaviours)	Prohibiting sales of solvents to people under 18 to reduce use for intoxication
Environmental restructuring	Changing the physical or social context	Providing on-screen prompts for GPs to ask about smoking behaviour
Modelling	Providing an example for people to aspire to or imitate	Using TV drama scenes involving safe-sex practices to increase condom use
Enablement	Increasing means/reducing barriers to increase capability (beyond education and training) or opportunity (beyond environmental restructuring)	Behavioural support for smoking cessation, medication for cognitive deficits, surgery to reduce obesity, prosthesis to promote physical activity

TABLE 6.3: Policy categories, reproduced from Michie et al. [150] used under Creative Commons Attribution License

Policy Category	Definition	Example
Communication / marketing	Using print, electronic, telephonic or broadcast media	Conducting mass media campaigns
Guidelines	Creating documents that recommend or mandate practice. This includes changes to service provision	Producing and disseminating treatment protocols.
Fiscal measures	Using the tax system to reduce or increase the financial cost	Increasing duty or increasing anti-smuggling activities
Legislation	Making or changing laws	Prohibiting sale or use
Environmental / social planning	Designing and/or controlling the physical or social environment	Establishing support services in workplaces, communities etc.

6.1 Reflections on this chapter

At this point in my thesis I had my transfer of status. It was a pivotal moment where I began to view reporting guidelines differently, Charlotte Albury joined my supervision team, and the direction of my thesis took a qualitative turn.

In my introductory chapter I presented why I view reporting guidelines to be part of a complex behaviour change intervention, but I have not always viewed them this way. When I first discovered reporting guidelines in 2014, I would have described them simply as recommendations for authors to adhere to when writing-up research. I would have talked about checklists, full guidance, and journal policies. But I would not have zoomed in to the details of those parts: I would not have scrutinized the contents of resources, their design, nor the intricacies of editorial policy. I also would not have zoomed *out* to view the broader picture of how guidelines are created and disseminated, or the connections between these parts.

But after reading accounts of authors struggling with the guidance content, access, formats, workflows, confidence, and the behaviour of others, I began to see things differently. Instead of revering guidelines as perfect publications set in stone, I began to see them as resources whose content, structure, and layout could be optimized. I began to see details; how a single word could lead to confusion, the nuances between *standards* and *recommendations*, between *drafting*, *revising*, and *checking* manuscripts. Beyond the guidelines themselves, I began to see the broader system that creates and disseminates them. I began to recognise the complexity arising from the number of different stakeholders involved, differences between guidelines, the skills and prior knowledge an author must have to act on the guidance, and variation in how, when, and why guidelines are used.

This simultaneous zoom-in and zoom-out felt confusing, and I struggled to communicate my thoughts to my initial supervision team - Gary, Jen and Michael - who worked for EQUATOR and shared a different, perhaps more traditional, view of reporting guidelines.

Discovering the MRC guidance for complex interventions was my first turning point. Its description of complexity resonated and helped me understand how an intervention could be a *system*, not just a single *thing*. From there, I discovered the world of behaviour change and took a few tentative steps. My transfer assessors foresaw my thesis becoming increasingly qualitative, felt I would benefit from a qualitative supervisor, and recommended Charlotte Albury. Charlotte's expertise emboldened my exploration of methods less familiar to EQUATOR and I, and under her guidance my tentative steps into qualitative behaviour change territory became more certain.

Chapter 7

Following the behaviour change wheel guide: Workshops with reporting guideline experts

7.1 Introduction

Having identified influences affecting whether authors adhere to reporting guidelines (chapters 3 - 5) and selected the Behaviour Change Wheel as a framework [150] (chapter 6), my next step was to use the framework to unpick these influences, diagnose what drives them and how they could be addressed. In my introduction I describe how the reporting guideline system grew organically and I justified why I wanted to redesign this system and the guidelines themselves using evidence and behaviour change theory.

In “The Behaviour Change Wheel - A Guide To Designing Interventions”[154], Michie et al. suggest eight steps to help intervention designers understand behaviour and identify intervention and implementation options. The guide is aimed at intervention designers who are not behaviour science specialists, and so offered a practical way to include stakeholders from the reporting guideline ecosystem. I wanted to do this because I expected that input from experts with intimate knowledge of reporting guidelines would lead to more ideas, and that these ideas may be more likely to gain traction and have

impact.

I had to be mindful of how much time I could expect stakeholders to give to my project. The eight stages outlined by Michie et al. would take many hours and require background familiarity with the COM-B model, intervention functions and policy categories. This seemed like too much to ask of strangers uninvested in this work, and so I decided to begin by involving reporting guideline experts with whom I already had a relationship, and who were already interested in the project: members of the UK EQUATOR Centre. See chapter 1 for an introduction to the EQUATOR Network, and see chapters 8 and 10 for how I sought input from wider stakeholders and authors later.

In this chapter, I describe how I led members of the UK EQUATOR Centre through the intervention design process outlined in “The Behaviour Change Wheel - A Guide To Designing Interventions”[154]. I describe the methods and results for each stage in turn.

7.2 Methods

Of the 7 members of the UK EQUATOR Centre I invited, 6 took part in the workshops. These members had experience in developing and disseminating reporting guidelines, and in training authors how to use them.

I led the workshops and actively participated too. I felt justified drawing on my own experience from my earlier DPhil work, as an author, and developer of tools to help authors (see chapter 2). Together, we completed online worksheets from Michie et al’s book @[154]. Hence my paradigm was constructivist in that knowledge was “constructed between inquirer and participant through the inquiry process itself”[155]. Constructivism “rejects the idea that there is objective knowledge in some external reality for the researcher to retrieve mechanistically” and instead, “the researcher’s values and dispositions influence the knowledge that is constructed through interaction with the phenomenon and participants in the inquiry”[155].

On one hand I expected that my experience would be an asset, and would contribute to our aim of understanding and addressing the phenomena we were interested in. On

the other hand, I wanted to ensure that my opinions did not restrict the group, that I remained open-minded, and that I captured the thoughts of other workshop participants accurately. And so I used a number of established techniques to enhance trustworthiness and facilitate discussion.

I used Lincoln and Guba's criteria for trustworthiness [156], which asserts that for a study to be trustworthy, the researcher must show that the findings are credible ('true'), transferable (applicable to other contexts), dependable (consistent and repeatable), and confirmable (shaped by participants, not by the researcher's bias or motivation). I describe the techniques I used to achieve each criterion in Table 7.1.

TABLE 7.1: Techniques for establishing trustworthiness. Based on Lincoln and Guba's Evaluative Criteria [156]

TECHNIQUE	IMPLEMENTATION
Techniques for establishing credibility	<i>Member-checking</i> Lincoln and Guba argue that member checking is the most important way to establish validity of an account [156]. All workshop participants could edit the worksheets during and after each workshop. At the end of each session, I would invite workshop participants to confirm that their thoughts were reflected in the file, and I invited participants to comment on my written account of the workshops.
Techniques for establishing transferability	

TECHNIQUE	IMPLEMENTATION
<i>Thick description</i>	Although perhaps most relevant to ethnographic studies, I nevertheless drew on relevant aspects of ‘thick description’. I encouraged participants to describe ideas in detail and to document disagreement and context. Our aim was to document ideas in sufficient detail so they could be transferable across different reporting guidelines and stakeholders.
Techniques for establishing confirmability	
<i>Reflexivity</i>	I wrote down my own ideas before each workshop, and made notes at the end of each workshop to reflect on the process. I would invite participants to consider the worksheet on their own or in pairs. I did this so that everybody would engage with the task and have an opportunity to think before being influenced by others. We then discussed ideas as a group, with members agreeing, disagreeing, and bouncing off each other. I withheld my own ideas until the end of a session or task, whereupon I would invite discussion on any ideas that I felt hadn’t been covered already. This allowed me to contribute my experience from previous DPhil work, as an author, and as a software developer and give others a chance to discuss my thoughts, without biasing or narrowing discussion.
<i>Audit trail</i>	All worksheets were stored on the University One Drive account which performs automatic saves and versioning. This created an audit trail which meant I could look at how a worksheet changed over time as participants added and edited content.

To encourage open discussion, I encouraged participants to rise above their own pre-conceptions and reassured them that there were no wrong answers, that all ideas were valid and should be documented. To facilitate rich discussion [157] I used open-ended questioning, left space for participants to talk, and followed Michie et al.'s worksheets which structure inquiry around frameworks, models, and taxonomies.

We met 7 times between December 2021 and May 2022, and each online meeting lasted around 2 hours. All participants had access to the Behaviour Change Wheel book. We established some ground rules which were that no idea was a bad idea, we should favour evidence over preconceptions, and that we should aspire to challenge our own assumptions and be open minded as far as possible. We didn't seek consensus. Instead, we kept note of any disagreements that could not be resolved by discussion.

I explained the objectives and any background theory at the start of each step. I will now summarise each step and our discussions. Our co-edited worksheets are included in Appendices E–K. I purposefully use “we” in this chapter to reflect that my voice is included.

7.2.1 Ethics

Oxford University's Medical Sciences Interdivisional Research Ethics Committee judged ethical approval unnecessary.

7.2.2 Reporting

I used SRQR [30] to draft and check this chapter (See appendix D).

7.3 Results

7.3.1 Step 1. Defining the problem in behavioural terms

Michie et al.'s first step is to define the problem in terms of *who* needs to do *what*. For example, *weight loss* is not a behaviour, but *increasing physical activity* is, and could be specified further as *walking 10,000 steps a day*. We were all in alignment here: *we*

want researchers to include important details in their articles, in line with the relevant reporting guidelines. Our notes from the workshop are in Appendix E.

7.3.2 Steps 2 & 3. Selecting and defining the target behaviour

“Behaviours do not exist in a vacuum but occur within the context of other behaviours of the same or other individuals” write Michie et al. [154] when explaining that the desired behaviour need not be the behaviour that you target. For instance, if you want a child to eat more fruit (desired behaviour), you may seek to influence what food their parent buys (target behaviour).

Step 2 involves generating a longlist of candidate target behaviours that could bring about the desired behaviour, and then selecting which behaviour(s) to focus on.

We thought of many target behaviours involving authors, such as “reading the guidance in full” or “studying guidance (in an abstract sense)”. We also considered targeting people other than authors. For instance, “peer reviewers” and “editors” could “check articles against guidelines [and] tell authors what is missing”, and supervisors could “encourage” the use of guidelines.

This exercise helped us break down things that we had not questioned previously. For instance, we were forced to define what we actually meant by a “reporting guideline”, noting that guidance could be distributed across publications, checklists, or supplements. We wanted people to “use the *full* guidance (often reported in an *Example & Elaboration* document)” and “not just the checklist”.

Similarly, where we may previously have thought of “writing” as a single task, we began to consider how authors could “use guidance when planning”, “drafting”, “editing [their own work]”, or “checking”. We discussed how these behaviours required researchers to be open-minded to assistance and wondered if we could encourage them to “ask for help when writing”.

Ultimately we ended up with a longlist of 14 possible target behaviours, listed in Appendix F. The next task was to narrow our longlist by considering:

1. The expected impact if the behaviour were to be performed

2. How easy we expected behaviour change to be
3. The centrality of behaviour - how close it was to our desired behaviour
4. How easy the behaviour will be to measure

Criteria two, three, and four lead us to prioritise the behaviour of authors above that of editors or peer reviewers. We felt authors to be central as ultimately they have control over what gets written. Editors, peer reviewers and other stakeholders are less central because all they can do is ask an author to edit their writing. We also felt that changing the behaviour of editors and reviewers would be harder to instigate and to measure. Although journal policies are easy to audit, the actual behaviour of the editorial team and peer reviewers are not. In contrast, we felt authors' behaviour was easier to measure through literature audits, surveys, and web analytics for EQUATOR's website. When considering criteria one and two, we felt that authors were most likely to apply guidance successfully if they used it early when writing. Our workshop notes for step 2 are in Appendix F.

Step three of the Behaviour Change Wheel involved bringing these thoughts together and specifying the behaviour in more detail. Michie et al. suggest defining *who* needs to perform the behaviour, *what* do they need to do, *when* will they do it, *where* will they do it, how *often* will they do it, and *with whom* they will do it. Our final definition of our target behaviour was: **Researchers should use reporting guidance as early as possible in their research pipeline. They will do this for every piece of research, at their place of work, on their own but in the context of collaboration**

We broke this key behaviour into two sub behaviours:

1. Engage with reporting guidelines as early as possible (i.e., access and read them) and,
2. Apply the guidance to their writing as intended by the guideline developer.

By "research pipeline" we mean the many steps involved in a typical project which may include ideation, designing, writing a protocol, obtaining funding and ethics permission, drafting a manuscript, editing a manuscript, submitting a manuscript. Instead of specifying *which* stage reporting guidance should be used, we decided to specify that

we want authors to use guidance as “early as possible”. We did this for a few reasons. Firstly, guidelines differ in how easily they can be used for writing protocols or applications, not all disciplines have culture of writing and publishing protocols, and not all research projects will begin with a written funding application, ethics application, or protocol. Secondly, researchers will naturally come across reporting guidelines at different stages of their work. Should a researcher discover a guideline at the point of journal submission, then we would still want them to apply the guidance then, even though this is a relatively late stage. But by specifying “as early as possible”, we declare our hope that *next time* that same researcher may decide to use a guideline at an earlier point.

By “apply”, we refer to using guidance to plan, write or edit a written description of research (e.g., within a manuscript or application). By “the guidance” we mean the fullest form of the reporting guideline available. Ideally this will be an example and elaboration paper, but for some guidelines it will just be the checklist. Applying guidance may include the use of tools like templates or checklists. Participants discussed specifying “completing a reporting checklist” as a target behaviour, but decided against it as previous research showed that authors who complete checklists upon submission don’t necessarily edit their manuscript or comply with guidelines. Participants also recognised that focussing on checklists may be problematic because checklists appear administrative, are used *after* a manuscript has been written, at which point authors are least able or motivated to edit their work. Nevertheless, participants recognised that checklists will continue to be an important part of how reporting guidance are disseminated, and so they are included within the term “apply the guidance”, without being named.

We specified authors should apply guidance “as intended by the guideline developer” because thematic synthesis revealed authors may misinterpret guidelines and believe they are adhering when they are not. Our workshop notes for step 3 are in Appendix G.

Hence our target behaviour definition specifies the *who, what, when, where, how often*, and *with whom* but is broad enough to account for differences between reporting guidelines and researchers’ working practices.

7.3.3 Step 4. Identify what needs to change

This step involved identifying what needs to change in the person and/or environment to achieve our target behaviour. Michie et al. [154] provide a questionnaire to facilitate this step, called the *COM-B Questionnaire*, which asks you to consider each COM-B domain in turn by considering 23 items like “To perform the target behaviour, authors would... have to know more about why it was important”, “...have more time to do it”, and “...have more support from others” etc.

Michie et al. [154] emphasise the importance of evidence in this step, recommending that data should be “collected from as many relevant sources as possible” and “triangulated”, as a consistent picture of behaviour from multiple sources will “increase confidence in the analysis”. Consequently, after first brainstorming using the worksheet (see appendix H), we went through the influences I identified in chapters 3 - 5 to support or refute our thoughts and to ensure we did not miss anything. We consolidated our thoughts with my descriptive themes from my qualitative evidence synthesis, additional codes from my review of survey content, and my findings from evaluating EQUATOR’s website into 32 influences that we felt needed to change for our target behaviour to occur. We used the COM-B questionnaire to label each influence as being driven by capability, opportunity, or motivation. We grouped thoughts and data so that each influence pertained to a single behavioural driver. I added context from our discussion and my previous chapters where I felt appropriate. I circulated this list to workshop attendees for comments and revisions. You can see this list in Appendix I.

7.3.4 Steps 5 & 6. Identify Intervention Functions and Policy Categories

Having defined our target behaviour and identified what needs to change for that behaviour to occur, the next step was to consider *how* to achieve those changes. Michie et al. [154] stress the importance of “considering the full range” of possible intervention functions and policy categories available. See chapter 6 for a fuller introduction to these terms but, briefly, an intervention function is a “category of means by which an intervention can change behaviour” and policy categories are options for delivering those functions. For example, the function *modelling* could be delivered through a communication campaign or a service.

Michie et al. [154] recommend using the APEASE criteria to prioritise options, which stands for Affordability, Practicability, Efficacy, Acceptability, Side-effects, and Equity. They note that whilst effectiveness is key, the other criteria must also be considered as “behaviour change interventions operate within a social context”.

We saw enablement, education, training, persuasion, modelling, and environmental restructuring as favourable intervention functions, ranking well on all APEASE criteria. Of these, enablement was viewed as a particularly low-hanging fruit that would likely be effective and welcomed by the research community.

We found the remaining intervention functions problematic. Incentivization and restriction (e.g., rewarding guideline adherence with funding or reduced article processing charges, or punishing non-adherence by withholding these benefits) were seen as inequitable because as long as guideline adherence is harder for some researchers than others, less-experienced, poorly resourced researchers would be at a disadvantage, as would those working in disciplines where reporting guidelines are poorly designed or harder to apply. Conversely, we felt that enablement has potential to reduce existing inequalities. In addition to being inequitable, participants voiced that restriction or punishment would be unacceptable to researchers, as would coercion (the threat of punishment). Furthermore, threats without enforcement may become known as pointless administration, lose effectiveness, and erode trust.

Regarding policy categories, we saw environmental planning as the most affordable, effective, acceptable, safe, and equitable. When talking about the *environment*, we were really talking about the *digital* environment, within which EQUATOR has control over its own website but not websites or digital services run by others. Improving and extending the website was viewed as an affordable investment, that would be welcomed by authors (thus acceptable), and a practical way to increase the proportion of authors engaging with reporting guidelines that would increase equity without side effects.

We recognised communication as a favourable policy category that is already utilised. Communication channels included the website, mailing lists, social media, conferences, and publications. Whilst communication was rated as affordable, practicable,

acceptable, safe, and equitable, its effectiveness may be limited. Although communication could promote awareness of reporting guidelines are available, what they are, and how they can be used, EQUATOR members noted that the efficacy of a communications campaign may be undermined if the campaign directs authors to a website that is difficult to use, or guidelines that are difficult to access or understand. Thus communication on its own might not be sufficient, unless proceeded by improving the digital environment.

Service provision was also favoured, as long as the service was financially sustainable. So too were guidelines, specifically guidance to help reporting guideline developers create and disseminate resources. Participants felt that legislation, regulation, and fiscal measures would not be acceptable to researchers and were not practical options for EQUATOR to use.

Our notes from steps 5 and 6 are in Appendices J and K.

7.3.5 Step 7. Identify behaviour change techniques

The aim of step 7 was to identify possible behaviour change techniques by systematically considering items from a taxonomy of 93 techniques [158] for each intervention function chosen in step 5. Michie et al suggest doing it this way because “the process of designing behaviour change interventions usually involves first of all determining the broad approach that will be adopted and then working on the specifics of the intervention design. For example, when attempting to reduce excessive antibiotic prescribing one may decide that an educational intervention is the appropriate approach. Alternatively, one may seek to incentivise appropriate prescribing or in some way penalise inappropriate prescribing. Once one has done this, one would decide on the specific intervention components”. I interpret this passage as Michie et al. [154] suggesting that intervention designers will select only one or two intervention functions, and then consider all behaviour change techniques relevant to that function. However, picking a “broad approach” didn’t feel helpful to us. Given that a system already exists for disseminating reporting guidance, we found ourselves considering how we could use multiple

intervention functions to refine this system. Consequently, because we were not limiting our approach to one or two intervention functions, it meant we could not filter the list of behaviour change techniques and had to consider most of them.

Additionally, although the taxonomy is designed to be applicable to a range of contexts and intervention types, it didn't always feel like a perfect fit for our needs. Some techniques are explicitly health-focussed (e.g., *Body changes*, *Pharmacological support*, and *Information about health consequences*) whereas for others the link with health interventions came from the examples provided. For example, the technique *practical social support* lists the example of "Ask the partner of the patient to put their tablet on the breakfast tray so that the patient remembers to take it". It's not immediately obvious how this technique could generalise to our target behaviour. The taxonomy developers perhaps acknowledge this limitation, suggesting in their discussion that the list can be viewed as a "core" taxonomy that can be modified or extended according to context [158].

Hence, given that we didn't want to choose a "broad approach", and given that this step required familiarity with the taxonomy and how it can be adapted to our context, doing this step with EQUATOR staff would have been very time consuming, as others have also found [159]. EQUATOR staff asked for a more pragmatic approach.

Instead, I did this step on my own. I was well placed to do this because I was familiar with the taxonomy and the Behaviour Change Wheel Guide, I had led all workshops, and I was most familiar with the influences we were trying to address. I went through each intervention function the group had favoured in step 5 (Enablement, Education, Training, Persuasion, Modelling, and Environmental Restructuring), considered all relevant behaviour change techniques, and used the APEASE criteria to choose ones I felt were appropriate.

Whilst I did that, EQUATOR members considered each of our favoured intervention functions and brainstormed 24 ideas to address the 32 influences identified in step 4. I describe these ideas fully – alongside ideas captured through focus groups with other stakeholders – in the next chapter. I then deduced the behaviour change techniques these ideas used and cross checked them against my list of favoured behaviour change

techniques, adding any I had missed. The final list of 39 techniques is shown in Table 7.2.

TABLE 7.2: Favoured techniques from the Behaviour Change Technique taxonomy v1.0 [158]

Intervention	Possible Behaviour Change Techniques
Function	
Enablement	Social support (practical) Adding objects to the environment Restructuring the physical environment (which I took to include the digital environment) Reduce negative emotions Self-monitoring of outcome(s) of behaviour Framing/reframing Identification of self as role model Salience of consequences Anticipated regret Vicarious consequences
Education	Information about social and environmental consequences Information about health consequences (which I took to mean <i>other</i> people's health) Feedback on behaviour (e.g., feedback on guideline use) Feedback on outcomes of the behaviour (e.g., feedback regarding quality of reporting from editors, reviewers, or colleagues) Self monitoring of outcome(s) of behaviour (e.g., checking one's own work against guidance) Information about emotional consequences (e.g., telling authors they how they will feel) Information about others' approval

Intervention

Function	Possible Behaviour Change Techniques
Training	Instruction on how to perform the behaviour Demonstration of the behaviour Feedback on the behaviour Feedback on the outcome(s) of the behaviour Self-monitoring of outcome(s) of the behaviour
Persuasion	Credible source Information about social and environmental consequences Information about health consequences Feedback on behaviour Feedback on outcome(s) of the behaviour Monitoring outcome(s) of behaviour by others without feedback Identification as self as role model Information about emotional consequences Salience of consequences Information about others' approval Social comparison Framing/reframing Remove aversive stimulus
Modelling	Demonstration of the behaviour
Environmental	Adding objects to the environment
Restructuring	Prompts/cues Restructuring physical environment

7.3.6 Step 8. Identifying delivery options

I then had to decide how to deliver our favoured behaviour change techniques, intervention functions and policy options. Michie et al.'s guidance for this step is more open ended [154]; although they offer delivery options for communication as an example intervention function, there is no framework or systematic approach to this step and instead they recommend that designers consider the delivery options that they have

at their disposal. For EQUATOR, that meant developing a new service or training programme, running a communication campaign, developing new guidance, or improving their website. It was important that my choice was feasible within my funding window, and I wanted it to have a lasting impact after my DPhil was finished.

Creating a new service felt unsustainable as it would likely stop once my funding ran out. Developing training or guidance for guideline developers could be useful and acceptable but not achievable within my time constraint. I could have developed a communications campaign, but EQUATOR members felt this was something they could do independently and that it should be done after addressing other influences – a communications campaign promoting a guideline would be less impactful if the guideline is hard to understand or use.

Instead, redesigning reporting guidelines and key parts of EQUATOR's website felt like the perfect choice for multiple reasons. We decided that displaying redesigned guidelines as a webpage would make them easy to disseminate, easy to access, and would enable functionality not available in traditional publications. These redesigned guidelines could become part of the existing EQUATOR Network website. We also felt that EQUATOR's home page was particularly important to focus on, as it is the most visited page of the website where the highest proportion of visitors disappear.

Improving guidelines and the website aligned with our prioritised options. Planning the digital environment was our highest ranked policy category. We could use the website to deliver many of our favoured intervention functions (enablement, education, persuasion, modelling). It spoke to my skills as a software developer. Finally, these changes could be made within the time limit of my DPhil and would have a lasting impact.

7.4 Discussion

Following Michie et al.'s guide [154] helped us to specify our target behaviour, understand the behavioural drivers behind 32 influences on our target behaviour, select intervention functions, policy categories, behaviour change techniques, and delivery options.

More fundamentally, the process helped the workshop participants to begin thinking about reporting guidelines as a behaviour change intervention. The process helped us break down the differences between tasks (e.g., writing vs. editing vs. reviewing research articles), users (e.g., inexperienced vs. experienced researchers, editors, reviewers), resources (e.g., explanation and elaboration documents vs. checklists). It helped EQUATOR staff to look at the current system objectively and it helped us challenge our preconceptions.

One of the most interesting parts of this process was witnessing an unexpected change of opinion amongst EQUATOR staff. Before beginning this study, a common refrain heard around the office was that in order for reporting guidelines to be successful editors had to start enforcing them and refuse to publish research that did not adhere. So it was fascinating to witness workshop participants unanimously rating restriction and coercion as their least favourite options.

I think two things happened here. Firstly, having discussed the challenges authors face when using reporting guidelines, participants felt that forcing authors to use them would be unacceptable to researchers, impractical for editors, and inequitable as some authors would face larger hurdles than others. Secondly, participants reassessed things they had taken for granted, and realised that there are many low-hanging fruit that could make guidelines easier to find and use, and that these fruits were growing in their own orchard.

Using a framework helped us to systematically consider a full range of options, many of which may not have come to mind naturally. However, sometimes it was difficult to get participants to think “outside of the box”. The default was to think about how the existing system could be improved, and it was difficult to imagine a world where we could be starting from scratch. In one sense this was an opportunity, as improving a system that already exists is easier than creating something totally new. But it could also be seen as a limitation, as our imagination may have been constrained by what already exists.

We may also have been limited by group-think [160]. All workshop participants had worked for the EQUATOR Centre for many years and had many shared opinions and

experiences. Including other stakeholders in these workshops would have helped address this but would have been impractical to coordinate. To mitigate this, I decided to gather input from guideline developers, publishers, and authors through separate pieces of work which I describe in chapters 8 and 10.

In conclusion, by working through Michie et al's suggested approach to applying the Behaviour Change Wheel [154] in a series of workshops with members of the EQUATOR Network, we defined our target behaviour as "*Researchers should use reporting guidance as early as possible in their research pipeline. They will do this for every piece of research, at their place of work, on their own but in the context of collaboration*". We broke this down into two sub behaviours: 1) engage with reporting guidelines as early as possible (i.e., access and read them) and, 2) apply the guidance to their writing as intended by the guideline developer. We identified 32 influences that could affect this target behaviour. We favoured Enablement, Education, Training, Persuasion, Modelling, and Environmental Restructuring as intervention functions, and we favoured Environmental Planning, Communication, Service Provision, and Guidelines as policy categories. I identified 39 behaviour change techniques that could be used, and decided my focus should be redesigning reporting guidelines and the EQUATOR Network home page.

7.5 Reflections on this chapter

At times these workshops felt a little awkward to run. I was very aware of being "just" a student and my imposter syndrome was strongest when conversations questioned EQUATOR's identity. I'm used to a corporate approach where a company's vision and objectives are set at a high level and all work is expected to support those goals. Although EQUATOR does have organisational objectives, they have not been updated since its formation and its staff are not tethered to them. As with any research group, staff are free to pursue their own research interests. This academic freedom is important, but because staff are not aligned behind a core set of values or objectives, some of our workshop discussions touched on bigger questions like *what does EQUATOR stand for?, what should it stand for in the future?*, or the differences between the EQUATOR

Network, the *UK EQUATOR Centre*, and us as a research group. One of the biggest debates was around whether EQUATOR (and reporting guidelines) should be advising researchers on how to design studies. Design and reporting go hand in hand: one reason reporting is important is to allow the reader to assess the design. However, the two are separate research tasks, done at different times, and perhaps by different people. Ultimately the group decided reporting and design should be linked but kept separate, and that EQUATOR should stick to its founding objectives of improving reporting. This distinction has settled comfortably in my head, but perhaps not so much in others', especially those whose interests lie in experimental design and who perhaps see EQUATOR and reporting guidelines as trojan horses for disseminating design advice.

This is one way in which the interests of participants influenced workshop dialogue and outputs. As have conversations outside of the workshops. We are all colleagues, and continually talk about our work. At the time of these workshops a few colleagues were involved in a study assessing adherence to reporting guidelines. When they assessed the same manuscript they often disagreed whether an item was reported properly. In an attempt to solve this problem, another colleague created an assessment tool that deconstructs reporting guideline items into their smallest parts and asks assessors to confirm the presence and absence of each sub item. Implicit in these conversations is a perception that reporting is either good or bad. Information is either present or absent. On first glance this makes sense and at the time of the workshops I never really questioned it. But upon revising this chapter I realise that my thinking may have shifted since. Does this black and white mentality always hold? Some reporting items are more subjective than others, and a black and white approach to assessment makes less sense when guidelines explicitly acknowledge this subjectivity. For example, this section is partly an attempt to fulfil SRQR's item 6 (Researcher characteristics and reflexivity) which also appears in JARS-Qual, where it warns "It may not be possible for authors to estimate the depth of description desired by reviewers without guidance". To me, this reads as "good reporting is in the eye of the reader" and is an acknowledgement of subjectivity. What might satisfy one reader may not satisfy another. So a black-and-white approach to adherence might not always be possible or best. If I had raised these

questions during the workshop, they might have led to additional ideas or lines of dialogue.

What other assumptions may have shaped my workshops? In writing this section I realised that describing my “stance” is difficult because some assumptions only became clear to me when they were challenged or shifted. Describing my stance feels easier when I have something to describe it against. Perhaps I find this difficult because I’m inexperienced, or perhaps because my community — EQUATOR and the reporting guideline crowd — are quite homogenous in their own stances and so I had less to contrast myself against until my research led me to question and disagree with prevailing attitudes. Implicit assumptions have undoubtedly influenced this chapter (and others). As I gain experience, I hope these assumptions will surface as I continue to reflect on them.

Chapter 8

Generating ideas to address influences: focus groups with reporting guideline developers, advocates, and publishers

8.1 Introduction

In chapter 7 I described how I led reporting guideline experts from the EQUATOR UK Centre through the stages recommended in Michie et al.'s book "The Behaviour Change Wheel - A Guide to Designing Interventions" [154]. This guide helped us define our target behaviour, identify 32 influences that had to change for our target behaviour to occur, and to prioritise lists of intervention functions, policy categories, possible behaviour change techniques, and delivery options. However, these lists were abstract and still had to be *actualised*. For instance, although we had prioritised *Restructuring the physical environment* as a possible behaviour change technique, *how* could the environment be restructured? Similarly, we prioritised Enablement, Persuasion, or Education, but *how* and *where* could these functions be applied? We had vague notions that redesigning reporting guidelines, the EQUATOR website, developing new guidance and communication campaigns might be useful, but no tangible plans.

Thus my next objective was to gather concrete ideas on how these abstract concepts could be realised to address reporting guideline limitations. In chapter 7 I explained how I began this process with EQUATOR staff as part of step 7 of the workshops, when I asked them to consider how we could use our preferred intervention functions to address influences. This was an opportunity to invite input from broader stakeholders. Although EQUATOR is a key part of the reporting guideline landscape, they are *only* a part of it. Ultimately, all stakeholders within the academic system influence the impact of reporting guidelines. Guideline developers and publishers arguably do so most directly, and so it was important to draw on their experience and opinions. It hadn't been feasible to include these stakeholders through all stages of the Behaviour Change Wheel approach as the time commitment was too great, and it would have required stakeholders to become familiar with the framework and its terminologies which was too big an ask. In contrast, brainstorming ideas was a convenient and important stage to include them in. I expected that seeking input from a more diverse group would lead to *more* ideas and that those ideas would be more likely to gain traction.

In this chapter I explain how I went through this brainstorming process by 1) running workshops with EQUATOR staff members and 2) running focus groups with other stakeholders before describing the combined results.

8.2 Methods

The purpose of this study was to elicit ideas from experts familiar with reporting guidelines, on how to address influences that may affect adherence. My methods had two parts: 1) brainstorming ideas with EQUATOR during workshops and 2) extending this list through focus-groups with guideline developers, publishers, and guideline advocates. I then describe the combined results.

8.2.1 Workshopping Ideas with EQUATOR

I described the workshops I ran with EQUATOR in chapter 7, including the techniques I used to encourage rich discussion and to navigate my position as a participating researcher, contributing ideas myself whilst facilitating others to share their voices. In

step 7 of those workshops, I asked participants to consider each intervention function in turn and suggest ideas employing it. After the workshop, I then labelled each idea with the influence(s) it would be addressing to create an “ideas table” with two columns.

The left hand column listed influences. These came from the consolidated list of *what needs to change for our target behaviour to occur* that came out of step four of the Behaviour Change Wheel workshops in the previous chapter. The influences were initially worded as statements (e.g. “Researchers may not know when reporting guidelines should be used”). To make discussions fluid, I rephrased each to a question (e.g. “How can we ensure researchers know when reporting guidelines should be used?”), and included context and elaboration.

The right hand column listed the ideas of how each influence could be addressed, coming from step 7 of the workshops. The table had 32 rows, one for each influence, and I gave each row a full page so there was a lot of white space to add ideas to, and so that rows could be presented individually on a computer screen (see Figure 8.2). I invited workshop participants to review and edit this table, thereby co-producing a list of influences and linked ideas.

8.2.2 Focus groups with external stakeholders

Focus groups are researcher-facilitated group discussions that use conversation as a form of data collection [161]. A key element of focus groups is interactions between participants as they agree, disagree, challenge, and ‘feed off’ of each other [161]. I chose focus groups because I expected this interaction to lead to more ideas being generated than if I interviewed participants in isolation. Focus groups are also a practical way to collect data from larger groups of people. This is in contrast to in-depth interviews which are more useful in eliciting rich detail about individuals’ perspectives [161].

Before describing my focus group methods in detail, I will briefly outline two key differences to the workshops. The workshops described above were part of the sessions described in the previous chapter, where I explained how I actively contributed because I felt my experience would be useful. However, by the time the focus groups with external stakeholders began I felt I had contributed everything I had to give, and my priority

for this study was for stakeholders to shape and build upon what EQUATOR and I had done. Therefore, whereas I actively contributed during the EQUATOR workshops, I tried to remain a passive facilitator during the focus groups. This was the first key difference. The second was that whereas in the workshops, EQUATOR staff and I spent many hours, across multiple sessions, working through influences one at a time, this was not feasible in the focus groups because it would have required multiple sessions and participants had limited time. Instead, focus group participants selected influences to discuss.

8.2.3 Sampling

To seek variation and ideas from a broad range of stakeholders, I invited a purposive sample including the developers of popular reporting guidelines, publishing professionals, and academics that have studied reporting guidelines. My invitation email is in Appendix M. I asked participants to extend the invite to others they felt would be appropriate. Because the Behaviour Change Wheel requires input from experts with insight into the intervention, I decided against recruiting naïve authors for this study (although I did elicit their opinions in a subsequent study — see chapter 10).

Following best-practice, I used information power [162] to guide my desired sample size. I decided to use information power above data saturation because, although the latter is commonly requested by editors and reviewers, many have argued against its (often poorly defined) use beyond its original application in grounded theory [164]. Instead, I followed Braun and Clarke's advice to use pragmatic 'rules of thumb' to anticipate a lower sample size that could potentially generate adequate data, and then make an in-situ decision about when to stop collecting data [164]. Information power is one such set of rules. Malterud et al. posit that the more relevant information a sample holds, the fewer participants are needed. They argue that sample size sufficiency depends on five factors: 1) whether the study's aim is narrow or broad, 2) whether samples are considered dense (they have a lot of relevant experience or knowledge of the phenomena) or sparse, 3) whether the study is well supported by theory, 4) the quality of dialogue, and 5) whether data will be compared between participants/groups. My aim was narrow

and well defined. My sample was dense in that participants knew a lot about how reporting guidelines are disseminated but also showed variance in terms of which guidelines they work on and which parts of the academic system they represented. I used the Behaviour Change Wheel as an applied theory. I used open questioning to encourage strong dialogue (I elaborate on this later), and I was not planning a cross-case analysis. All of these choices put my study towards the high-power end of Malterud's spectrum. Therefore, I deemed my information power sufficiently strong to justify initially recruiting 15-20 participants across 4-5 groups.

By monitoring the number of edits to the co-produced file I could be confident that my information power was good (i.e. that my sample was generating new ideas). I used the dialogue criteria from information power to decide when to stop recruiting. Once groups began to add fewer and fewer comments, I judged that the benefit of continued recruitment was insufficient given time constraints.

8.2.4 Materials

I used the ideas table to prompt discussion in the focus groups. However, because ideas generated by previous participants could bias or limit the creativity of current participants [165] I initially hid them by turning the text in the ideas column white before sharing the document. I would reveal the text only after participants had exhausted their own imagination (see Figure 8.2 for an example). All participants could edit this file to record their own ideas or elaborate on other people's ideas. At the end of each focus group, I then turned the ideas column white again, ready for the next group to continue the process. In this way, each group built upon the output of the previous groups without being biased by it.

8.2.5 Focus group sessions

I conducted focus groups between May-July 2022 online using Zoom. Before each focus group, I asked participants to spend some time thinking about influences and what needed to change. I did this because I wanted participants to get into the frame of mind and come to the focus group "armed" with influences they were ready to discuss.

The screenshot shows a Microsoft Word document titled "1600_02_08_22.docx - Saved". A table is displayed with two columns. The left column contains the question: "How can we ensure researchers understand the guidance as intended?". The right column contains a numbered list of three items. The text in the right column is white, indicating it is currently hidden.

How can we ensure researchers understand the guidance as intended?	<ol style="list-style-type: none"> 1. Researchers may not understand the guidance as developers intended, or at all 2. Researchers may not understand concepts, terms, or words within the guidance, or they may understand them differently to how the developers intended. 3. Researchers may have nowhere to turn for help should they not understand something
--	---

(A) Ideas initially hidden

The screenshot shows the same Microsoft Word document and table. The text in the right column is now black, indicating it has been revealed.

How can we ensure researchers understand the guidance as intended?	<ol style="list-style-type: none"> 1. Researchers may not understand the guidance as developers intended, or at all 2. Researchers may not understand concepts, terms, or words within the guidance, or they may understand them differently to how the developers intended. 3. Researchers may have nowhere to turn for help should they not understand something <ul style="list-style-type: none"> • Define key words or terms in a glossary or tooltips that are easy to find, use and are written in plain language. • Provide training courses • Allow researchers to give user feedback and ask questions • Update guidance in response to user feedback to clarify confusing terms • Provide examples (real or generated) of good and bad reporting as researchers report that examples help them understand what the guidance means • "Reporting champions" within institutions that can help researchers • Developers could consult different user groups when creating guidance to make sure all users understand it as intended • Consider replacing confusing words or concepts e.g., TIDIER removed the word "context" because users struggled to understand it.
--	---

(A) Ideas revealed

FIGURE 8.2: An example row in the Ideas Table that was co-edited by participants. The left hand column contained influences that needed to be addressed for our target behaviour to occur. Existing ideas to address each influence were initially hidden (by turning the text white), and only made visible once participants had discussed their own ideas.

I was also interested to see whether participants would contribute influences that I did not identify in chapters 3 - 5.

Each focus group lasted 2 hours. Following standard practice, I began by introducing myself and the project in a way that I hoped would help participants relax and to think open-mindedly, not defensively [166]. I explained where the list of influences had come from, and that the influences were in reference to reporting guidelines *in general*, and not necessarily a comment on any guideline in particular. I encouraged participants to think beyond the guideline documents themselves, and to consider all stakeholders and resources involved. I explained the goal was to brainstorm as many ideas as possible, and not worry about whether ideas were good or bad [165].

It was not possible for a single focus group to cover all rows within a reasonable amount of time, so I allowed participants to select which influences they wanted to discuss. I did this by giving them a few minutes to read through the influences in the left hand column, raise any additional influences they felt were missing, and mark those that they wanted to talk about. I occasionally selected items to discuss myself, either because they had been neglected by previous groups or because I expected participants to have insight into them.

For each influence discussed, I would explain it and allow participants to ask questions. Inspired by the Think/Pair/Share teaching method to encourage engagement within classrooms [167], I then asked participants to spend a minute reflecting on the required change and brainstorming solutions on their own before discussing them as a group. I encouraged this solo reflection because I wanted all participants to engage with the problem.

To facilitate discussion, I asked open ended questions, often drawing on intervention functions from the BCW by asking questions like “how could this be easier to do?” or “how could we change how people feel about this?”. I did this when participants ran out of ideas, or when they got fixated on a particular type of intervention, in which case I would reassure participants that their fixated solution was already documented and that it would be useful to think of alternatives.

Once participants had exhausted their own ideas, I revealed the ideas identified by previous groups by changing the colour of text from white to black. Participants could then edit and extend the text until it reflected all of their thoughts too. Ideas were never removed from the document, but participants could add concerns or disagreements if they wanted to. Editing the file in this way allowed participants to document their thoughts in their own words.

After each focus group I made notes on how the session went and reflected on what I could have done differently. I made a copy of the ideas document and then turned the text in the ideas column white again, ready for the next group. Taking copies after each group created a paper trail of how the document had evolved after each session. I counted the number of additions so that I could monitor how many new ideas had been added and, therefore, whether I could stop data collection.

8.2.6 Data processing and analysis

I used qualitative description for my analysis [168,169], which involved aggregating and summarising ideas. I imported the final ideas document into NVivo [170] and applied descriptive codes to ideas. If a sentence contained multiple ideas, I coded each idea separately. I also coded the influences and stakeholders that were related to each idea. I did not interpret data as doing so would erase the views captured during co-production.

In the next chapter I describe how I used the Behaviour Change Wheel to label ideas according to their intervention function. However, I wanted the results of this chapter to be accessible and useful to the reporting guideline community (I have since shared this chapter with a few guideline development groups), and so I decided to group ideas inductively in ways that felt cohesive and made the results easy for my intended audience to understand and act upon. For example, I aggregated “ask authors to cite reporting guidelines” and “display citation metrics on reporting guideline resources” into a group about “Citations”, even though they target different influences (discoverability and perceived trustworthiness) and employ different intervention functions (education and persuasion). I labelled each group with the influence(s) it addressed, and the stakeholders it involved.

I discussed and refined my coding, aggregating and summarising with one of my supervisors, Jennifer de Beyer. I sent the aggregated, summarised ideas to focus group participants and EQUATOR staff members, inviting them to check it reflected their ideas faithfully.

8.2.7 Reflexivity & Trust

In chapter 7 I described my active role within my workshops with EQUATOR and I argued that my subjectivity was an asset within the workshops. In contrast, I tried to remain objective when running focus groups with external stakeholders, in order to capture the perspectives of participants without influencing them. My research paradigm for the focus groups was post-positivist, in that I considered that ideas were “out there”, but that differences in context, experience, and opinion would affect what I (and participants) observed, understood, and concluded.

I wanted to ensure my results could be trusted as an account of participants’ views. Lincoln and Guba [156] argue that for a study to be trustworthy, the researcher must show that the findings are credible (‘true’), transferable (applicable to other contexts), dependable (consistent and repeatable), and confirmable (shaped by participants, not by the researcher’s bias or motivation). Lincoln and Guba propose a number of techniques to achieve these criteria, and I describe the techniques I used in Table 8.1.

TABLE 8.1: Techniques for establishing trustworthiness. Based on Lincoln and Guba’s Evaluative Criteria [156]

TECHNIQUE	IMPLEMENTATION
Techniques for establishing credibility	

TECHNIQUE	IMPLEMENTATION
<i>Member-checking</i>	<p>Lincoln and Guba argue that member checking is the most important way to establish validity of an account [156]. Accordingly, I invited participants to comment on my synthesised results, asking for feedback on the structure of categories, my interpretation of their data, and my findings and conclusions. I also invited participants to comment on the product of my data analysis in the form of itemized information and condensed notes.</p>
<i>Peer debriefing</i>	<p>Throughout the design, data collection, analysis and reporting, Charlotte Albury acted as a disinterested peer. By questioning my reasoning and exploring my assumptions, she helped me become aware of biases, perspectives I was taking for granted, and assumptions I was making. Jennifer de Beyer acted as a disinterested peer during data analysis.</p>
Techniques for establishing transferability	
<i>Thick description</i>	<p>I aspired to report my results with context by indicating when ideas were common or rare, and who they originated from when I felt this was particularly relevant. I reported disagreements, provided quotes, and relationships between ideas.</p>
Techniques for establishing confirmability	

TECHNIQUE	IMPLEMENTATION
<i>Audit trail</i>	I referred to audio recordings of the focus groups whenever I needed to clarify parts of the document. I kept versions of raw data collected from all stages. I made a note of my own ideas before commencing data collection, documented all stages of the workshops I held internally with the EQUATOR Network, and kept copies of the co-produced file after every focus group. I kept a copy of my coding in NVivo, and versions of the unitized information and summaries that I sent to participants before and after member checking. This audit trail meant that I could be certain of which stages of research ideas originated from.
<i>Reflexivity</i>	I kept personal notes throughout planning, data collection, analysis and reporting, in an attempt to remain aware of my own perspectives and positions, and how they may influence my research.

8.2.8 Ethics & Data Management

The study was approved by the Medical Sciences Interdivisional Research Ethics Committee (R80414/RE001). Participants gave informed consent by completing an online form [171] having read the participant information sheet (see Appendix N). Participant's edits to the co-produced file were anonymous. I recorded the audio of focus groups so that I could refer to them during analysis. All data and recordings were kept on secure university storage.

8.2.9 Reporting

I used SRQR [30] when outlining this chapter, and again to check my reporting during revision (see appendix L).

8.3 Results

8.3.1 Units of study

I held 7 focus groups involving 16 participants in total. Participants included guideline developers ($n=11$), publishing professionals ($n=3$), and academics that study reporting guidelines ($n=2$). Although I had intended to include 4 or 5 participants per focus group, in practice it was difficult to coordinate participants across time zones, and so sessions only had 2 or 3 participants.

Of the 23 invitations that I sent in total, 7 received no response. Of the 15 guideline groups I invited, 5 took part, 4 wanted to but were unable to coordinate a time, 5 did not respond, and 1 group declined; they felt their guideline did not need updating because it was highly cited. Because I invited prospective participants to share the invitation I have no way of knowing my exact recruitment rate.

Before the focus groups, my workshops with EQUATOR had generated a list of 24 ideas, which formed the initial “ideas table” presented to the first focus group. Focus group participants then extended this list to include 128 ideas to address 32 influences. Focus group participants did not identify any new influences. The seventh focus group added just two new ideas and extra context to an another existing two, and so I decided to stop data collection. I grouped ideas into 28 broader ideas, which I categorised according to whether they could be considered before developing a reporting guideline, when developing a reporting guideline, when writing guidance down and creating resources, when disseminating resources, or on an ongoing basis. Participants identified 10 stakeholders that could enact these ideas: funders, ethics committees, institutions, publishers, equator network, guideline developers, registries, preprint servers, conference organisers, and societies.

8.3.2 Synthesis and summary

Here I describe the aggregated ideas originating from the focus groups in this chapter and step 7 of the workshops from the previous chapter, as described at the beginning of my methods section. Although I described the people taking part in focus groups as *participants*, I did not use this term in the previous chapter because I took part in the

workshops with my colleagues and so the line between researcher and participant was blurred. Because focus group participants were building upon our workshop outputs, the results presented here reflect the ideas of focus group participants *and* workshop members. Consequently, I use the term *stakeholder* in this section to reflect that the results presented originate from both focus group participants *and* workshop members, which included myself.

The broad ideas are summarised in Table 8.2 and described below. I report them fully in Appendix O, where they are labelled with the sub-ideas they contain, the influence(s) they address (from Appendix I) and stakeholders involved.

TABLE 8.2: Broad ideas and the influences they target, categorised by when they should be considered. Influences were derived and consolidated in previous chapters, and are reported fully in Appendix I)

IDEAS	INFLUENCES
Before developing a reporting guideline	
Create reporting guidance for early stages of research	Researchers may not have tools for the job at hand Researchers may not encounter reporting guidelines early enough to act on them
Avoid confusing authors with too many reporting guidelines	Researchers may not know what reporting guidelines exist Researchers may struggle to reconcile multiple sets of guidance Researchers may not know what reporting guideline is their best fit
When developing a reporting guideline	
Avoid prescribing structure	Researchers may struggle to reconcile multiple sets of guidance Researchers may struggle to keep writing concise

IDEAS	INFLUENCES
Keep reporting guidelines agnostic to design choices	<p>Researchers may feel restricted if reporting guidelines prescribe design</p> <p>Researchers may feel afraid to report transparently</p> <p>Researchers may expect the costs to outweigh benefits</p> <p>Researchers may not know what reporting guidelines are</p>
Describe reporting items fully	<p>Researchers may not know whether a reporting guideline applies to them</p> <p>Researchers may not know how to report an item in practice</p> <p>Researchers may not know how to do an item</p> <p>Researchers may not know what to write when they cannot report an item</p> <p>Researchers may struggle to keep writing concise</p> <p>Researchers may not know why items are important</p> <p>Researchers may not care about the benefits of using a reporting guideline</p> <p>Researchers may be asked to remove reporting guideline content</p> <p>Researchers have limited time</p> <p>Researchers may struggle to keep writing concise</p>

IDEAS	INFLUENCES
Describe each reporting guideline fully	<p>Researchers may not know whether a reporting guideline applies to them</p> <p>Researchers may not know what reporting guidelines exist</p> <p>Researchers may not know what reporting guideline is their best fit</p> <p>Researchers may not know when reporting guidelines should be used</p> <p>Researchers may feel that checking reporting is someone else's job.</p> <p>Researchers have limited time</p> <p>Researchers may feel patronized</p>
Keep guidance short	<p>Researchers have limited time</p> <p>Researchers may expect the costs to outweigh benefits</p>

When writing guidance down and creating resources

Make resources ready-to-use	<p>Reporting guideline resources may not be in usable formats</p> <p>Researchers have limited time</p>
Make reporting guidelines easy to understand	<p>Researchers may misunderstand</p> <p>Researchers may not understand the language</p>
Use persuasive language and design	<p>Researchers may feel afraid to report transparently</p> <p>Researchers may feel patronized</p> <p>Researchers may not believe stated benefits</p> <p>Researchers may feel that checking reporting is someone else's job.</p>

IDEAS	INFLUENCES
Create additional tools	<p>Researchers may not have tools for the job at hand</p> <p>Researchers may not encounter reporting guidelines early enough to act on them</p>
Make resources easy to discover and find	<p>Researchers may not know what reporting guidelines exist</p> <p>Researchers may not know what resources exist for a reporting guideline</p> <p>Guidance may be difficult to find</p> <p>Researchers may not know what reporting guideline is their best fit</p>
Make information digestible	<p>Researchers have limited time</p> <p>Researchers may expect the costs to outweigh benefits</p> <p>Guidance may be difficult to find</p>
When disseminating resources	
Describe reporting guidelines where they are encountered	<p>Researchers may not know what reporting guidelines are</p> <p>Researchers may not know when reporting guidelines should be used</p> <p>Researchers may not know what benefits to expect</p> <p>Researchers may not know why items are important</p> <p>Researchers may feel that checking reporting is someone else's job.</p>
Make resources accessible	Reporting guidelines may be difficult to access
Show and encourage citations	<p>Researchers may not know what reporting guidelines exist</p> <p>Researchers may not believe stated benefits</p>

IDEAS	INFLUENCES
Provide testimonials	<p>Researchers may not know what benefits to expect</p> <p>Researchers may not believe stated benefits</p> <p>Researchers may expect the costs to outweigh benefits</p> <p>Researchers may not care about the benefits of using a reporting guideline</p> <p>Researchers may feel afraid to report transparently</p>
On an ongoing basis	
Budget for reporting	<p>Researchers have limited time</p> <p>Researchers may not consider writing as reporting</p>
Create rewards	<p>Researchers may not care about the benefits of using a reporting guideline</p>
Create discussion spaces	<p>Researchers may misunderstand</p> <p>Researchers may feel patronized</p> <p>Researchers may not believe stated benefits</p> <p>Researchers may not know how to report an item in practice</p> <p>Researchers may not know how to do an item</p>
Create ways to catch authors earlier	<p>Researchers may forget to use reporting guidelines at earlier research stages</p> <p>Researchers may not encounter reporting guidelines early enough to act on them</p> <p>Researchers have limited time</p> <p>Researchers may not know when reporting guidelines should be used</p>

IDEAS	INFLUENCES
Endorse and enforce reporting guidelines	Researchers may not know what reporting guidelines exist Researchers may expect the costs to outweigh benefits
Evidence the benefits	Researchers may not believe stated benefits
Make reporting guidelines appear as a priority	Researchers may not believe stated benefits Researchers may not care about the benefits of using a reporting guideline
Promote reporting guidelines	Researchers may not know what reporting guidelines are Researchers may not know what reporting guidelines exist
Install reporting champions	Researchers may not know what reporting guidelines are Researchers may not know what benefits to expect Researchers may misunderstand Researchers may not know when reporting guidelines should be used Researchers may not know why items are important

IDEAS	INFLUENCES
Provide additional teaching	Researchers may not consider writing as reporting Researchers may misunderstand Researchers may not know why items are important Researchers may not know how to report an item in practice Researchers may not encounter reporting guidelines early enough to act on them Researchers may not care about the benefits of using a reporting guideline
Make updating guidelines easier	Reporting guidelines can become outdated Researchers may misunderstand

In the summary below I have occasionally mentioned which stakeholder group an idea came from, but only when I felt like it added useful context. I have chosen not to label the rest for three reasons. Firstly, stakeholders were all editing the same file, so some ideas would be revisited multiple times by different stakeholders who would build upon the thoughts of previous contributors. Consequently, it wasn't possible to attribute an idea to a single stakeholder, as it may have been the product of multiple stakeholders. Secondly, just because a stakeholder didn't edit an idea in the document didn't mean that the idea hadn't also occurred to them. Hence allocating an idea to a stakeholder just because they were the one that wrote it down may be misleading. Thirdly, I didn't consider labelling the *origin* of an idea to be useful because I didn't see it as an indication of that idea's *quality*; to the contrary, I had explicitly encouraged participants not to worry about whether an idea was "good".

Before developing a reporting guideline

Create reporting guidance for protocols and applications Stakeholders suggested “developing [reporting guidance] for protocols”, funding, and ethics applications to encourage authors to consult reporting guidance earlier in their work when they are more likely to have the time, motivation, and ability to reflect and act on it.

Avoid confusing authors with too many reporting guidelines “We need fewer, better, reporting guidelines” wrote one stakeholder, when discussing how authors may struggle to identify which reporting guidelines apply to their work.

Acknowledging that reporting guideline developers may duplicate each other’s work unwittingly, stakeholders wrote that developers should consult EQUATOR’s register of reporting guidelines under development before creating a new guideline. Stakeholders noted that “EQUATOR cannot prevent guideline developers from creating duplicate guidelines” but could “improve the registration process for reporting guidelines that are under development”, better “highlight [reporting guidelines] that are under development in the main search results” and could “create options and instructions to encourage developers to extend existing guidance instead of duplicating. This could go in the new guidance for guideline developers”.

Stakeholders had ideas of how this “extending” could be done. They suggested that developers could tailor existing guidance to a particular niche by making different “versions” or “extensions” of existing guidance “e.g., STROBE split into STROBE Cohort, Case-control etc”. If only a few items need to be edited or added, stakeholders suggested creating “modular” guidance instead of duplicating an entire guideline. Stakeholders spoke of modules in two different ways. Firstly, new reporting guidelines can be created to substitute for particular items in existing reporting guidelines. Stakeholders named TIDIER as an example of this strategy, which can be substituted for item 5 in CONSORT. Stakeholders identified a second kind of modularity in the JARS guidelines, where a general reporting guideline covering all quantitative psychology research can be mixed-and-matched with modules for specific designs (non-experimental and experimental designs, with or without random assignment, or special modules for longitudinal, n-of-1, or replication studies [172]).

When discussing how modules or extensions could be “harmonized” so that they “speak to another”, participants suggested resources could have compatible structure, and should not “use different wording for what is essentially the same item”. Participant’s recommendation to “use similar terminology [across] related guidelines” extended to the title; STROBE-nut, STROBE-ME and STROBE-RDS are more readily identifiable as STROBE extensions than STREGA, ROSES-I, or STROME-ID are. Naming can also indicate when a guideline has been revised: “ARRIVE 2.0 is recognisable as a replacement of ARRIVE”, whereas “TIDIER Placebo appears to be an extension but should be called TIDIER 2.0”.

When developing a reporting guideline

Avoid prescribing structure Stakeholders suggested that “[reporting] guidelines should avoid prescribing structure as [it] may clash with journal guidelines”. A good example of a structure clash is abstract subheadings; most journals have strict requirements about whether abstracts should have subheadings and what those headings should be. Authors can struggle if reporting guidelines suggest an alternative structure.

Keep reporting guidelines agnostic to design choices Stakeholders discussed how designing research is a separate task to reporting it, and that many authors will only encounter reporting guidelines *after* the manuscript has been written, at which point design advice is less useful.

Nevertheless, stakeholders cited multiple reasons for including design opinions in reporting guidance. Some suggested it should be included so that authors can “learn for next time”, and others wrote that consequences of design choices justify why an item was important to report, but acknowledged that justifying items in this way can be problematic if developers do not consider all contexts or study types in which their guidance may be used.

Some participants argued that design advice could be removed from reporting guidelines entirely. This was proposed as a solution when considering how authors may feel

afraid to report transparently if what they did goes against the design recommendation, or may feel restricted if forced to use a reporting guideline that prescribes design choices.

Instead, developers could encourage authors to “explain reasons for methods choices, [which may] be legitimate”, noting that “the consequence of not choosing one [design] option over another, even if the choice is a rarely used option, may not have major consequence on the results of the study”. One stakeholder wrote that authors may feel reassured if told that “editors and peer-reviewers may not judge as harshly when they understand the rationale for the choice”. Stakeholders wrote that developers should “encourage transparent reporting over and above good design”, and that authors could be encouraged to “describe what they did in plain language to make it clear - if what they did doesn't quite fit with standard terminology (e.g., if they didn't really do theoretical sampling or aren't sure if they did theoretical sampling) then just describe what they did, how they made sampling decisions”.

Discussions about removing design advice also arose when considering how including it elongates guidance, potentially deterring authors from reading it. To solve this, stakeholders suggested reporting guidelines could link to design resources “elsewhere”. For example, if reporting guidance were presented on a website, authors could be given options to “display or hide” design advice by choice, or depending on whether the author was “designing [research], applying for funding, or drafting” a manuscript.

Describe each reporting guideline fully Stakeholders identified that many influences could be addressed by editing the introductory text that authors may read before using a reporting guideline or checklists. This text may appear on journal guideline pages, the EQUATOR website, or on the guideline or checklist itself.

For example, stakeholders suggested that reporting guidelines and resources should “clearly state what kind of research the guidance applies to” when considering how authors may struggle to identify which reporting guideline they should use. Additionally, stakeholders suggested guidelines could “point researchers to other guidelines if more

relevant, perhaps using ‘if, then’ rules e.g., ‘if you did X then use Y instead of this guideline’ ”. If no better guideline exists for a particular type of study, stakeholders suggested to “warn the user that they can still use this guideline but that they may need to ignore certain reporting items e.g., ‘This guideline is for studies that did X. You can still use it if you did Y, but you will need to ignore items 4, 6, and 8-10’ ”.

To ensure that a reporting guideline for writing manuscripts can also be used for writing protocols, developers could specify “which items need to be considered at protocol/planning stage, and which at results reporting stage.”

Stakeholders warned that developers should be mindful when using words like *standard* instead of *guideline* when introducing a reporting guideline, as these words may influence how prescriptive the user expects the resource to be. In choosing their wording, stakeholders suggested developers should be honest and clear about a guideline’s “aim”. Developers could also “be clear about what reporting guidelines *are not*”. For instance, “when they are not design guidelines or critical appraisal tools” this could be specified and authors could be directed to other tools instead.

Noting that it might not be obvious how or when to use guidance, stakeholders suggested being explicit about this. For example, “tell authors that they don’t need to fill out templates [or checklists] sequentially but can use an order that matches their workflow or decision making. Put this instruction on the template, checklists, and other tools. Example from PRISMA – population and subgroup items are separated in the checklist but go together when thinking/making decisions.”. Similarly, authors could benefit from explicit suggestions of how to use reporting guidelines as a team, perhaps by asking “their co-researchers to check their reporting”. Finally, stakeholders suggested telling authors how much time the guideline is expected to take them to read and use, why it can be trusted, and where they can read about its development.

Describe reporting items fully Many influences prompted stakeholders to suggest that specific content should be included for every reporting item. For example, stakeholders noted that authors need to know what to write, and that a brief description could go in checklists and a longer description in the full guidance. Stakeholders

suggested this description should include what to write if they didn't or couldn't do something to make it "easy for researchers to report things they are embarrassed about". In these instances, developers could suggest authors "explain reasons" for their choices or "consider the item in their discussion section as a possible limitation" if necessary. Stakeholders also wrote that developers could suggest what to write when an item doesn't apply.

Stakeholders suggested explaining "why [an item] is important and who it is important to" as this isn't always obvious to authors. To make guidelines faster to use, stakeholders suggested indicating which items are most important, perhaps "prioritise[ing] certain items as essential vs. recommended, like ARRIVE 2.0 essential 10" and indicating conditions that make items less or more important, including circumstances that make the item non-applicable.

To help authors who want to keep writing concise or need to reduce word counts, stakeholders suggested advising "when an item can be put into a table, figure, box, supplement, or appendix etc" and to "explain the pros and cons of different options e.g., whether content will be peer reviewed or indexed" by search tools.

Many stakeholders suggested including examples to help authors understand and apply guidance. These examples could include both "good and bad reporting" with an explanation of "how [bad reporting] could be improved". To help authors who are afraid to transparently report limitations, stakeholders suggested including examples of imperfect research that is perfectly-reported. Examples could also include "reporting in different contexts", from different disciplines, "in multiple languages", and concise reporting e.g., reporting "TIDIER items nicely in a table".

Because examples from the published literature can be "difficult to [...] find and list", stakeholders suggested that examples could be "real or generated" by developers. Examples should be "easy to find" from within the guidance resources. Others suggested a searchable "bank" of examples, or ways to showcase "exemplar papers (e.g., badges)".

Finally, stakeholders discussed the pros and cons of including item-specific design advice, procedural instructions, and appraisal advice, and the different ways of doing it

(see *Keep reporting guidelines agnostic to design choices*).

Keep guidance short “Make guidance shorter” wrote one stakeholder, and “Make checklists shorter”, wrote another, when considering how “guideline length [...] may put researchers off” as “long guidelines appear more complex and time consuming” and can challenge word limits.

One stakeholder posited that guidelines could be shorter if developers were “realistic about what they ask for”. Another challenged developers to “try using a guideline from start to finish and see how the manuscript ends up”. Others wrote that guidelines could be shortened by linking out to content that wasn’t directly related to reporting, like design or appraisal advice (see *Keep reporting guidelines agnostic to design choices*). Guidelines that are “very long or [have] lots of optional items” could be split into multiple versions (see STROBE as an example in *Avoid confusing authors with too many reporting guidelines*).

Others suggested presenting guidance online with non-essential content collapsed so that guidance appeared short but authors could choose to “display or hide” content. Stakeholders noted that “itemisation can make guidance more digestible, but can also make it harder to get the bigger picture” and makes text appear longer.

When writing guidance down and creating resources

Make resources ready-to-use Stakeholders suggested that resources should be in ready-to-use formats. For instance, “checklists should be editable (not PDFs)”.

Make reporting guidelines easy to understand “Make guidance easier to understand” was written as a solution to help authors who misinterpret, or can’t understand, guidance. Stakeholders suggested developers could “make guidance as ‘plain language’ as possible” or “create plain language versions of existing guidance”, whilst being “mindful of language that may appear patronizing”.

Stakeholders suggested defining “key words or phrases in a glossary or tooltips”, using consistent terms across related resources, translating guidelines and examples, and

ensuring these translations are easy to find by making them “searchable”, “linked properly” to each other, and “more evident on [the] EQUATOR site”.

Stakeholders recognised a need to collect and respond to feedback from “international researchers” representative of the user base across disciplines and institutions, and that this could be sought when developing guidance to “test” it, but also on an ongoing basis to “continually revisit the items in the guidelines that may be confusing or difficult to implement”. As an example, one guideline developer contributed that collecting feedback had led them to remove “the word ‘context’ because users struggled to understand it”.

Use persuasive language and design When discussing how authors feel about reporting guidelines and checklists, stakeholders reflected on how the wording and presentation of resources can influence authors’ perceptions of it.

Stakeholders noted that a poorly formatted checklist, lacking “visual appeal/graphic design”, can “appear outdated” or “larger than it is”. Stakeholders suggested that developers could use design to foster “feelings of simplicity”, and “engage graphic designers” if necessary.

Reassurance seemed especially important to stakeholders when considering how to motivate authors to transparently report items that aren’t perfect. Stakeholders suggested that developers could use language and tone of voice to cultivate “a feeling of confidence, not judgement”, perhaps reminding “researchers that all research has limitations”. Stakeholders also wrote that reassurance could also be given by including notes aimed at reviewers, and cited JARS as an example “which explains to reviewers why an author may make certain choices”, thereby educating reviewers whilst reassuring authors that they won’t be penalised for transparency. Stakeholders suggested rewording text and tone of voice to make reporting guidelines and checklists “appear less like ‘red tape’”, and to reassure authors that “reporting guidelines are just that: guidelines!”.

Create additional tools Stakeholders noted that checklists may be easier to use if they are editable (see *Make resources ready-to-use*), and if authors could complete them with

“the relevant text, rather than [the] page/section number” which can be frustrating to keep updated. Although one stakeholder, a publisher, noted that checklists completed with text are difficult to double check, and therefore it would be most useful to include the text *and* the page number.

To aid writing, stakeholders suggested “templates for drafting” manuscripts, interactive forms and “writing tools (e.g., COBWEB)”, “tools for creating figures and tables like PRISMA’s flowchart generator”, and tools for generating text, like TIDIER’s tool for generating intervention description [173]. However, one stakeholder warned that structured writing tools “provide opportunities for inclusion [of reporting items] but there is always the risk that they exclude more [important items] that are outside the boundaries of the template”.

To encourage authors to consider reporting guidance earlier in their research, stakeholders considered to-do lists with items “in the order they are done”, or “embed[ing] items into data collection tools” like software for systematic reviewers.

To help reviewers check reporting, stakeholders suggested creating “tools for co-researchers to check each others’ work”, creating extra guidance for peer reviewers, providing text that can be pasted in reviewer feedback forms to “request additional information” for poorly reported items, or even building a “reviewer tool that generates a report”. Stakeholders noted that tools should be presented in ways that “better differentiate” how and when they should be used “e.g., resources for writing vs. checking vs. reviewing”.

Make resources easy to discover and find Stakeholders had ideas for how to help authors discover and find resources. Stakeholders wrote about hyperlinks being an important way for authors to discover related guidelines and tools. These could be “horizontal links between related guidelines” and between guideline “documents and tools” like checklists. All resources should “link to one another”, ideally “item by item” so that checklist items could link directly to the relevant section of full guidance. Stakeholders noted that hyperlinks can become out of date, and so stakeholders should “fix broken links”.

Stakeholders also wrote that resources could be hosted on “a convenient place, such as a unified website”. EQUATOR’s website is one such place, and participants suggested making it “easier to navigate” and its “search tool more prominent and easier to use”. Although some authors will use such search tools, stakeholders recognised that those who browse may benefit from “curated” collections of reporting guidelines or a “manageable list of related, commonly used guidelines”. As an example, an author writing a review article may be helped by a page listing “PRISMA, MOOSE, ENTREQ, PRISMA-ScR etc.” and that “this page could be kept up to date as guidelines are revised”. Another stakeholder (a publisher) warned, however, that journals may not want to link to these pages if the collection has “a much broader scope” and includes guidelines that the journal doesn’t endorse.

When discussing how to help authors who are less familiar with study designs, one stakeholder suggested creating “tools to help researchers identify study designs (e.g., a questionnaire)”, and another, already familiar with such a tool previously developed by EQUATOR suggested it should “use plain language”.

Make information digestible Stakeholders acknowledged that authors’ needs may differ between tasks (e.g., drafting an article vs demonstrating compliance), and that authors may use guidance in different ways; some will read the whole thing from start to finish, whilst others will dip in and out as-and-when they need. Consequently, stakeholders wrote that “having different options available that meet the needs of different users is vital” and that authors should be able to consult guidance in ways that “work for them”.

One suggested way of doing this may be to “structure guidance using navigation menus and subheadings” so that “it is easy to find the information you need”, making reporting guidelines faster to use and less overwhelming. Another noted that checklists can also be designed, citing TIDIER as “a nice example” that has “integrated the intervention and placebo into one table” with the active intervention and placebo in adjacent columns.

Dynamically hiding and showing content was floated here again (as it was in *Keep guidance short*), with one stakeholder suggesting that users could “filter out” irrelevant content, to only see instructions for their task (e.g., planning, writing, reviewing) or specific to their study. This could be done with a “decision tree” or “branching questions” to determine specific features of the study (“e.g., a systematic review *with* network meta-analysis of individual participant data”). Answers to these questions could then be used to “modify” items to create “personalised guidelines”, or to generate a “customised reporting checklist” that includes all “main and relevant extension items”.

Dynamic content was also seen as a favourable way to embed guideline extensions, with the aim of making them easier to discover without overwhelming the author. For example, noting that “some guidelines ‘fit together’...e.g., PRISMA and PRISMA-Abstracts”, stakeholders wrote that PRISMA-Abstracts could be “embedded” as collapsed content that interested authors could expand.

When disseminating resources

Describe reporting guidelines where they are encountered When first introducing reporting guidelines stakeholders suggested telling authors what reporting guidelines are, “when and how best to use” them, and what benefits to expect. This information could go wherever authors are advised to use reporting guidelines (like journal instruction pages, registries), EQUATOR’s website, social media campaigns, at the start of the guidelines, and could go at the beginning of checklists too “in case people don’t read the whole [guideline] paper”. This introductory text could be “short, sweet, and to the point”. Benefits could be even more prominent by putting them “in a box, or [by using] font or positioning”.

Make resources accessible Stakeholders wrote “Ensure guidance is open access” so that all authors can access it. Stakeholders also noted that if guidance is published under a permissive license, then others can reuse the content to extend the guidance or build new tools.

Show and encourage citations Displaying citation counts on the EQUATOR Network website (or other websites where authors search for reporting guidelines) was described as a way to “provide social proof” and convince authors that guidelines are credible.

To generate these citations, stakeholders suggested explicitly asking “researchers to cite the guideline they used”. Stakeholders wrote that if an author cites a guideline they have used, then readers may discover the guideline from that authors’ article.

Provide testimonials Stakeholders suggested providing “testimonials” as a way to tackle a few different barriers using education and persuasion. Stakeholders suggested providing “quotes from authors/researchers who felt that reporting guidelines helped their work and who have had positive experiences” such as making “writing easier” or helping “with co-authorship communications”. Stakeholders proposed that testimonials could bring benefits to life, thereby making them more believable.

To make authors care more about research waste caused by poor reporting, stakeholders suggested testimonials from “research consumers for whom an item is important”, or quotes that illustrate “how detrimental poor reporting is for end users”.

Stakeholders wrote that testimonials from decision makers (like editors, reviewers, and grant-givers) could communicate their “preference for transparent reporting” and convince authors that reporting will be checked. If these testimonials conveyed that transparency is valued above perfectionism, participants suggested this could reassure authors. Stakeholders also suggested collecting positive testimonials from such “nervous researchers”.

Finally, stakeholders suggested collecting testimonials from researchers “with a range of experience”, including “experienced researchers who have benefited by changing their practices”. Diverse case studies would help engage a diverse user base, and challenge assumptions that reporting guidelines are too patronizing for experienced researchers or too complicated for inexperienced ones.

On an ongoing basis

Budget for reporting Stakeholders noted that “researchers need budget to allocate time to writing” and that “funders could encourage proper financial/time budgeting for writing”, as could research supervisors.

Create rewards Stakeholders suggested “offer[ing] some sort of tangible reward/benefit” to motivate guideline use, creating new rewards when necessary. Ideas included “publishers offering a fast-track review/discount”, “badges on published articles” or platforms “like publons”, or “a certificate after completing training”.

Create discussion spaces Multiple barriers lead stakeholders to suggest “create[ing] spaces for researchers to connect with other researchers to celebrate and share experiences”. These spaces could include “forums, meetings, tea clubs, [and] clinics both in real life and virtual”. Such spaces could help authors solicit help and could act as social proof, as seeing “others using and talking about the guidance” may be motivational.

Online discussion spaces were also considered a useful way to gather feedback from users directly (by asking for it) and indirectly (by monitoring discussions). Stakeholders wrote that feedback channels “could be useful to guideline developers”, and may also “cultivate a feeling of community ownership” by “communicating an invitational attitude”, thereby making guidelines appear less bureaucratic.

Create ways to catch authors earlier Stakeholders thought of ways to “try and shift the time at which researchers discover or use guidelines”, hypothesising that “it’s more likely that guidelines will save them time” if used earlier or “at the right time” and “not just upon submission”.

Most simply, stakeholders suggested “telling” or “encouraging” authors to use reporting guidelines for planning or drafting research (and not just for demonstrating compliance upon submission). Building upon this, stakeholders suggested organising the EQUATOR Network website to make it obvious which stages of work resources can be used for.

Stakeholders suggested “including [reporting guidelines] in the university teaching and training curriculum and text books” so that students learn about them before running or writing up their first study.

Stakeholders suggested funders, ethics committees, and writing training programmes could advertise reporting guidance for early research outputs like funding applications and protocols (as previously described in *Create reporting guidance for early stages of research*).

When considering whether authors may need reminders to use a reporting guideline for their next study, stakeholders suggested publishers and EQUATOR could use “email reminders” or strategies used by e-commerce sites “like when you buy something from an online business... then they work hard to gain your custom again”.

Endorse and enforce reporting guidelines Stakeholders suggested “encouraging more journals to endorse guidelines” and drew a distinction between endorsing reporting guidelines and promoting them on a website, social media, or email (see *Promote reporting guidelines*). Endorsement was described as a long term commitment to recommend or encourage guideline use, requiring buy-in from organisational leaders, and possibly changes to policies, instructions, infrastructure, and workflows. Promotion, conversely, was described as ephemeral and does not require organisational changes.

Stakeholders drew another distinction between endorsement and enforcement, whereby enforcement meant reporting guidelines are “a requirement” or “condition”. Enforcement was further divided into enforcing checklist completion or, noting that checklists may not always accurately reflect manuscript content, checking text for adherence to guidance.

When considering *who* could enforce guidelines, stakeholders noted that reporting guidelines could be “a requirement for publication”, “for registration (where applicable) (e.g., clinical trial registries, PROSPERO)”, or when submitting conference abstracts. “Ethics committees and funding organisations” could require applicants adhere to guidelines by, for example, completing “SPIRIT for clinical trial submissions”, or could

ask applicants to declare they will use a guideline when writing their results. To facilitate enforcement, stakeholders suggested that the software academics use to provide funders with updates could ask for completed checklists. One stakeholder suggested that reporting guideline adherence should be a condition of university employment and that a “digital dashboard [may] help audit[ing] and monitoring”. Noting that enforcement requires resources, stakeholders suggested enforcement policies could “focus on main [reporting guidelines]”.

Evidence the benefits Stakeholders suggested that benefits may be more believable if they were evidenced. This could be “evidence that [reporting guidelines] improve the completeness and transparency of the output”. For quantifiable personal benefits, the suggestion was to collect and report data on “acceptance rates, publishing speed, writing speed”. One stakeholder posited that “more transparent reporting / structured reporting may lead to faster editorial processes as it becomes easy for peer reviewers and editors to review papers about their study” and another suggested that providing “statistics about processing times of articles that follow / don’t follow reporting standards” would help evidence this claim and “emphasise to researchers that clear reporting will minimise the number of times others contact them for clarification”. However, some stakeholders were sceptical whether data on acceptance rates would show any benefit at all: “Likelihood of being accepted might not be heavily influenced – bad research well reported would still be rejected”.

For personal benefits that are harder to quantify (like increased confidence), stakeholders suggested providing case studies or testimonials (see *Provide testimonials*).

Make reporting guidelines appear as a priority A few stakeholders suggested making reporting guidelines more prominent within journals’ submission workflows to make them appear more important, and that perceived importance may assist enforcement/endorsement policies. Notes included that “reporting guidelines could be more prominent on journal author guideline pages”, or that “if the journal uses any sort of structured peer review (e.g., specific questions related to methodology) to tell authors this explicitly [on author guideline pages and within review feedback] and link it to the reporting guideline content”. A third suggestion was that “when journals

'stitch' or 'build' together the manuscript pdf (including cover letter, manuscript main text, appendices, etc.), prioritise the reporting guideline or move it earlier in the pdf".

However, a few stakeholders (publishers) warned that prioritising guidelines on author instruction pages "is complicated as these [pages] already have to do a lot", although guidelines could be more prominent if the pages "were better organised and/or filterable".

Promote reporting guidelines Stakeholders suggested many bodies that could help spread the word that reporting guidelines exist. For example, "professional societies" could "advertise" reporting guidelines, despite not having a role in the funding, regulation or dissemination of research.

Promotion could occur online. Most obviously, on stakeholder guidance web pages. "Email campaigns, social media, blogs" could be useful channels to "share and connect with others [and] drive traffic to guideline website[s]", but "these require time and energy" from the reporting guideline community to set up and manage.

"Conferences and workshops at institutions" were cited as channels to promote reporting guidelines offline, as were "seminars, webinars, and presentations" especially in "hard to reach countries/fields". These events were described as useful ways "to assist in the interpretation and use of the guidelines" and could be opportunities for "universities/funders/journals [to speak] together about the importance of reporting guidelines".

To reach students, participants suggested that universities could include reporting guidelines in their curricula, learning materials, or through reporting champions (see *Install reporting champions*).

Stakeholders wrote that "promotion can begin before a guideline has been published so that researchers know about guidelines being developed" and suggested "the provision of a time buffer/phasing period for updating new reporting guidelines which would allow researchers to have information about these new guidelines".

Install reporting champions Stakeholders wrote that publishers, universities, funders, and ethics committees could have members to promote and facilitate the usage of reporting guidelines, and used terms like “champions” and “EQUATOR ambassadors”. Within publishers, funders, or ethics boards, a champion’s responsibility may be to “expand knowledge/awareness of guidelines”. Within institutions, champions could also “help researchers” by “providing feedback on writing”. Early Career Researchers may feel most comfortable talking with a champion from “an accessible level (e.g., post-doc, library staff)”. Participants suggested this could follow a local network model (UKRN was cited as an example) with EQUATOR as the central organiser, and could utilize existing reproducibility networks.

Provide additional teaching Stakeholders proposed additional teaching as a way to promote reporting guidelines, make them easier to use, and to communicate the impacts of poor reporting. Education and training could be general (EQUATOR’s publication school was cited as an example) or could be “guideline-specific”, and could be delivered in person or online, as courses, videos, or text.

In addition to learning about a particular guideline, stakeholders suggested that students could learn about “writing as a process” and “workflows for documenting and communicating research”. This was considered useful as “researchers don’t necessarily understand that reporting is a stage in the research process”. Curricula could include “methods studies that indicate the research waste” to teach students why reporting matters, or students could learn for themselves by attempting “to replicate a study or do a systematic review to discover how poorly research is currently reported”.

To gain experience in using reporting guidelines, stakeholders suggested students could develop “research protocols (as Bachelor or Master Theses) using reporting guidelines” and that these could “be assessed based on the compliance with the appropriate reporting guideline” in addition to “other criteria more related with methodology”. To make this easier, stakeholders suggested structuring courses around reporting guideline items “for example: a course on [randomized controlled trials] covering every single SPIRIT or CONSORT item”.

Make updating guidelines easier Stakeholders acknowledged that feedback from authors “could be useful to guideline developers”. Some stakeholders went further, expressing that guidance and websites should be updated “in response to user feedback or changes in the field”. Others suggested that “developers could consult different user groups when creating guidance” and “engage as many health professions as possible” so that “professional cultural issues can be usefully accommodated.”

However, stakeholders were not forthcoming on how to go about gathering this feedback. One wrote “provide ways for researchers to give feedback to guideline developers” without suggesting any ways to do this. Another simply asked “how can we enable users to give feedback on guidance?”. Many guideline developers noted that they required access to extra funding to evaluate, refine, and update their resources. One developer suggested that “minor updates could be made without publishing a new article” if the guidance were disseminated on a website.

8.4 Discussion

8.4.1 Summary of results

My objective was to identify ideas to address influences affecting reporting guideline adherence by running workshops with the EQUATOR Network and focus groups with guideline developers, publishers, and other experts. As I had hoped, including perspectives from a range of stakeholders led to more ideas: EQUATOR staff thought of 24 which was then expanded to 128 through the focus groups.

Ideas employed all intervention functions. There were ideas to consider before, during, and after creating reporting guidelines, ideas to consider when creating resources, ideas about tools to assist application, ideas about ongoing activities to support or promote guidance use, and ideas about refining guidance over time in response to feedback. Many of these ideas could be enacted by guideline developers, publishers, and EQUATOR, but participants also saw opportunities for ethics committees, funders, academics, registries, and syllabus writers; stakeholders who are typically less frequently considered.

This is the first time that guideline developers from different groups have come together with publishers and academics to consider reporting guidelines as a system. I had a good response rate from guideline groups and responses were generally supportive, even if they were unable to take part. Multiple guideline developers volunteered their guideline to be a “guinneapig” and expressed support for my work. All stakeholders were open minded to the influences I presented except for one developer who expressed scepticism that reporting guidelines were anything but perfect, and requested additional evidence of authors’ negative experiences.

8.4.2 Implications

These results will be of interest to the reporting guideline community. Guideline developers may find inspiration here when writing or revising guidance. Many of the ideas could be enacted by the EQUATOR Network. At the time of writing, the EQUATOR Network was in the process of updating its advice for guideline developers, which I hope will be informed by these results. The ideas generated may also be of interest to publishers, funders, and Universities.

Although only a few reporting guideline development groups took part in this study, most ideas identified were abstract enough to generalise to most reporting guidelines, which tend to be developed and distributed in similar ways; (e.g., all development groups will have to consider what guidance to create, its scope, how to communicate it clearly and how to disseminate it). Some ideas may even generalise to other interventions to encourage good research practices (e.g., to communicate personal benefits, and not to patronize researchers).

8.4.3 Limitations

The study may have benefited from more diversity in participants’ backgrounds and expertise. I could have recruited participants from funders, ethics committees, or registries, behaviour change experts or experts familiar with user experience of websites or written documents. All participants were western and proficient in English. Although this is partly a consequence of who writes reporting guidelines, I could have sought

input from publishing houses that cater to non-western authors. Broadening the participant pool in this way could have led to more ideas.

My focus groups were smaller than I had planned. Some would consider these group sizes too small to be called focus groups, and may instead call them paired interviews, dyads, or triads ([174,175]). A hallmark of focus groups is that they use “group interaction to produce data” [176], and these interactions may include sharing experiences and challenging each other. I did not feel the small group sizes to be a limitation in this study for two reasons. Firstly, because participants were co-editing a file and building upon the thoughts of previous groups, participants could react and respond to participants from previous groups. Secondly, participants had deep understandings of the topic (evidenced by sessions overrunning and participants dwelling on a single topic) which meant that even pairs of participants had plenty to discuss, share, and debate. If I had condensed participants into, say, 3 groups of 5-6 participants, each participants would have had less time to speak and I anticipate that many ideas would have gone un-spoken.

8.4.4 Future work

I purposefully did not seek input from naïve authors as this study required input from experts familiar with reporting guideline dissemination. Instead, I describe how I sought input from authors in chapter 10. I also purposefully did not ask stakeholders to prioritize or rank ideas. In chapter 7 I explained that prioritization is subjective and, therefore, should be done by stakeholders separately.

8.4.5 Conclusions

In this chapter I have described how ran workshops with reporting guideline experts from the UK EQUATOR Centre to generate 24 ideas to address 32 influences affecting whether authors adhere to reporting guidelines. I then described how I ran focus groups with publishers, reporting guideline developers, and advocates, to expand this list to 128 ideas. In the next chapter I describe how I decided which ideas to act on and how I turned them into intervention components.

8.5 Reflections on this chapter

As with chapter 3, I decided to keep notes on my experience of writing this chapter using the SRQR guideline. I began by writing paragraphs for each guideline item in the order they are presented in the guideline. The resulting text lacked a narrative thread, and I found it difficult to explain how this chapter related to the previous one. As a result, my first draft required a significant re-write.

I also kept a diary of my experiences running these focus groups. The only time I felt uncomfortable was when talking with senior academics from the reporting guideline community. I did not want to make them feel like I was criticising their work. In my methods section I explained this concern was because I did not want participants to feel defensive or to limit discussion in any way. However, on reflection, I was also aware that these people may one day be my peer reviewers, collaborators, or even employers. So not only did I not want to criticise them, but I wanted to leave a good impression. I navigated these concerns by talking about reporting guidelines in an abstract way (as opposed to focussing on *their* guideline) and by relying heavily on the behaviour change wheel to structure discussion when necessary. I've already explained how the framework helped participants to brainstorm ideas systematically and not feel defensive, but on reflection I realised it also allowed *me* to relax, as it neutralised my questions that may otherwise have seemed personal or accusatory.

Chapter 9

Defining intervention components and developing a prototype

9.1 Introduction

In this chapter I describe how I brought together the outputs of all previous research chapters to create an intervention planning table containing intervention components. I then describe how I turned this table into a working prototype ready to be tested by authors.

In chapters 3 - 5 I described how I identified influences affecting whether authors adhere to reporting guidelines. In chapter 7 I described how ran workshops with reporting guideline experts at the UK EQUATOR Centre to define a target behaviour, consolidate influences, prioritise intervention functions and policy categories, and how I decided to prioritise redesigning reporting guidelines and the EQUATOR Network's website hope page. In chapter 8 I described how I asked guideline developers, publishers, and other stakeholders to build upon our workshop outputs by brainstorming solutions to barriers.

Consequently, at this point in my thesis I had multiple lists; a list of influences, lists of intervention options, and a list of ideas. Not all ideas were relevant to the EQUATOR Network or aligned with our prioritised intervention options, and although ideas

referenced influences, the links between ideas, influences, intervention functions, and behaviour change techniques were not fully explicit.

What remained was to select ideas I could implement, define them as intervention components, then build a prototype I could test with authors. Michie et al. [154] define an intervention component as:

1. a designed element that uses one or more behaviour change technique,
2. theorized to work through one or more intervention functions
3. to target one or more behavioural drivers.

Michie et al. [154] describe this step as “identify[ing] intervention content in terms of which [behaviour change techniques] best serve intervention functions”, where a behaviour change technique is “an active component” that is “observable, replicable, [and] irreducible”. Defining intervention content in this way is useful because it helps intervention developers understand why the component was added (or removed), how it is theorised to work, and how its effectiveness may be tested.

Once defined, I wanted to turn this list of intervention components into a prototype that could be tested by authors. This prototype would have to include a) a redesigned reporting guideline and b) a redesigned version of the EQUATOR Network home page that testers could navigate to find the redesigned reporting guideline. I wanted the prototype to be real enough so testers could use it as they would a real website. This meant I had to make decisions about content, design, and also the underlying software used to create and host the website.

I wanted to include EQUATOR staff and guideline developers in the redesign process for three reasons. Firstly, because EQUATOR staff will need to take responsibility of what I build if it is to be successful, the underlying architecture must be easy for them to understand and maintain and affordable to run without funding. Secondly, I expected their experience and expertise to help decision making and by prioritising their opinions I hoped to keep my own subjectivity in check. Thirdly, I wanted to ensure they understood the intervention and felt like it was “ours” not “mine”.

I decided to focus on redesigning one reporting guideline instead of many: the Standards for Reporting Qualitative Research (SRQR). Focussing on one guideline meant I could work alongside its creators to fully understand its nuances and allowed me spend time on design rather than repeatedly applying a design to multiple guidelines (even a simple design choice – like how to format an item – takes time to apply, and this time multiples with more guidelines). I chose SRQR because I was familiar with it, understood it, and its lead developer expressed interest in collaborating. Additionally, because it is applicable to all medical qualitative research I expected its users to be plentiful and varied. For example, authors may include students, clinicians, public health experts, quantitative researchers dabbling in qualitative methods within a positivist framework, or experienced qualitative researchers working within other paradigms. This broad and varied author base would be useful for a subsequent qualitative feedback study (see next chapter).

In summary, the aim of this chapter was to create an intervention in preparation for a subsequent qualitative user feedback study. My objectives were to:

1. Create a list of intervention components by combining the outputs of previous chapters.
2. Redesign the EQUATOR Network home page and SRQR guideline.

9.2 Methods

9.2.1 Defining intervention components

I started the 128 ungrouped ideas originating from my focus groups (chapter 8). I then cross-checked the list of favoured behaviour change techniques (from step 7 of chapter 8) with the list of influences (from step 4) to ensure I had considered all options. This sparked 8 more ideas (see Appendix O), bringing my starting list to 136 ideas.

I labelled each idea with:

1. The influence(s) it addresses (derived from chapters 3, 4, 5, and step 4 of chapter 7).
2. The behavioural driver it targets (from step four of chapter 7).

3. The intervention functions it uses.
4. The behaviour change techniques it uses, selected from Michie et al.'s taxonomy [158].

I asked a colleague with experience in behaviour change taxonomies to double check my coding, and we resolved disagreements through discussion.

For each idea, I then considered:

5. Whether it was implementable by redesigning guidelines or by modifying the EQUATOR website. If not, I removed it from the list.
6. Whether its intervention function aligned with the intervention options EQUATOR and I had prioritised during our workshops (see workshop steps 5 & 6 in chapter 7). If not, I removed it from the list.

To give structure to this list, I grouped intervention components according to the sub-behaviours they targeted: 1) engaging with guidance and 2) applying it (see section on specifying the target behaviour in chapter 7). To provide extra context and clarity, I described each component in more detail and how it compares to the status quo. In doing these comparisons I made generalisations about how popular reporting guidelines are written and disseminated by drawing on evidence from my qualitative synthesis (3), looking at how the EQUATOR website is currently, and referring back to workshop discussions (from chapter 7).

9.2.2 Redesigning the SRQR reporting guideline and EQUATOR Network home page

Co-design meetings and process

Throughout the creation process I drew on the principles of co-design: “active collaboration between stakeholders in the design of solutions to a pre-specified problem” [177]. In this chapter I describe how I worked alongside SRQR’s lead developer and reporting guideline experts from the UK EQUATOR Centre, and in the next chapter I describe how I continued these principles by seeking feedback from authors. I invited the EQUATOR staff who participated in the workshops (chapter 7) to form a design advisory group.

The group and I met three times between November 2022 and January 2023. In our first meeting, we decided how the home page and redesigned guideline should join and whether any interim pages were required.

In our second meeting, the design advisory group and I sketched ideas for how the home page and reporting guideline page could be laid out and for the positioning of intervention components. These sketches were wireframes: simple illustrations focussing on space allocation, functionalities, and intended behaviours. Wireframes do not include styling. They have no colours, images are represented as blank boxes, and lines represent blocks of text.

After the second meeting, once participants had agreed on a layout, I created an alpha version of the new home page and guideline page. These were real webpages, viewable in a browser, but I used dummy text and images because I wanted to solicit feedback on layout, structure, and functionality, not on the content. This is a common practice in web development, as people can become distracted by wording or stylistic choices. I used a web annotation tool called Pastel to collect feedback from the design advisory group [178] and then refined the alpha version based on this feedback.

In our third meeting, we co-created text for the home page. We began by listing the intervention components the text needed to address. The text needed to explain what reporting guidelines were, how they can be used, and the benefits they bring to authors. The design advisory group drafted text on their own before discussing and editing as a group. We also discussed style and imagery in this meeting. Again, we began by listing the intervention components the images needed to address. These included communicating what reporting guidelines are, who should use them, communicating simplicity and confidence. I invited design group members to contribute websites and images they admired for inspiration. We ended up discussing websites run by the National Health Service [179], the International Organisation for Standardisation [180], and the National Institute for Health and Care Excellence (NICE) [181]. We looked through examples of free-to-use images from a number of libraries [182–185]. I also consulted usability best-practices [186] and NICE’s style guide for guideline developers [187]. I then populated the alpha version with the text and images discussed in this meeting.

After the third meeting I began redesigning the SRQR guideline. I got written permission from SRQR's publisher (the Association of American Medical Colleges) and lead developer, Bridget O'Brien. I began redesigning SRQR by pasting the text from SRQR's Explanations and Examples supplement [30] into Microsoft Word and rearranging item content into categories identified in chapter 8 (see Appendix O *Describe reporting items fully*): what to write, how/where to write it, what to write if the item was not/could not be done, why the item is important and to whom, and examples. I edited sentences to speak directly to authors and to use active voice, e.g., "Describe X" instead of "X should be described". This shortened the text and made it clearer that the primary audience is authors.

For composite items I split the sub-items into bulleted lists. E.g.

For each X, describe:

- A
- B
- C

I rearranged conditional sub-items to read as "If X, then describe Y", instead of "Describe Y if X". I moved definitions into a glossary and contextual information into notes. I edited the tone of voice to add reassuring language. I asked SRQR's developer to provide feedback on the redesigned guideline and made refinements based on her comments.

After development, I double checked the intervention against my intervention planning table (see chapter 9) to ensure I had included all components. I invited another round of feedback from the design advisory group and made more refinements.

Trustworthiness

As before, I used established, best-practice techniques to facilitate open discussion in the design meetings [156]. I made it clear there were no right or wrong answers. I reflected on my own opinions before meeting, and held them back until the design advisory group had finished talking. When disagreements arose, I took time to explore and

understand both sides.

I made heavy use of my intervention planning table (see chapter 9) to make decisions based on theory over and above personal preferences. Instead of asking EQUATOR members whether they liked something (e.g., *Do you like this font?*), I asked whether they felt it reflected the intended intervention function instead (e.g., *Does the font convey simplicity?*). By prioritising theory above preference, and others' opinions above my own, I reduced my subjectivity.

Target audience

Throughout our meetings we kept our target audience in mind. Building user personas is a common best practice in web development [188–190]. For example, as part of their work to help NHS staff adopt new technologies and become a digital workforce, Health Education England identified archetypes: digital creators, end users, embedders, change drivers, and shapers [191]. They then developed imaginary personas around these archetypes, each with their own needs, motivations, and challenges. Developing personas can be a significant task. Some product teams go as far as interviewing 5 or more representatives of each persona in an iterative, co-creation process. This was not feasible within my DPhil and so my persona generation process was less formal but still evidence based. Referring to my website service evaluation (chapter 5), I recalled most website visitors are new and visit only once, and so I inferred many will be naïve to reporting guidelines and the EQUATOR Network. Website visitors come from all over the world, so I assumed some may speak English as a second language (or not at all). I had no way of knowing how much research experience EQUATOR's website visitors have, but I assumed the distribution would reflect real-world numbers, where students and early career researchers outnumber experienced professors. Although the EQUATOR website has content aimed at editors, librarians, and guideline developers, this content is rarely accessed and these users are not part of my target behaviour (see chapter 7), and so I did not create personas for them.

Consequently, two personas crystallised in my mind: inexperienced and experienced authors.

1. Inexperienced author:

- may not be familiar with EQUATOR or reporting guidelines
- may find authoring / publishing difficult
- may need help in finding, understanding, and applying a reporting guideline
- if English is their second language, they may be less proficient in reading/writing in it

2. Experienced author:

- may be familiar with EQUATOR and/or reporting guidelines
- may have used checklists before, but may not realise full guidance exists, and may have never used reporting guidelines for drafting.
- may feel patronized or restricted by reporting guidelines, especially if they believe them to be design requirements
- may see checklists as red tape necessary for publication, or they may be reporting guideline “converts” already convinced of their value.

Many authors may fall somewhere between these two personas. In discussion with the design advisory group, I decided to prioritise inexperienced authors because I believed them to be more numerous, and because qualitative evidence suggests inexperienced authors may benefit more from reporting guidelines but experience more barriers when using them (see chapter 3). By focussing on helping inexperienced authors, I hoped any spill-over effects may also help experienced authors.

User personas typically include motivations. There is an important difference between user motivations - what authors are wanting to do when they visit EQUATOR’s website - and our target behaviour - what *we* want authors to do when they visit EQUATOR’s website. Because most visitors come from journal websites or submission systems and because checklists are the most accessed resource (chapter 5), I assumed most authors visit the website because they want to fill out a reporting checklist as part of journal submission. In contrast, my target behaviour specifies authors should read the full reporting guidelines *as early as possible*, i.e., when planning or drafting research, way before

they submit to a journal. I could have designed a website to focus entirely on drafting manuscripts, but this would have abandoned authors seeking checklists. Instead, I took a more pragmatic approach and decided to continue catering for authors seeking checklists, but to nudge them towards using the full guidance for drafting research in the future.

Making decisions about system architecture

When considering architecture options I spoke to the UK EQUATOR Centre and University departmental staff members about financial constraints, data security concerns, and to understand what website building tools and technologies they felt comfortable with. I then looked at the designs we had created on paper, considered the functionality those designs would require, and selected architecture options to support those functions.

9.3 Results

9.3.1 Intervention components

From the initial list of 136 ideas I identified 62 intervention components that could be implemented by redesigning guidelines and by improving the EQUATOR Network home page. Together, these components use 17 behaviour change techniques and 6 intervention functions to target 32 influences. Of these 62 intervention components I managed to include 46 when redesigning SRQR and EQUATOR's home page. The remainder were not feasible.

The full intervention component table is long and disrupted the narrative of this chapter, and so I moved it to Appendix P. In this chapter I've included shorter tables for each intervention part, and summarised included components labelled with their intervention function and employed behaviour change technique. I'll now go through parts of the design and summarize the components in each.

9.3.2 Intervention description

Site organisation and search

Some intervention components required considering how authors would navigate and search the website (see Table 9.1). My first meeting with the design advisory group focussed on how the redesigned home page and guideline should link, and whether we needed any steps in between the two. On EQUATOR's existing website, authors starting on the home page must navigate through 5 webpages to reach the full reporting guidance; the home page, guideline database page, PubMed, a publication, and then sometimes a supplement. From my website service evaluation (chapter 5), we knew many authors leave at each step (see Figure 9.1). Therefore, we wanted our redesigned home page to link directly to the most frequently used guidelines, thereby reducing this journey to 2 steps, with the aim to increase the proportion of authors reaching the full guidance (see Figure 9.2).

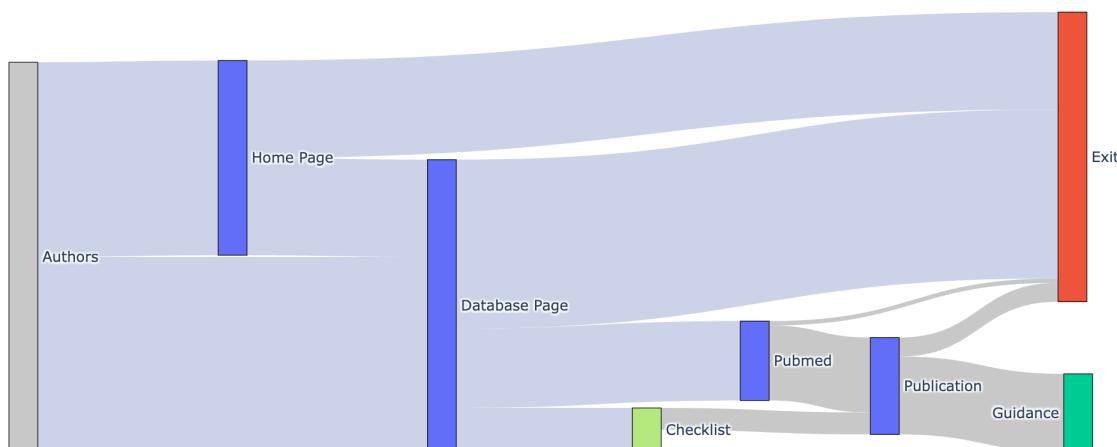


FIGURE 9.1: A simplified layout of the existing EQUATOR Network website, as of the 5th of April, 2023. Users must navigate through up to 5 different web pages to reach reporting guidance. The proportion of users navigating between each step is shown in the width of the links. Links in grey are estimated proportions. “Exit” means leaving the website. For simplicity, I have not included the 7% of authors viewing pages within the “Library” subdirectory, nor the 1% visiting other content categories.

We also decided to display the full guidance texts on the website whenever licences allow. Currently, EQUATOR's website directs authors to different publishers and guideline development group websites. Sometimes the full guidance is published as a supplementary material, and checklists can appear as supplements, tables, or figures. Bringing the full guidance texts to a single website makes them easier to find, means we can

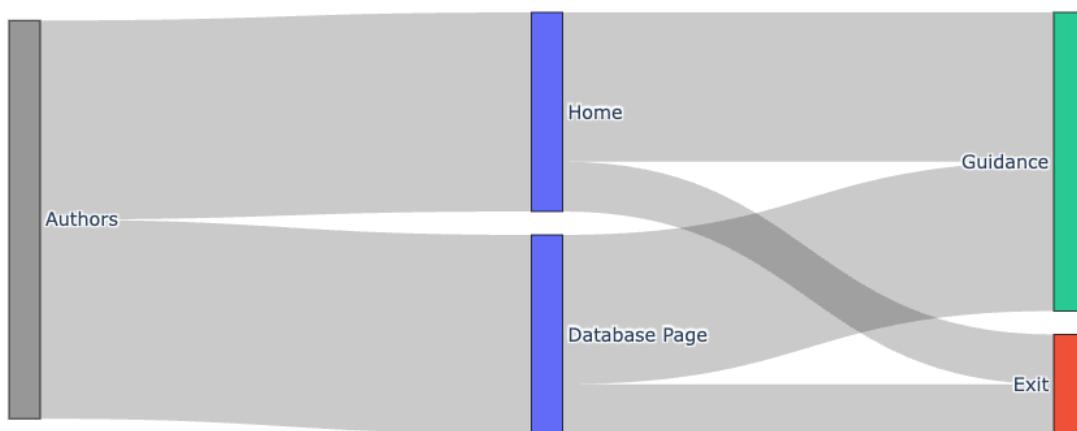


FIGURE 9.2: The layout of the new web-intervention. Users must now only need to navigate to 2 pages to access reporting guidance. The proportion of users navigating between each step is shown in the width of the links. All links in are estimated proportions, based on a realistic aim to reduce the exit rate from the home page and database page by 50%.

ensure all have a consistent, usable format, and allows us to track authors behaviour. But this is only possible when guidance is published under a permissible licence.

When discussing how authors would navigate the website, we decided to cater to both users that search and users that browse. The website features a search function adhering to user experience norms (the search button on EQUATOR's original website is difficult to find), and each web page has enhanced meta data to optimise search engine performance. To help browsing authors, we decided to ensure all related guidelines and resources link to each other. Previously, guidelines rarely linked to each other, and checklists rarely linked to full guidance. In the redesign, guidelines link to related guidelines, and all resources link to each other. In the future, EQUATOR may consider creating collections of related guidelines (e.g., a page featuring all guidelines relating to a study design, health condition, or theme), decision aid tools, or may embed related guidelines (e.g., PRISMA for Abstracts could be embedded within PRISMA).

TABLE 9.1: Intervention components pertaining to how the website is organised, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Centralised hosting	Enablement	Restructuring the physical environment
Search function on website	Enablement	Restructuring the physical environment
Links between related guidelines	Environmental Restructuring	Restructuring the physical environment
Search Engine Optimization	Enablement	Restructuring the physical environment
link all resources to each other	Environmental restructuring	Restructuring the physical environment

Intervention components that occur throughout

Some intervention components pervaded all parts of the redesign (see Table 9.2). These components centred on three themes: our target audience, nudging authors towards using guidelines earlier in their writing, and positioning reporting guidelines as helpful rather than administrative.

In the methods section I described how EQUATOR and I decided that our primary audience was authors (not editors, reviewers etc.), and *inexperienced* authors in particular. In my thematic synthesis (chapter 3) I found some authors felt reporting checklists were the responsibility of an editor or peer reviewer. In part, this might be because some reporting guidelines are written passively (“Reports should describe...” etc.). In contrast, my design speaks directly to the author (“When writing up *your* research...” etc.), and I used terminology with inexperienced authors in mind.

Many components nudge authors towards using reporting guidelines early in their writing. For instance, the first text authors see mentions “writing up” and the guideline

text is worded from the perspective of writing (as opposed to checking or reviewing). The home page and guideline page explicitly describe stages of writing where reporting guidelines can be used. The website layout forces authors to access the full guidance before the checklist, and the checklist button is positioned after the button for a template that can be used when drafting. The meta-data for search engines includes terms like “how to write” or “describe” so the pages appear in search engine results when authors search queries like ‘how to write up an epidemiology study’. In contrast, my thematic synthesis revealed that some authors may currently think reporting guidelines *are* checklists, and are meant to be used as part of journal submission. This is, after all, how most authors discover them and few guidelines contain any instruction to the contrary.

Finally, because my thematic synthesis found some authors describe reporting guidelines as administrative red-tape, I wanted to position reporting guidelines as trusted resources that will help, not hinder, researchers. I did this by creating content and an aesthetic that would foster trust and convey two personal benefits of using reporting guidelines: simplicity and confidence. I created this aesthetic by using a simple colour scheme with plenty of negative space. I tried to keep the tone of voice professional but simple and empowering. I avoided design or language that may invoke feelings of judgement, complexity, or administration. I tried to keep language concise and clear without being patronizing.

TABLE 9.2: Intervention components appearing throughout the website, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Remove branding and messaging that may invoke feelings of judgement, complexity, or administrative red-tape	Coercion (Removal of)	Remove aversive stimulus
Include design, features, and language to foster trust	Persuasion	Credible source, Social comparison

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Use tone of voice and design to communicate personal benefits; confidence and simplicity	Persuasion	Framing/reframing
Avoid patronizing language	Persuasion	Remove aversive stimulus
Address communications to authors	Education	Instruction on how to perform a behaviour
Reassure that all research has limitations to encourage explanation over perfect design	Persuasion	Social support (unspecified)

Home page

EQUATOR's current home page (see Figure 9.3) does not define what reporting guidelines are, or how or why to use them. Although it is full of content, the overall impression is one of complexity with no clear message. Intervention components appearing on the home page are listed in Table 9.3. Its limitations include:

- 1) No prominent description of what reporting guidelines are
- 2) No clear instruction on what tasks reporting guidelines can and cannot be used for
- 3) Search function hard to find (area A)
- 4) Decision tool for identifying which reporting guideline to use was hard to find (area B)
- 5) The page looked cluttered and unappealing
- 6) The tone of voice was functional. It was not particularly judgemental but not reassuring either.
- 7) There was little description of benefits of using a reporting guideline besides the mention of 'quality' and 'transparency' in the definition of EQUATOR, reference to 'high-impact research', 'improve your writing', and 'enhance your peer review' in the header.

- 8) No reassurance that most research has limitations
- 9) Frequently accessed guidelines are fairly easy to find (area C).

In contrast, the redesigned homepage aims to minimise the proportion of authors leaving the website without engaging with reporting guidelines. The home page explains what reporting guidelines are, what writing tasks to use them for, the personal benefits they offer, and why reporting is important for reproducibility. Short quotes from authors and editors, expounding the benefits of using reporting guidelines aim to convince sceptical authors. The redesigned home page can be seen in Figure 9.4. Intervention changes made to the homepage include:

- 1) Reporting guidelines are now clearly defined (areas A)
- 2) The site looks simple and has plenty of white space
- 3) Personal benefits are described explicitly and communicated through reassuring language and quotes (see areas B)
- 4) Search and browse buttons are easy to find (area C)
- 5) Frequently accessed guidelines are still easy to find (area D)
- 6) The site describes what tasks reporting guidelines can be used for, and differentiates tools by task (area E)

TABLE 9.3: Intervention components appearing on the website's home page, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Describe what reporting guidelines are where they are first encountered	Education	Instruction on how to perform the behavior
Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for	Education	Instruction on how to perform the behavior
Educate authors about writing as a process	Education	Instruction on how to perform a behaviour

INTERVENTION COMPONENT	FUNCTION	INTERVENTION
Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications)	Education	Information about emotional consequences, Information about others' approval, Information about social and environmental consequences
Include testimonials from researchers who were nervous about being punished for reporting transparently	Persuasion	Demonstration of the behavior
Gather and communicate evidence for benefits	Persuasion	Information about emotional consequences, Information about others' approval, Information about social and environmental consequences
Include testimonials from research users who benefit from complete reporting	Persuasion	Salience of consequence
Explain importance of complete reporting to the scientific community	Education	Information about social and environmental consequences

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Tell authors when to use reporting guidelines, or that reporting guidelines are best used as early as possible	Education	Instruction on how to perform the behavior

Introduction to the SRQR reporting guideline

Although some journals direct authors to EQUATOR's home page, others send authors straight to EQUATOR's database records for a particular reporting guidelines. Although EQUATOR's database pages for each guideline do include some explanation of when the guideline should be used, and links to related guidelines and tools, this information was quite hard to find and the pages lacked many other intervention components (see Figure 9.5). Limitations of the existing database pages include:

- 1) The actual guidance is hard to find. Area A includes 3 links. The first two send users to an article describing how SRQR was developed. The actual guidance appears in a supplement of that article, which is the third link in area A. The label "relevant URLs" is vague.
- 2) Little instruction regarding what the reporting guideline is or can be used for other than "Qualitative research"
- 3) Links to related guidelines that are hard to find or, as for SRQR, absent
- 4) No metrics around how many authors use this reporting guideline (e.g., citation counts)
- 5) The French translation of the guidance is well labelled and fairly easy to find (area B), but to the right of it is a box prominently labelled "Translations", and the link in here would actually take the user further away from the translated guidance.

EQUATOR's database pages direct authors to reporting guideline publications. Although these publications might include (or link to) the full guidance, most focus on lengthy descriptions of how the guidance was made and justifications of why it is

equator network

Enhancing the QUAlity and Transparency Of health Research

[Website translation help](#)

[Home](#) [About us](#) [Library](#) [Toolkits](#) [Courses & events](#) [News](#) [Blog](#) [Librarian Network](#) [Contact](#)

Your one-stop-shop for writing and publishing high-impact health research
[find reporting guidelines](#) | [improve your writing](#) | [join our courses](#) | [run your own training course](#) | [enhance your peer review](#) | [implement guidelines](#)

Library for health research reporting

The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

A [Search for reporting guidelines](#)
B [Not sure which reporting guideline to use?](#)
[Reporting guidelines under development](#)
[Visit the library for more resources](#)

C **Reporting guidelines for main study types**

	CONSORT	Extensions
Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Study protocols	SPIRIT	PRISMA-P
Diagnostic/prognostic studies	STARD	TRIPOD
Case reports	CARE	Extensions
Clinical practice guidelines	AGREE	RIGHT
Qualitative research	SQROR	COREQ
Animal pre-clinical studies	ARRIVE	
Quality improvement studies	SQUIRE	Extensions
Economic evaluations	CHEERS	

[See all 557 reporting guidelines](#)

Toolkits

Find practical help and resources to support you in:

- [Writing research](#)
- [Peer reviewing research](#)
- [Using guidelines in your journal](#)
- [How to develop a reporting guideline](#)

[View all Toolkits](#)

Endorsements

THE LANCET

Funding

Pan American Health Organization
World Health Organization Americas

Search
Search the whole website

[About this site](#) [Accessibility statement](#)
[Terms of use](#) [Contact](#) [Privacy policy](#)

EQUATOR highlights

27/02/2023 - Automatic translation now available on the EQUATOR Network website
We're delighted to announce that our partnership with the Pan American Health Organization (PAHO/WHO) has resulted in [http://www.equator-network.org](#) becoming multilingual

6/01/2022 - ICMJE Recommendations updated to include new section on preprints
The International Committee of Medical Journal Editors (ICMJE) has updated its official Recommendations for journals and authors and now includes a whole new section about preprints. The document now sets out more detailed advice for journals and authors regarding preprints, [...]

Interesting videos

EQUATOR Canada Publication School team educational video
The EQUATOR Canada Publication School team (consisting of patient partners and researchers) have launched an educational video resource, titled "How do I publish a paper? The introductory video provides viewers with practical guidance on how the publication team, consisting of patient/public partners and research team members, can work together to define roles and contributions throughout the publication process.

Centre for Journalology Speaker Series video
The Canadian EQUATOR Centre, home to the Centre for Journalology, has started a new Speaker Series for 2021. The first session in the series was held on the 14th of January and was presented by Dr Lisa Caulley, Dr Robert Frank, Hassan Khan, and Alicia Ricketts, from the Centre. They spoke about recent research on reporting guidelines, planned research examining open science hiring practices, and journal transparency. The recording of the talk is freely available.

News

CONSORT and PRISMA websites down – alternative ways to access the checklists
1/03/2023

Automatic translation now available on the EQUATOR Network website
27/02/2023

EQUATOR Network Newsletter January 2023
31/01/2023

EQUATOR Network Newsletter October 2022
31/10/2022

Launch of the EQUATOR Canada Publication School team educational video
9/09/2022

[Sign-up to our newsletter](#)
to keep up-to-date with the latest developments by email.

Latest guest blogger

What is the little thing you can do to increase reproducibility, replicability and trust in science?
5/10/2020

How can reporting quality interfere with reproducibility issues and overall trust in science results? With that question in mind, we participated in the Reproducibility, Replicability and Trust in Science conference organised by the Wellcome Genome Campus from 9 to 11 [...]

NDORMS
NUFFIELD DEPARTMENT OF ORTHOPAEDICS,
RHEUMATOLOGY AND MUSCULOSKELETAL SCIENCES
 UNIVERSITY OF OXFORD

FIGURE 9.3: The existing EQUATOR home page, as captured on 5th of April, 2023. Labelled areas are limitations and area described in the previous paragraph.

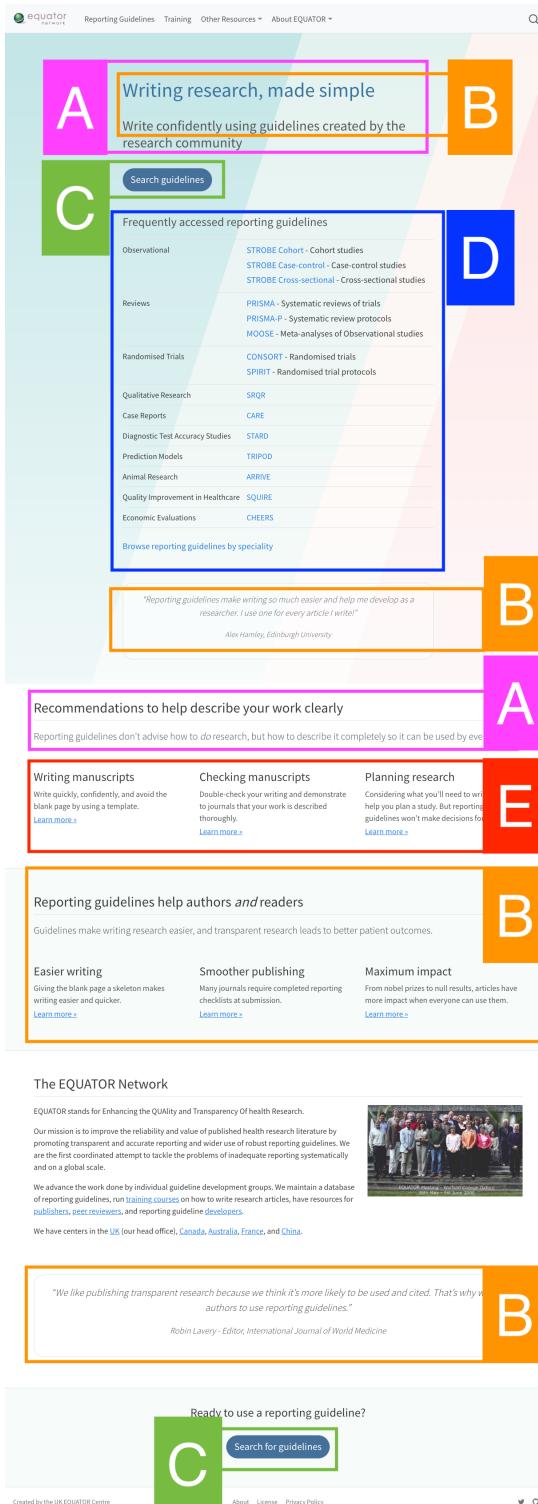


FIGURE 9.4: Redesigned intervention home page.

[Translate](#)


eQUATOR
network

Enhancing the QUAlity and Transparency Of health Research

[Home](#) [About us](#) [Library](#) [Toolkits](#) [Courses & events](#) [News](#) [Blog](#) [Librarian Network](#) [Contact](#)

Home > Library > Reporting guideline > Standards for reporting qualitative research: a synthesis of recommendations

Search for reporting guidelines

Use your browser's Back button to return to your search results

 **Standards for reporting qualitative research: a synthesis of recommendations**

Reporting guideline provided for?	Reporting of qualitative research studies. (i.e. exactly what the authors state in the paper)
Full bibliographic reference	O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-1251.
Language	English
PubMed ID	24979285 
Relevant URLs (full-text if available)	The full-text of this reporting guideline is available at: SRQR guideline
Detailed descriptions of each reporting item and examples are available in the supplementary material at: http://links.lww.com/ACADMED/A218	

Availability in additional languages

Reporting guideline acronym	SRQR
Study design	Qualitative research
Applies to the whole report or to individual sections of the report?	Whole report
Record last updated on	January 18, 2023

Reporting guidelines for main study types

Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Study protocols	SPIRIT	PRISMA-P
Diagnostic/prognostic studies	STARD	TRIPOD
Case reports	CARE	Extensions
Clinical practice guidelines	AGREE	RIGHT
Qualitative research	SRQR	COREQ
Animal pre-clinical studies	ARRIVE	
Quality improvement studies	SQUIRE	Extensions
Economic evaluations	CHEERS	

Translations

Some reporting guidelines are also available in languages other than English. Find out more in our [Translations section](#).

About the Library

For information about Library scope and content, identification of reporting guidelines and inclusion/exclusion criteria please visit [About the Library](#).

Visit our [Help page](#) for information about searching for reporting guidelines and for general information about using our website.

Library index

- [What is a reporting guideline?](#)
- [Search for reporting guidelines](#)
- [Browse reporting guidelines by specialty](#)
- [Reporting guidelines under development](#)
- [Translations of reporting guidelines](#)
- [EQUATOR Network reporting guideline manual](#)
- [Reporting guidelines for animal research](#)
- [Guidance on scientific writing](#)
- [Guidance developed by editorial groups](#)
- [Research funders' guidance on reporting requirements](#)
- [Professional medical writing support](#)
- [Research ethics, publication ethics and good practice guidelines](#)
- [Links to other resources](#)
- [About the Library](#)

Search

[About this site](#) [Accessibility statement](#)
[Terms of use](#) [Contact](#) [Privacy policy](#)

The UK EQUATOR Centre is hosted by the Centre for Statistics in Medicine (CSM), NDORMS, University of Oxford. The EQUATOR Network website and database is provided by the UK EQUATOR Centre.


NDORMS
NUFFIELD DEPARTMENT OF ORTHOPAEDICS,
RHEUMATOLOGY AND MUSCULOSKELETAL SCIENCES


UNIVERSITY OF
OXFORD

FIGURE 9.5: The existing EQUATOR page for SRQR, as captured on 5th April, 2023. Labelled limitations are described in the previous paragraph.

needed. This is true for SRQR (see Figure 9.6). Limitations of these publications may include:

- 1) Not all reporting guidelines describe what reporting guidelines are or what they can be used for, and these descriptions can be hard to find (areas A).
- 2) The actual guidance or checklist may be relegated to a box, table (area B), or a linked supplement (see area C).
- 3) Reporting guideline publications may not reassure authors that most research has limitations, and that transparency is OK
- 4) Publications may not be written with a reassuring tone of voice. Guideline developers often justify reporting items by emphasising the negative impact of research waste. This may be how developers justify their work to themselves, editors, reviewers, or readers. As a result, to a naïve author considering using the guidance, the tone of voice may come across as judgemental.
- 5) Benefits to the user may be hard to find or (as with SRQR) not described at all. Benefits to *others* are more likely to be described, including a focus on how transparent, complete reporting benefits the research community or, conversely, how poor reporting is wasteful.
- 6) Instruction on when reporting guidelines do/do not intend to prescribe structure, or instruction may be hard to find (see area D) or missing.
- 7) Instructions on whether a reporting guideline intends to be a strict standard vs. ‘just’ a guideline may be hard to find (see area D) or missing.
- 8) Links to related resources only include those that were created before the reporting guideline was published. Some guidelines don’t include any links.
- 9) No clear instruction on whether to use the guideline in a situation that it wasn’t designed for, but when no better guidance exists.

In my redesign, I wanted authors to be able to find the full guidance directly, without having to navigate a database page or lengthy development article. However, I still wanted to include an introduction and FAQ section to help authors quickly understand what the guideline is, what it *isn’t*, whether it applies to them and, if it does, to convince them to use it (see Table 9.4). Because half of visitors bypass the home page and arrive directly on these guideline-specific pages, I decided to duplicate some home page



FIGURE 9.6: The SRQR publication, captured on the 5th April, 2023. Areas A-D show the scope, checklist, link to the guidance, and its strictness.

components aimed at naïve authors such as basic information on what reporting guidelines, how to use them, when to use them, and why. Intervention changes included in reporting guidelines' introductions and FAQs (shown in Figure 9.7) include:

- 1) Clear description of what the reporting guideline is, what it can and cannot be used for, the benefits to the author and to society, and how and when it can be used. (area A)
- 2) Description of whether the reporting guideline is intended to be a standard or 'just' a guideline (area A)
- 3) Tools are clearly differentiated by task (area B)
- 4) Related guidelines and other resources are linked. These links can be updated as and when newer guidelines are published (area C)
- 5) Clear instruction on whether a reporting guideline can be used in a situation that it wasn't designed for, but where no better guidance exists (area D)
- 6) Links to translations (area E)
- 7) Reassuring language throughout, and reassuring quotes from editors, readers, and authors (e.g., area F)
- 8) Citation metrics (area G)
- 9) An estimation of how long guidance will take to read (area H)
- 10) Advice on how or where to report items so as not to breach word count limits and when reporting guidelines do or do not intend to prescribe structure (area I)
- 11) Citation information (area K)
- 12) Information on how the guidance was developed and why it can be trusted (area L)

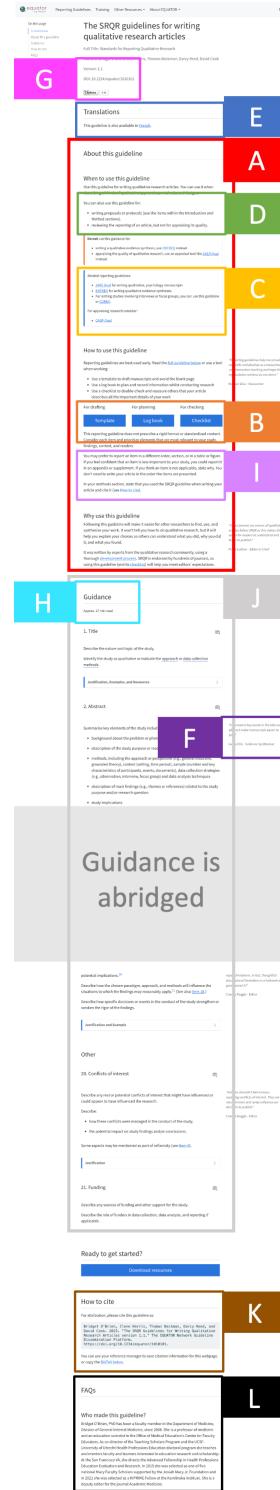


FIGURE 9.7: Redesigned intervention reporting guideline introduction and FAQ. Labelled areas include intervention components and are described in the preceding paragraph.

TABLE 9.4: Intervention components appearing in the introductory material preceding the reporting guideline, alongside the intervention functions and behaviour change techniques (BCT) they employ. Italicized items are duplicated from previous tables. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Instruct authors to cite reporting guidelines so readers may learn about them	Education	Instruction on how to perform the behaviour
Describe the scope of a reporting guideline at the top of every resource	Education	Instruction on how to perform the behaviour
Use if-then rules to direct authors to more appropriate and up-to-date guidance when available	Education	Instruction on how to perform the behaviour
Explicitly state when no better guidance exists for a use case	Education	Instruction on how to perform the behaviour
Provide translations	Enablement	Instruction on how to perform the behaviour
Tell authors how long the guidance will take to read	Education	Instruction on how to perform the behaviour
Explain when reporting guidelines do not intended to prescribe structure	Education	Instruction on how to perform the behaviour
Reassure when guidelines are <i>just</i> guidelines	Persuasion	Social support
Explain how the guidance was developed and why it can be trusted	Education	Credible source
<i>Tell authors when to use reporting guidelines, or that reporting guidelines are best used as early as possible</i>	Education	Instruction on how to perform the behaviour
<i>Describe what reporting guidelines are where they are first encountered</i>	Education	Instruction on how to perform the behaviour

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
<i>Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for</i>	Education	Instruction on how to perform the behaviour

Reporting guideline content

Reporting guidelines comprise reporting *items*, and an example from SRQR is in Figure 9.8. Limitations of how items are typically described include:

- 1) Text is unstructured, so it is difficult to immediately identify what needs to be written.
- 2) Text uses verbose, passive language
- 3) The text appears long and difficult to digest
- 4) Terms are not always defined
- 5) Not all reporting items are justified
- 6) Not all items include instruction of what to write if the item could not/was not done.

When redesigning items, although I did not change the meaning of guidance, I did give additional structure and edit phrasing. To increase clarity, I used plain language, consistent terms, and added definitions when I felt necessary. I wanted to cater to authors who read from start to finish as well as those who dip in and out, and so I created navigation menus so authors can jump to items. Intervention components applying to the entire reporting guideline content (which includes but is not limited to reporting items) appear in Table 9.5.

For each reporting item, I prioritized instructions of what to write. I placed remaining content in an expandable box, thus making the guidance appear shorter at first glance. Each expandable box includes why the reporting item is important and to whom, examples, links to related advice, and any other content that isn't an instruction of what to write. Each item links to its own discussion channel. Two intervention components

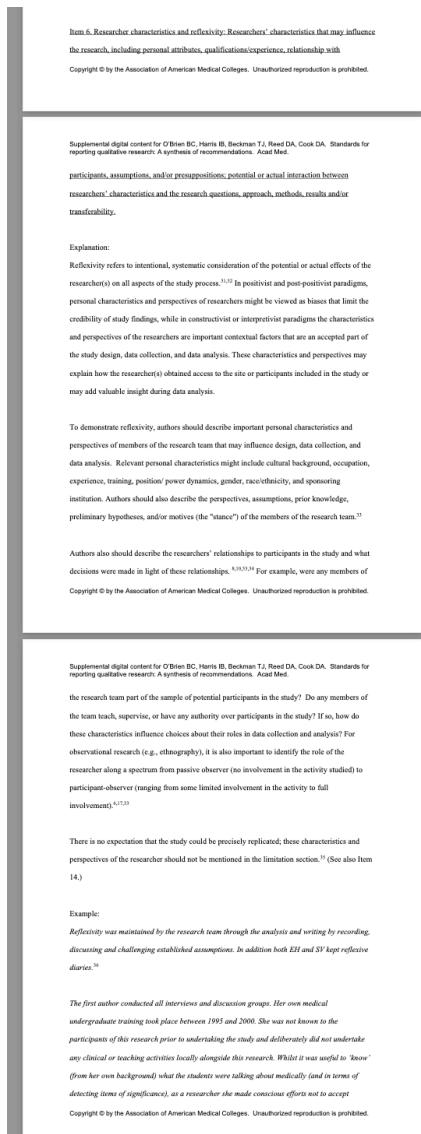


FIGURE 9.8: An example item from the SRQR guideline.

required presenting design advice separately to reporting advice, but this was not necessary as SRQR made few design assumptions or recommendations. Intervention components applying to reporting items appear in Table 9.6.

Figure 9.9 shows a redesigned reporting item. Intervention changes include:

- 1) Content is separated into what to write (area A), why information is important (area B), examples (area C), and any additional background information (not shown).
- 2) Areas B and C are presented as expandable content, so the only instruction immediately visible is what to write (area A). This means that the guidance is easier to digest and less intimidating.
- 3) Definitions are presented as pop-ups for technical terms
- 4) Quotes provide reassurance and persuasion (area D)
- 5) Language is direct and edited for clarity and brevity
- 6) Each item has its own discussion page (see Figure 9.10, linked to from the top right of area A)

TABLE 9.5: Intervention components pertaining to the content and organisation of reporting guidelines, alongside the intervention functions and behaviour change techniques (BCT) they employ. A reporting guideline's content includes (but is not limited to) reporting *items*. Components specific to reporting items appear in Table 9.6. Italicized items are duplicated from previous tables. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Cater to different kinds of user (readers vs dippers) by structuring guidance with headings, itemisation, hyperlinking to particular sections, and with optional content	Environmental restructuring	Restructuring the physical environment

INTERVENTION COMPONENT	FUNCTION	INTervention
		BCT
Make guidance appear shorter by removing superfluous information, hiding optional content, splitting long guidelines, using concise language, and separating design advice	Environmental restructuring	Restructuring the physical environment
Decrease fear of judgement by making reporting guidelines design agnostic	Coercion (Removal of)	Remove aversive stimulus
Present design advice separately to reporting advice	Coercion (removal of)	Restructuring the physical environment
Use plain language	Enablement	Instruction on how to perform the behaviour
Define key terms	Education	Instruction on how to perform the behaviour
Use consistent terms	Enablement	Instruction on how to perform the behaviour
<i>Educate authors about writing as a process</i>	Education	Instruction on how to perform a behaviour
<i>Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications)</i>	Education	Information about emotional consequences, Information about others' approval, Information about social and environmental consequences

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
<i>Include testimonials from research users who benefit from complete reporting</i>	Persuasion	Salience of consequence
<i>Include testimonials from researchers who were nervous about being punished for reporting transparently</i>	Persuasion	Demonstration of behaviour
<i>Gather and communicate evidence for benefits</i>	Persuasion	Information about emotional consequences, Information about others' approval, Information about social and environmental consequences

TABLE 9.6: Intervention components appearing in reporting guideline items, alongside the intervention functions and behaviour change techniques (BCT) they employ. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Structure guideline items to make them quicker to digest	Enablement	Restructuring the physical environment
For each item, explain why the <i>information</i> is important and to whom (not just what constitutes “good” design)	Education	Information about social and environmental consequences
Provide links to other resources that explain how an item can be done	Education	Instruction on how to perform the behaviour

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
For each item, provide clear instruction of what needs to be described	Education	Instruction on how to perform the behaviour
For each item, provide examples of reporting in different contexts	Modelling	Demonstration of the behaviour
Create spaces for authors to discuss reporting guidelines with others	Persuasion	Social comparison, Credible source, Adding objects to the environment
Provide clear instruction of what needs to be described when an item was not done, could not be done, or does not apply	Education	Instruction on how to perform the behaviour
Provide instruction as to how and where information can be reported without breaching word count limits or making articles bloated.	Education	Instruction on how to perform the behaviour
Encourage explanation even when choices are unusual or not optimal	Education	Instruction on how to perform the behaviour

FAQ

At the end of the guideline, an FAQ section includes information that was too lengthy to be fully discussed in the page's introduction (see Table 9.7). This includes a longer description of how the guideline's development (and link to the original publication), citation instruction, and advice on how to respond if an editor or co-worker asks you to remove guideline content.

6. Researcher characteristics and reflexivity

Describe how roles and identities of research team members influence choice of research approach, data collection, and data analysis.⁵

Describe the perspectives, assumptions, prior knowledge, preliminary hypotheses, and/or motives (the "stance") of the members of the research team.

Describe the researchers' relationships to participants in the study and what decisions were made in light of these relationships.⁶

If your research was observational (e.g., ethnography), describe the role of the researcher along a spectrum from passive observer (no involvement in the activity studied) to participant-observer (ranging from some limited involvement in the activity to full involvement).

There is no expectation that the study could be precisely replicated; these characteristics and perspectives of the researcher should not be mentioned in the limitation section. (See also [Item 14: Data Analysis](#))

A

Coming from quantitative research it felt strange to include such personal descriptions of the team. But I came to realise why it was useful to others, and that peer reviewers would expect it."

Tim Westland - Researcher

B

Justification and Examples

Jump to:

- [Why readers need this information](#)
- [Examples](#)

Why readers need this information

In positivist and post-positivist paradigms, personal characteristics and perspectives of researchers might be viewed as biases that limit the credibility of study findings, while in constructivist or interpretivist paradigms the characteristics and perspectives of the researchers are important contextual factors that are an accepted part of the study design, data collection, and data analysis. These characteristics and perspectives may explain how the researcher(s) obtained access to the site or participants included in the study or may add valuable insight during data analysis.

C

Examples

Reflexivity was maintained by the research team through the analysis and writing by recording, discussing and challenging established assumptions. In addition both EH and SV kept reflexive diaries.

The first author conducted all interviews and discussion groups. Her own medical undergraduate training took place between 1995 and 2000. She was not known to the participants of this research prior to undertaking the study and deliberately did not undertake any clinical or teaching activities locally alongside this research. Whilst it was useful to 'know' (from her own background) what the students were talking about medically (and in terms of detecting items of significance), as a researcher she made conscious efforts not to accept potentially common assumptions at face value.

[Back to top](#)

D

FIGURE 9.9: A redesigned item from the SRQR reporting guideline. Labelled intervention changes are described in the paragraph above.

The screenshot shows a web page from the EQUATOR network. At the top, there is a navigation bar with links for 'Reporting Guidelines', 'Training', 'Other Resources', and 'About EQUATOR'. A search bar is also present. The main title of the page is 'Discussion for SRQR item: Techniques to enhance trustworthiness'. Below the title, there is a section titled 'Techniques to enhance trustworthiness' with a descriptive text. Further down, there is a 'Justification, Examples, and Resources' section, a 'Comments' section (which is currently empty), and a 'Footnotes' section containing a detailed list of techniques. A large blue rectangular area on the right side of the page is labeled with a white letter 'A'.

FIGURE 9.10: Intervention discussion page. Every reporting item now has its own discussion page where authors can ask and answer questions (area A), and provide feedback to guideline developers.

TABLE 9.7: Intervention components pertaining to the website's FAQ section, alongside the intervention functions and behaviour change techniques (BCT) they employ. Italicized items are duplicated from previous tables. See Appendix P for further descriptions and links to barriers each component addresses.

INTERVENTION		
INTERVENTION COMPONENT	FUNCTION	BCT
Provide advice regarding how to respond if asked to remove reporting guideline content by a colleague, editor, or reviewer	Education	Problem solving
<i>Explain how the guidance was developed and why it can be trusted</i>	Education	Credible source

9.3.3 System architecture

EQUATOR's existing website is built using Wordpress. Although EQUATOR staff can make basic content modifications using a content managing system, they rely on software developers for more significant alterations. I considered Wordpress and other DIY website builders (like Wix [192] or Squarespace [193]) but decided against them. Most come with expensive monthly subscriptions. Most offer a 'drag and drop' building experience which, although easy to use, is a laborious way of uploading and formatting large amounts of content. Should EQUATOR want to change how reporting items are presented (for example, reposition examples), they would have to manually edit every item for every reporting guideline. Additionally, many intended intervention changes required functionality not offered by these services (e.g., popup glossary definitions, discussion boards).

Creating a custom coded website also felt unacceptable. Html and javascript are two languages commonly used for creating websites. Although many software developers will be fluent in them, few early career researchers are, and fewer still would have the necessary experience to maintain databases or web servers.

I noticed that researchers are increasingly writing reproducible manuscripts using markdown, and therein lay a compromise that promised the functionality we needed

without technical expertise. Markdown is a simple language and takes minutes to learn. It uses asterisks, underscores, and carets to make text ****bold****, *_italic_*, or ^{^superscript^}. Headings, URLs, and references are similarly easy, and free editing software makes writing markdown feel like writing a Microsoft Word document. Markdown converts into lots of other formats, including docx (Microsoft Word files), LaTeX, PDF, and HTML.

Many researchers already use tools like RStudio [194] or Quarto [195] to convert markdown into other formats. I decided to use Quarto because it is open source, has great documentation, and its functionality can be extended with other programming languages commonly used by researchers and statisticians, like Python or Ruby. Quarto takes plain text (markdown) and converts it into a web page. Academics who write using Quarto or R Studio will be familiar with the idea of “knitting” documents together: taking text with data, figures, analyses, and references, all saved as separate files, and compiling them into a single, publication-ready manuscript. The process of creating a web page is similar. I’ve created a platform that takes a guideline, stored as plain text across multiple files, and converts it into an interactive web page. These guideline files include reporting items, a glossary, meta data, introductory information, and FAQ information.

Quarto websites do not require a database or a server. This means they are free to host, easy to maintain, and have fewer data security concerns. Because all reporting guideline content is saved as plain text files, these files can be version controlled on Github, an industry-leading version control system commonly used by academics, and each version can be cited. This means guideline developers will be able to release incremental updates (e.g., to clarify the wording of a reporting item) and authors will be able to cite the exact version of the guidance they used.

I’ve made all code available on Github [196]. I’ve used Github Pages [197] to host the website because it is free, secure, beginner friendly, configurable, and integrates seamlessly with Github’s version control system.

9.4 Discussion

The aims of this chapter were to create an intervention component table by drawing on the results of all previous chapters, and to turn these components into a prototype ready to test amongst authors. The intervention component table brings together the outputs of all of my previous thesis chapters and links them together; each component comes from an idea (chapter 8), and addresses one or more influences and behavioural drivers (chapters 3 - 5) using intervention options ranked favourably by EQUATOR staff (chapter 7). My design process relied on this table and I used it again in my next chapter as an analysis framework. Hence this table, although relegated to the appendix because of its size, is a focal point for my thesis: it crystallises all preceding chapters, and formed the basis for all of my subsequent work.

The table includes components derived from all ideas pursuant to the intervention options EQUATOR staff and I prioritised in our workshops. Because I filtered out ideas *not* in line with our priorities, the list is not exhaustive. Stakeholders using the same set of initial ideas would create different components because of their different priorities and implementation opportunities. For example, a funder may have generated more components related to reporting guidance for applications or protocols, or relating to their application criteria and submission systems. Stakeholders with the power to grant approval (for funding/ethics/publication) may have components related to enforcement or reward. Consequently, although I hope my list will help readers understand the intervention changes I have made and why I have made them, I would encourage other intervention designers to go through this process themselves instead of using this list verbatim.

Throughout the intervention component table I have drawn comparisons between the proposed components and existing reality. Sometimes these comparisons are vague; I use terms like “some” or “rarely”. Where possible, I refer to images or examples. Some examples came from my qualitative synthesis (chapter 3), others arose organically in my workshops (chapter 7) when participants shared long-standing frustrations with reporting guidelines or EQUATOR’s website. Other times, after discussing a barrier or idea, we would look at a few guidelines to see how things are done currently. So my

comparisons were ad-hoc, and I have included them only to provide context to the proposed changes. I considered making this comparison formal by systematically coding BCTs employed by EQUATOR's website and popular guidelines. Ultimately, I decided this would not be helpful nor practical. With so many components and so many guidelines, this would have taken too long. Secondly, I doubted this audit would influence my subsequent design, but would merely quantify how my redesigned intervention differs to the current set-up. Should a quantifiable comparison be useful, it would be more appropriate to do it after the redesigned intervention has been refined in response to user feedback, once its design is stabilised.

Using this intervention component table I have created functional prototypes of a redesigned reporting guidelines and the EQUATOR Network home page. These new designs include 46 of the 62 intervention components I identified.

Using a framework and a systematic method helped the design advisory panel and I to make decisions based on evidence and theory, and to reduce the influence of our own subjectivity. Instead of relying on personal preference, we tried to ensure choices reflected the function we were trying to employ. For example, when choosing a background image, instead of asking "do you like this one?", the questions became "what feelings do you think this image conveys? Does it communicate simplicity?". Reflection was also useful when exploring disagreements between panel members. For example, when sketching layouts for the home page, some panel members drew a single, prominent search button. Others drew a plethora of options like "view guidelines by speciality", "view guidelines A-Z". Discussion revealed that whereas some staff prefer to search directly for what they want with laser-like focus, others prefer to explore, especially when they are not certain what they want or what the website is about. In this instance, the final design takes both use cases into account, but other times we resolved disagreement by referring to the intervention planning table or to similar websites. Hence using a framework and exploring disagreements as a group helped mitigate personal preferences.

However, many decisions required a degree of subjectivity and, as lead researcher, designer, and developer, often these decisions landed on my shoulders. I tried to mitigate

this by prioritizing other people's ideas over my own, and providing many opportunities for feedback. But the result undeniably has my "stamp". If someone else had built it using the same table of intervention components then some things might be the same (like simplifying the user journey from 5 steps to 2, or the conventional home page layout and search button position) but other things would look different (like the choice of wording and images).

My design may have benefited from input from other stakeholders. I describe how I obtained feedback from authors in my next chapter, but I would have liked to include authors, publishers, funders, and other stakeholders from the start of the design process. If EQUATOR decide to take my designs forward, these consultations could still take place, but they were not feasible within the time constraints of my PhD.

Input from user experience experts and graphic designers would also be useful. We found images to be a time consuming pain point. None of us had the skills to create professional looking graphics ourselves, and we found most free stock images were generic and did not communicate what we needed.

My experience of working with SRQR's lead developer, Bridget, was positive; she was supportive, liked the result, and she was interested when my process revealed gaps in SRQR item description. For example, often there was no guidance of what to write if a reporting item was not or could not be done. Some items did not explain why they are important and to whom. Filling these blanks required time and input that SRQR's development team were unable to give at present, and so I left these gaps unfilled for now. I anticipate other guidelines will have similar gaps. I hope other guideline developers will be as open-minded as Bridget was, but I expect others may feel less able or motivated to engage with a redesign, or may feel protective over their writing and resistant to change.

In addition to filling these gaps, making my redesign "live" would require further technical work. Some of these tasks are administrative and have no behaviour change impact, but there are still 16 intervention components outstanding. These components were too difficult or time consuming to include at this stage. For example, I intended to include more examples of reporting items, including examples of concise reporting

or transparently reported “imperfect” items. EQUATOR could add these intervention components at a later date. Other components may require significant design and development work, most notably creating ready-to-use checklists and templates.

If EQUATOR chooses to adopt these changes and apply them to other guidelines, hundreds of thousands of authors would access these redesigned resources each year. These redesigned resources have potential to benefit authors directly, and also to help other stakeholders. For example, publishers may find enforcing guidelines easier if my redesigned resources prove more user friendly. Guideline developers will benefit from having a ready-to-use dissemination platform based on evidence, with built-in channels for collecting feedback from authors. This feedback may help guideline developers refine their resources further, and could act as evidence to support future funding applications.

9.4.1 Conclusions

By linking components with barriers, functions, and behaviour change techniques, I have justified components using evidence and described how they are theorized to work. This table will help other intervention developers and stakeholders understand what changes I have made and why. It also helped me redesign the EQUATOR Network’s home page and the SRQR guideline. These new designs include 46 of the 62 intervention components I identified, and many of the remainder could be added later. Although my designs may have benefited from including other stakeholders, I explained how I facilitated open discussion, prioritised other’s opinions, and used my intervention component table to make decisions. In the next chapter I explain how I refined these designs further by interviewing and observing authors.

9.5 Reflections on this chapter

This was the hardest chapter to write. Initially I included it in my focus groups chapter, then I split it apart into separate chapters for defining intervention components and developing the intervention which I subsequently rejoined. I found writing difficult because the chapter did not obviously fit into a traditional research package. I had

objectives, but no research questions. I had an approach and outcomes, but they didn't feel like methods and results and when my approach was iterative it felt difficult to force it into a linear narrative. There were no reporting guidelines to help me and so instead I turned to the internet to seek general advice for structuring and writing academic text. I landed on Rachael Cayley's blog, *Explorations of Style* [198] and Helen Sword's book *Stylish Academic Writing* [199]. In particular, I found Rachael's description of *Reverse Outlines* as a technique for unpicking and restructuring messy text useful for revising this chapter. Reading these made me realise how little I knew about writing as an art and as a skill. I've never enjoyed writing but I had always thought I it was skill I had ticked and would never forget. But reading Rachael's blog and Helen's book made me realise how little I knew and how much I had forgotten since my only formal writing training in GCSE English 20 years ago! I'd forgotten the nuts and bolts of structuring arguments, paragraphs, or sentences. I'd forgotten the heuristics of clear writing, and developed an unhealthy reliance on passive verbs and relative clauses. When writing my early chapters I tried using a reporting guideline as a skeleton and I then filled in the blanks. I made no outlines, no structural scaffold beyond the guideline. Because I had no reporting guideline for this chapter I had lost my crutch, and my inexperience in structuring writing became apparent.

This realisation was interesting to me as it revealed barriers beyond reporting guidelines. If readers cannot understand my work it's probably not *just* because I failed to adhere to a reporting guideline, but also because I'm an inexperienced writer and my writing can be clunky and imprecise. When workshop members talked about educating authors about writing as a skill (chapter 7) the idea did not resonate with me until now, and I realise that the website I've created contains only superficial advice about writing in general. If I had reflected on my own struggles with writing before making it, perhaps I would have tried harder to address this problem in the redesign. That I am already reconsidering design choices points to two further realisations. Firstly, it's a reminder of the subjectivity that is inevitably part of any design process, despite my efforts to remain systematic and objective. Secondly, it reveals my hope that the website will be improved in response to evolving needs, context, evidence, and understanding, ideally derived from user feedback, which is the subject of my next chapter.

Part III

Refining the intervention

Chapter 10

Refining the intervention: interviewing authors to identify deficient intervention components

10.1 Introduction

Having defined intervention components and built a prototype (chapter 9) I wanted to refine the home page and SRQR guidance page (which I refer to as “the website” in this chapter) by getting feedback from authors.

The MRC guidance on complex interventions, the Person Based Approach, and the Behaviour Change Wheel all stress the importance of including service users in the design of complex interventions [145,146,150]. Service users can help identify deficiencies which, if addressed, would make the intervention more successful [145]. Involving service users can also help researchers better understand influences, and whether intervention components are functioning as intended.

In this study, I wanted to address limitations I had identified my thematic synthesis (chapter 3), where I found testing reporting guidelines with users is rarely done, and often suffers from thin description and unrepresentative participants lacking diversity. In contrast, I wanted to obtain rich data from a diverse group of authors. In chapter 9 I

positioned SRQR as a good guideline to test with users because its wide scope includes qualitative methods used by many researchers, from many fields, with varying levels of experience. Thus SRQR offered potential to recruit a diverse sample.

I did not aspire to perfect the website as I, like many in software developers, did not believe an optimal design exists. Variation between users, evolving trends, changing technological and contextual landscape means that a one-size-fits-all optimized design is unlikely [200]. However, although perfection does not exist, designs can none-the-less be improved through feedback and many designers use an iterative process. The purpose of this study was to collect evidence to inform future iterations. Instead of asking authors to suggest improvements (I suspect this strategy would have led to superficial suggestions or blank faces), I was more interested in identifying *deficiencies*. I defined a deficiency as any website element that, if modified, could better facilitate the website's intended target behaviour (for researchers to use reporting guidance as early as possible in their research pipeline). I used the term *deficiency* over and above *barrier* or *facilitator* because it includes both (if a facilitator could be improved, it is deficient). In the future, EQUATOR and I will be able to identify modifications to address these deficiencies.

10.2 Methods

The purpose of this study was to identify deficiencies in a redesigned reporting guideline and EQUATOR Network home page (“the website”).

My objectives were to:

- explore the experience of a diverse sample of authors, and
- to identify and understand deficiencies.

10.2.1 Sampling strategy

My purposive sample of authors varied in their:

- years of academic experience,
- subject area,

- their first language, and
- their country of residence.

This variation was important because my thematic synthesis (chapter 3) suggested inexperienced authors may have the most to benefit from reporting guidelines, but also face the most hurdles. Inexperience may be due to early career stage, being new to a field or study design, or new to academic writing. My synthesis also suggested language barriers could hinder adherence, and my service evaluation of EQUATOR's existing website (chapter 5) revealed a highly international user base.

Authors were eligible to participate if they were currently engaged in research utilizing qualitative methods, and if they were able to attend an online interview conducted in English.

I recruited through four channels:

1. I posted on X (called Twitter at the time).
2. I advertised through Penelope.ai [82], the manuscript checker I created before starting my PhD. Many medical journals offer the manuscript checker to submitting authors. BMJ Open is the largest of these journals, and enjoys a large, international, author base.
3. I invited researchers from a research consultancy in the Philippines.
4. I wrote to Chinese researchers who had published qualitative research, and I asked them to share my recruitment advert. One of the researchers I contacted posted the advert on internet forums used by Chinese students.

My recruitment advert is in appendix R. The advert said I was “looking to speak with people performing qualitative research about a new website”. It did not specify what the website was about, who it was for, or whether it would help with their job, because I wanted authors to be naïve when first viewing the website. This mimics the real world, where authors might be sent to EQUATOR’s website by a journal with only minimal information on what to expect from it.

All recruitment channels invited authors to signal their interest by email. To check applicants’ eligibility as qualitative researchers, I asked them to describe their research

methods in a few sentences over email. I excluded applicants if their descriptions made no reference of a qualitative method.

I sent all eligible applicants the participant information sheet and consent form. I used JISC Online Surveys [171] to obtain consent and ask the following demographic questions:

- How many years have you done research?
- Please describe your research in a couple of sentences
- What is your first language?
- What country do you work and live in?

I offered participants \$50 as reimbursement in return for an expected 2 hour commitment. This was delivered as an Amazon voucher to UK participants, and a bank transfer to international participants. My email templates and information sheet are in Appendices S and T.

Time and money limited my target sample size to 10 participants. As argued by Nielsen and Launder [201], small samples (fewer than 10) are often sufficient to identify the majority of deficiencies. In chapter 8 I introduced information power [162] as a concept to guide sample size in qualitative research, and I drew upon it again in this study. I maximised information power firstly by using methods to elicit rich information from each participant. Secondly, I used my table of intervention components (9) as an analysis framework. Hence I anticipated a sample of 10 to sufficiently inform at least one design iteration at the end of data collection.

10.2.2 Procedures

I wanted my study to resemble the way authors will experience the website in real life. Firstly, interview sessions took place online using Microsoft Teams. This meant participants could view the website on their own computer, using their normal browser, in their usual place of work. This would allow me to identify problems like slow-loading over bad internet connections or display problems on different screen resolutions, whilst avoiding complications stemming from asking a participant to use an unfamiliar computer or browser.

Secondly, I wanted to replicate the experience of encountering a new website as a naïve user, gradually exploring content, and then reading and applying guidance to one's own writing. I did this by using a variety of methods:

1. 5 second test to capture initial reactions [202]
2. Think aloud to capture real-time exploration within the interview session [203]
3. Plus-minus task to capture exploration between interview sessions [204]
4. A writing evaluation to explore interpretation [112]
5. Semi structured interviews throughout. [205]

I have outlined the order of data collection methods in Table 10.1.

TABLE 10.1: Data collection methods and when they occurred.

STAGE	METHOD
Session 1	<ol style="list-style-type: none"> 1. Five second test 2. Semi-structured interview 1 - prior experience of reporting guidelines 3. Think aloud 1 - the home page 4. Semi-structured interview 2 - the home page 5. Think aloud 2 - the top of the SRQR page 6. Semi-structured interview 3 - the SRQR page
At home, between sessions	<ol style="list-style-type: none"> 8. Plus-minus task 9. Writing using SRQR
Session 2	<ol style="list-style-type: none"> 10. Interview covering the plus/minus annotations 11. Interview covering the writing sample 12. Semi-structured interview 4 - closing thoughts

My interview schedule (appendix U) included the verbal instructions for each task and topic guides for each semi-structured interview. I tested the interview schedule by doing a mock interview with a student at Oxford University.

I began sessions by introducing myself as part of a team creating a new website. To encourage open and truthful feedback I reassured participants the best way they could

help was by being honest, not to worry about critiquing the website or offending me, and to share positive *and* negative feedback. I asked participants to tell me a little about themselves to relax into the interview and help them feel comfortable talking, before moving on to the first task: the 5 second test.

Five second test

Until this point, participants had no idea what the website was about. My recruitment materials and interview introduction made no mention of writing nor reporting guidelines, and so participants were unaware of the website's purpose.

The five second test is an “in the moment” survey method [202]. By sharing my screen, I showed participants the top of the home page for five seconds before removing it and asking questions. The test limits exposure to five seconds because although a participant can absorb much information (colours, words, shapes), five seconds is rarely sufficient to make sense of everything as a whole. The aim is to capture participant’s immediate reactions to salient design elements (like images, large words) before they have a chance to consider the content more critically. Furthermore, this five second limit was relevant because my website service evaluation found many authors leave EQUATOR’s website (chapter 5) within five seconds without interacting with it.

This test was appropriate for the top of the home page because, as per best practice, the area has little text, all relevant content is visible in one frame, and I asked few questions:

1. What do you think the website is about?
2. How do you think this website may affect your work?

If participants answered the first question with “reporting guidelines”, I asked “what do you think reporting guidelines are?”. If participant’s answers did not mention writing, I asked what stages of research they might use the website.

I designed these questions to explore three intervention ingredients: describe what reporting guidelines are, how they can be best used, and their benefits. These are the main ingredients featured at the top of the home page.

Semi structured interview 1 - prior experience with reporting guidelines

After the five second test it was no longer necessary to keep participants masked to the website's purpose, and so I asked participants about their prior awareness of, or experience with, reporting guidelines. I asked which guidelines they had used and what they had used them for.

Think aloud 1 - home screen

Website designers often ask participant's to "think aloud" as they complete a task or view a website as a way of exploring participants' thought processes [206,207]. Think aloud as a method was first described by cognitive psychologists Ericsson and Simon [208]. Their strict approach viewed verbalizations as "indicators of what information was heeded and in what order, a sort of time stamp of the contents of short-term memory" [203]. As user experience testers adopted the method, they used it more flexibly to additionally capture participants' thoughts, feelings, and expectations [203]. Whereas cognitive psychologists use the method to understand cognitive processes, usability testers use it to "support the development of usable systems by identifying system deficiencies". Because "building robust models of human cognition is not a central concern", Ericsson and Simon's strict approach is less appropriate, and usability testers use a more flexible and pragmatic approach to data collection and interpretation [203].

As per best practice [203] I began by explaining the task, and giving instruction to continually verbalize a train of thought. I then demonstrated by sharing my screen, opening up a different website, and "thinking aloud" for a minute. Participants then shared their screen as they explored the home page. Whenever participants stopped talking, I would prompt them to continue by asking "What are you thinking?". I acknowledged participants' verbalizations with neutral sounds like "uh-hu" and "mmm", which encourage further talk, but do not show agreement or disagreement. These verbalizations and prompts are also considered best practice [203].

Semi-structured interview 2 - home page

Once participants had explored the entire home page, I asked participants about any intervention components they had not talked about in the think aloud, about their overall

opinions, and whether their understanding had changed since first viewing it.

Think aloud 2 - SRQR guideline page

I asked participants to find the relevant guideline for reporting qualitative research, and then to continue thinking aloud as they explored it. The top of the SRQR page included information about the guideline, such as its scope, and the number of journals endorsing it. Participants thought aloud as they continued down the page. Because the SRQR guidance is long, I stopped participants from thinking aloud once they reached the first reporting item.

Semi-structured interview 3 - SRQR guideline page

I then used semi-structured interview questions to explore any intervention components in the introduction missed by the think aloud. Moving down to the reporting guideline itself, I asked questions to explore participants' expectations of four key features within in guidance: defined words (signified by a dotted underline), footnotes (signified by a superscript number), links to discussion boards (signified by an icon), and drop down content (signified by a chevron icon). I pointed to an example of each and asked participants what they expected to happen if they clicked on it.

This marked the end of the first interview session. I then explained the plus-minus and writing tasks participants needed to complete before the second session.

Plus-Minus task:

The full SRQR guideline is too long to cover in a single interview session, and I wanted to capture participant's experience of reading and applying the guidance in a realistic context, so I looked for methods that would allow me to collect data whilst participants read and applied reporting items in their own time, as part of their normal work.

In their review of methods to solicit text evaluations from readers, de Jong and Schellens [204] distinguish between evaluation goals: selection (whether readers will engage with the text), comprehension, application (being able to apply information in a real world setting), acceptance (including credibility), appreciation, relevance, and completeness.

I was not interested in selection (participants had no option other than to engage with the text), and my study scope did not extend to SRQR's relevance nor completeness. I *was* interested in comprehension, acceptance, appreciation, and application.

de Jong and Schellens describe methods to target comprehension, acceptance, appreciation in isolation, but because my interest included all three, I chose a nonspecific method that could explore them all; the Plus-Minus task. In this task, readers are asked to annotate a document with plus and minus signs to signify positive and negative reading experiences and then discuss annotations retrospectively.

I asked participants to select and annotate 2 or 3 reporting items relevant to whatever they happened to be writing up in the time between interviews. I created duplicates of the SRQR guidance page and gave participants unique URLs so they did not see each other's annotations. I used a web annotation tool called Hypothes.is [209]. Participants could optionally add comments alongside their plus and minus signs. Participants explained their annotations in the second interview.

As de Jong and Schellens note [204], the plus-minus method is advantageous over other nonspecific methods (like thinking aloud whilst reading) because it collects data without disturbing participants' natural reading process. Additionally, it was useful in this study as participants could make annotations in their own time, as part of their normal work pattern, and then discuss them retrospectively in the second interview.

Writing Evaluation

Although the plus-minus task will detect text that participants consider incomprehensible, it cannot detect whether participants comprehend guidance correctly or whether they are able to apply it to their writing. To address this, I used a writing evaluation. I asked participants to use the reporting items they labelled in the plus minus task when writing up their own research in between interview sessions, and to send me what they wrote before the second interview. I read the excerpts and noted reporting items (and sub items) as present or missing.

In the second interview, inspired by Davies et al.'s SQUIRE guidelines evaluation [112], I asked participants to identify parts of their writing pertaining to reporting items. When

I considered an item (or sub item) to be missing, I asked the participant whether they had reported this information. If they felt they had, I asked them to point out where, and then explored any misinterpretations. If they had *not* reported information, I asked why.

Semi-structured interview 4 - closing thoughts

To end the second interview session I asked participants to describe their experience of using the reporting guideline, and to share any final thoughts.

Methods explored different intervention components.

Each method targeted multiple intervention components. For example, in the 5 second test, participants could only see the top of the home page. The text, images, and design in this section are there to communicate what reporting guidelines are, when they can be used, and that they will benefit authors. These functions come from three intervention components defined in chapter 9:

- *Describe what reporting guidelines are where they are first encountered,*
- *Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for, and*
- *Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications).*

Components that required participants to read text could be best explored in the plus minus-task, and I hoped the writing evaluation would reveal how participants interpreted and applied instruction. I hoped the think aloud would capture opinions on salient features, and the semi structured interviews would allow me to explore remaining, un-noticed, features. In table Table 10.2 I detail the intervention components I expected each method to explore. The intervention components are defined in chapter 9, where I also list the website elements related to each component.

TABLE 10.2: Methods used and the intervention components they explore. Intervention components are defined in chapter 9

METHOD	INTERVENTION COMPONENTS (defined in previous chapter)
5 Second Test	<ul style="list-style-type: none">• Describe what reporting guidelines are where they are first encountered,• Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for,• Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications),• Include design, features, and language to foster trust

METHOD	INTERVENTION COMPONENTS (defined in previous chapter)
Think Aloud	<ul style="list-style-type: none"> Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for, Instruct authors to cite reporting guidelines so readers may learn about them, Links between related guidelines, Centralised hosting, Search function on website, Describe the scope of a reporting guideline at the top of every resource, Use if-then rules to direct authors to more appropriate and up-to-date guidance when available, Explicitly state when no better guidance exists for a use case, Provide translations, Make guidance appear shorter by removing superfluous information, hiding optional content, splitting long guidelines, using concise language, and separating design advice, Cater to different kinds of user (readers vs dippers) by structuring guidance with headings, itemisation, hyperlinking to particular sections, and with optional content, Include testimonials from researchers who were nervous about being punished for reporting transparently, Remove branding and messaging that may invoke feelings of judgement, complexity, or administrative red-tape, Reassure that all research has limitations to encourage explanation over perfect design, Educate authors about writing as a process, Link all resources to each other, Gather and communicate evidence for benefits, Include design, features, and language to foster trust, Create spaces for authors to discuss reporting guidelines with others, Use tone of voice and design to communicate personal benefits; confidence and simplicity,

METHOD	INTERVENTION COMPONENTS (defined in previous chapter)
Interview	<ul style="list-style-type: none">• Describe what reporting guidelines are where they are first encountered,• Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for,• Instruct authors to cite reporting guidelines so readers may learn about them,• Describe the scope of a reporting guideline at the top of every resource,• Include testimonials from researchers who were nervous about being punished for reporting transparently,• Address communications to authors,• Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications),• Create spaces for authors to discuss reporting guidelines with others,• Use tone of voice and design to communicate personal benefits; confidence and simplicity,• Include testimonials from research users who benefit from complete reporting,• Define key terms,• For each item, explain why the <i>information</i> is important and to whom (not just what constitutes “good” design),• For each item, provide clear instruction of what needs to be described,• For each item, provide examples of reporting in different contexts,• Structure guideline items to make them quicker to digest,• Tell authors when to use reporting guidelines, or that reporting guidelines are best used as early as possible,• Provide instruction as to how and where information can be reported without breaching word count limits or making articles bloated.,• Explain when reporting guidelines do not intended to prescribe structure,• Provide advice regarding how to respond if asked to remove reporting guideline content by a colleague, editor, or reviewer,

METHOD	INTERVENTION COMPONENTS (defined in previous chapter)
+/- test	<ul style="list-style-type: none"> • Decrease fear of judgement by making reporting guidelines design agnostic, • Use plain language, • Define key terms, • Provide links to other resources that explain how an item can be done, • For each item, provide clear instruction of what needs to be described, • For each item, provide examples of reporting in different contexts, • Structure guideline items to make them quicker to digest
Writing	<ul style="list-style-type: none"> • Use plain language,
Evaluation	<ul style="list-style-type: none"> • For each item, provide clear instruction of what needs to be described
Not Explored	<ul style="list-style-type: none"> • Search Engine Optimization, • Use consistent terms, • Provide clear instruction of what needs to be described when an item was not done, could not be done, or does not apply, • Present design advice separately to reporting advice

I did not attempt to explore two intervention components. I did not expect participants to know about or comment on search engine optimization, especially as a large amount of optimization occurs in the website meta data and is thus invisible. Secondly, because I did not want to edit the *meaning* of the SRQR guidelines (just its layout), I did not want to add instructions about what to report when an item was not done, could not be done, or does not apply.

10.2.3 Data processing and analysis

I recorded video and audio transcriptions using Microsoft Teams. Because automatic audio transcription was not always accurate, I corrected them by rewatching the videos. I de-identified transcripts by replacing names with participant codes before importing them to NVivo for coding [170].

As with my focus group data in chapter 8 I used qualitative description to aggregate and summarise ideas [168,169]. I used my intervention component table (see chapter 9 and appendix P) as a framework to code transcripts line by line. I did this deductively; whenever a participant said anything about a component, I coded text to that component. Because some website features implemented multiple components (for example, an image can both educate and persuade), I sometimes coded text to multiple components. In this way, I created categories of codes, and each category was an intervention component. I gave codes equal weight across all methods (5 second test, think aloud, interviews e.t.c.), and mapped data from all methods onto the same framework.

Once all transcripts were coded, I grouped codes within categories into deficiencies. If a single component was deficient in multiple ways, I created a code group for each deficiency. If there was disagreement about a deficiency (e.g., some people disliked a component, but others liked it), then I created sub-groups within each deficiency. Although positive feedback did not directly address my objective of identifying deficiencies, I kept these codes because they provided context and counter-evidence to deficiencies.

Some participants spontaneously suggested modifications. In these instances, I coded the proposed modification and the underlying deficiency. Because some participants spontaneously shared prior experiences using reporting guidelines I coded these using

my list of influences from step 4 of chapter 7 as a framework. I decided to create new codes for any influences not previously identified.

In this way, I ended up with a list of deficiencies (my primary unit of analysis), and incidental lists of influences and possible modifications.

10.2.4 Reflexivity and Trust

As with my focus group chapter 8 I tried to remain objective during interviews, and I considered research paradigm to be post-positivist: my role was to identify deficiencies from what participants said, but I acknowledged that my own experience, perspective and opinion may affect what I observed, understood, and concluded.

As with previous chapters, I used a number of techniques to ensure credibility, transferability, dependability, and confirmability [156]. I describe these in Table 10.3.

TABLE 10.3: Techniques used in this study for establishing trustworthiness. Based on Lincoln and Guba's Evaluative Criteria [156]

TECHNIQUE	IMPLEMENTATION
Techniques for establishing transferability	<p><i>Thick description</i></p> <p>I aspired to report my results with context by indicating when ideas were common or rare, and who they originated from when I felt this was particularly relevant. I reported disagreements, provide quotes, and relationships between ideas.</p> <p>Interview sessions were long and used multiple techniques to elicit lots of data. I reported findings alongside relevant context, including participant demographics, and used context to reconcile disagreements.</p>
Techniques for establishing confirmability	

TECHNIQUE	IMPLEMENTATION
<i>Audit trail</i>	I referred to video recordings when I needed to clarify parts of the transcript. I kept all raw data, and a record of modifications made.
<i>Reflexivity</i>	<p>I kept a diary during data collection to note of ideas and my own feelings during. Because I created the website being tested, I felt it was important to reflect on any feedback that made me feel defensive, frustrated, or that I did not understand. In my previous experience, these moments of conflict are important as they often hint at a latent misunderstanding or deficiency.</p> <p>Receiving negative feedback does not always hurt me. For example, sometimes I expect negative feedback because it reflects a limitation or trade-off that I already know about. Other times negative feedback can feel like an “aha” moment, as I discover a problem I immediately understand, agree with, and can see a solution to. In contrast, when feedback feels <i>bad</i>, in my experience that is because I have misunderstood something, and my internal model of the situation is off. Although I found it easy to remain professional and neutral within the interviews, having a “thick skin” is not enough. In these moments of friction I made sure to delve deeper into the issue with the same participant or future ones.</p>

Techniques for establishing credibility

<i>Negative case analysis</i>	I purposefully explored negative feedback that I found unexpected or challenging (see reflexivity).
-------------------------------	---

TECHNIQUE	IMPLEMENTATION
<i>Peer debriefing</i>	CA acted as a disinterested peer throughout design, data collection, analysis and reporting. She questioned my reasoning she helped me become aware of biases, potential flaws, and assumptions I was making.

10.2.5 Ethics

Oxford University's Medical Sciences Interdivisional Research Ethics Committee deemed this study to be a service evaluation, and so judged ethical approval unnecessary.

10.2.6 Reporting

I used SRQR [30] when outlining this chapter, and again to check my reporting during revision (see appendix Q).

10.3 Results

10.3.1 Recruitment

I recruited participants between 21/03/2023 until 9/08/2023. The number of people who expressed interest, were eligible, consented, and participated are shown in Table 10.4. Eleven people participated. Two dropped out before the second interview, without giving a reason. Participants' characteristics are summarized in Table 10.5 and included variety in research experience (from 1 to 10+ years), subject area, country of origin, and first language. Six participants had never heard of reporting guidelines before – one had, but did not remember which. Three others had used a reporting guideline before, and one had used many reporting guidelines before. The first interview lasted between 45 minutes - 1.5 hours, and the second interview lasted 30-45 minutes.

TABLE 10.4: Recruitment and drop out of participants through different channels

CHANNEL	CONTACTED	INVITED						COMPLETED	COMPLETED
		No. PEOPLE EXPRESSED	INTEREST	ELIGIBLE	TO CONSENT	CONSENTED	INTERVIEW 1		
Penelope.ai	Unknown (23 were excluded because they did not want to attend an online interview conducted in English. A further 78 did not describe using qualitative methods when asked to describe their research).	144	43	43	13	8	7		
X	Unknown	1	1	1	0	0			

(1 did not reply)

CHANNEL	CONTACTED	INVITED						COMPLETED INTERVIEW 1	COMPLETED INTERVIEW 2
		No. PEOPLE EXPRESSED	INTEREST	ELIGIBLE	CONSENT	CONSENTED			
Email invitation	Unknown (2 emails sent, but were forwarded to an unknown number of recipients)	4	3	3	3	3	3	2	
Total	Unknown	149	47	47	16	11		9	

TABLE 10.5: Participant characteristics

ID	JOB TITLE	SUBJECT AREA	RESEARCH		PREVIOUS EXPERIENCE	
			EXPERIENCE (YEARS)	FIRST LANGUAGE	COUNTRY OF ORIGIN	WITH REPORTING GUIDELINES
1	Research consultancy	Health policy	4	English	Philippines	None

ID	JOB TITLE	SUBJECT AREA	RESEARCH			PREVIOUS EXPERIENCE WITH REPORTING GUIDELINES
			EXPERIENCE (YEARS)	FIRST LANGUAGE	COUNTRY OF ORIGIN	
2	Medical student	General qualitative medical research	1	English	Ghana	Had used PRISMA
3	Academic researcher	Clinical psychology and public health	7	Spanish	Ecuador	Had used COREQ
4	Academic researcher	Physiotherapy	10+	English	UK	Could not remember
5	Academic researcher	Medical ethics	7	English	India	None
6	Midwifery student	Sexual and reproductive health	3	Lango	Uganda	None
7	Academic researcher	Environmental Health	10+	English	South Africa	Had used JARS

ID	JOB TITLE	SUBJECT AREA	RESEARCH			PREVIOUS EXPERIENCE	
			EXPERIENCE	FIRST	COUNTRY OF	WITH REPORTING	GUIDELINES
8	Academic researcher	Physiotherapy	10+	English	Australia	Had used many reporting guidelines before	
9	Pre PhD student	Public health	1	Chichewa	Malawi	None	
10	PhD student	Child development	7	Chinese	China	None	
11	PhD student	Child development	4	Chinese	China	None	

10.3.2 Design Iterations

I had originally planned to finish data collection before making any changes to the website. However, the first five participants consistently mentioned similar deficiencies. After reflecting and discussing with UK EQUATOR staff, we agreed these deficiencies would likely affect many authors and diminish the website's success, and so we decided to iterate our design.

Briefly, the changes we made included:

- Editing the text at the top of the home page for clarity and to emphasise benefits. I edited it from “Writing research, made simple. Write confidently using guidelines created by the research community” to “Want help writing up research? Reporting guidelines help you describe research quickly, confidently, and completely.”
- Adding images and colour to make the home page more attractive and to convey meaning of the accompanying text (see Figure 10.1).
- Adding publisher logos to foster trust
- Reorganising the introduction to SRQR to make it appear shorter.

The results below include quotes and discussion pertaining to these changes, along with all other deficiencies.

10.3.3 Main findings

I identified 53 deficiencies. Appendix V lists the deficiencies identified for each intervention component with supporting quotes. I have chosen to describe a few deficiencies in more detail, either because I deemed them important, they were frequently mentioned (and, therefore, salient), or involved multiple parts of the website. I then provide a summary of the remaining deficiencies.

Deficient component: Describe what reporting guidelines are where they are first encountered

Relevant website features: Prominent definition on home page and guideline page.

Influences addressed: Researchers may not know what reporting guidelines are

The figure shows the home page of the EQUATOR Network website before and after initial revisions based on feedback from the first five participants.

(A) Before:

- Header:** The header includes the EQUATOR logo, Reporting Guidelines, Training, Other Resources, and About EQUATOR.
- Main Content Area:**
 - Section 1:** "Writing research, made simple". Subtext: "Write confidently using guidelines created by the research community". Call-to-action: "Search guidelines".
 - Section 2:** "Frequently accessed reporting guidelines". A table lists reporting guidelines across categories: Observational (STROBE Cohort, Case-control, Cross-sectional), Reviews (PRISMA, MOOSE), Randomised Trials (CONSORT, SPIRIT), Qualitative Research (SRQR), Case Reports (CARE), Diagnostic Test Accuracy Studies (STARD), Prediction Models (TRIPOD), Animal Research (ARRIVE), Quality Improvement in Healthcare (SQUIRE), and Economic Evaluations (CHEERS).
 - Section 3:** "Want help writing up research?". Subtext: "Reporting guidelines help you describe research quickly, confidently, and completely.". Call-to-action: "Search guidelines" and "What are reporting guidelines?".
 - Section 4:** "Frequently accessed reporting guidelines". A table lists reporting guidelines across categories: Observational (STROBE Cohort, Case-control, Cross-sectional), Reviews (PRISMA, MOOSE), Randomised Trials (CONSORT, SPIRIT), Qualitative Research (SRQR), Case Reports (CARE), Diagnostic Test Accuracy Studies (STARD), Prediction Models (TRIPOD), Animal Research (ARRIVE), Quality Improvement in Healthcare (SQUIRE), and Economic Evaluations (CHEERS). Call-to-action: "Search all 500+ reporting guidelines".
 - Logos:** Logos for SPONSORING PUBLISHERS (Wiley, Elsevier, BMJ, BMC), ENDORSED BY LEADING PUBLISHERS (SPONSORING PUBLISHERS), and ENDORSED BY LEADING PUBLISHERS (BMJ).
- Testimonials:**
 - "Reporting guidelines make writing so much easier and help me develop as a researcher. I use one for every article I write!" - Alex Harvey, Edinburgh University
 - "Reporting guidelines make writing so much easier and help me develop as a researcher. I use one for every article I write!" - Alex Harvey, Edinburgh University
- Helpful Sections:**
 - Recommendations to help describe your work clearly:** Subtext: "Reporting guidelines don't advise how to do research, but how to describe it completely so it can be used by everyone." Includes sections for Writing manuscripts, Checking manuscripts, and Planning research.
 - Reporting guidelines help authors and readers:** Subtext: "Guidelines make writing research easier, and transparent research leads to better patient outcomes." Includes sections for Easier writing, Smoother publishing, and Maximum impact.
 - Helping authors and readers:** Subtext: "Reporting guidelines make writing research easier, and transparent research leads to better patient outcomes." Includes sections for Easier writing, Smoother publishing, and Maximum impact.
 - The EQUATOR Network:** Subtext: "EQUATOR stands for Enhancing the Quality and Transparency Of health Research. Our mission is to improve the reliability and value of published health research literature by promoting transparent and accurate reporting and wider use of robust reporting guidelines. We are the first coordinated attempt to tackle the problem of inadequate reporting systematically and on a global scale." Includes a photo of a group of people and a list of centers in the US, Canada, Australia, France, and China.
 - Footer:** Created by the UK EQUATOR Centre. Links: About, License, Privacy Policy.

(B) After initial revision:

- Header:** The header includes the EQUATOR logo, Reporting Guidelines, Training, Other Resources, and About EQUATOR.
- Main Content Area:**
 - Section 1:** "Ready to use a reporting guideline?". Call-to-action: "Search for guidelines".
 - Section 2:** "Want help writing up research?". Subtext: "Reporting guidelines help you describe research quickly, confidently, and completely.". Call-to-action: "Search guidelines" and "What are reporting guidelines?".
 - Section 3:** "Frequently accessed reporting guidelines". A table lists reporting guidelines across categories: Observational (STROBE Cohort, Case-control, Cross-sectional), Reviews (PRISMA, MOOSE), Randomised Trials (CONSORT, SPIRIT), Qualitative Research (SRQR), Case Reports (CARE), Diagnostic Test Accuracy Studies (STARD), Prediction Models (TRIPOD), Animal Research (ARRIVE), Quality Improvement in Healthcare (SQUIRE), and Economic Evaluations (CHEERS). Call-to-action: "Search all 500+ reporting guidelines".
 - Logos:** Logos for SPONSORING PUBLISHERS (Wiley, Elsevier, BMJ, BMC), ENDORSED BY LEADING PUBLISHERS (SPONSORING PUBLISHERS), and ENDORSED BY LEADING PUBLISHERS (BMJ).
- Testimonials:**
 - "Reporting guidelines make writing so much easier and help me develop as a researcher. I use one for every article I write!" - Alex Harvey, Edinburgh University
 - "Reporting guidelines make writing so much easier and help me develop as a researcher. I use one for every article I write!" - Alex Harvey, Edinburgh University
- Helpful Sections:**
 - Recommendations to help describe your work clearly:** Subtext: "Your research is used by people from different disciplines and backgrounds for decades to come. Reporting guidelines are community-created recommendations for describing research so that everyone can understand, repeat, apply, and synthesise your work." Includes sections for Writing manuscripts, Checking manuscripts, and Planning research.
 - Reporting guidelines help authors and readers:** Subtext: "Reporting guidelines make writing research easier, and transparent research leads to better patient outcomes." Includes sections for Easier writing, Smoother publishing, and Maximum impact.
 - Helping authors and readers:** Subtext: "Reporting guidelines make writing research easier, and transparent research leads to better patient outcomes." Includes sections for Easier writing, Smoother publishing, and Maximum impact.
 - The EQUATOR Network:** Subtext: "EQUATOR stands for Enhancing the Quality and Transparency Of health Research. Our mission is to improve the reliability and value of published health research literature by promoting transparent and accurate reporting and wider use of robust reporting guidelines. We are the first coordinated attempt to tackle the problem of inadequate reporting systematically and on a global scale." Includes a photo of a group of people and a list of centers in the US, Canada, Australia, France, and China.
 - Footer:** Created by the UK EQUATOR Centre. Links: About, License, Privacy Policy.

FIGURE 10.1: The home page before and after initial revisions based on feedback from the first five participants.

Because it is important for website visitors to quickly realise the site contains resources for *writing up* research articles, as opposed to *designing* or *appraising* studies, I used the 5 second test to explore what participants understood the website to be about upon first impression. This was the first time participants saw the website. Before then, they had no idea what it would be about.

In the first iteration the heading was “*Research articles, made simple*”, but some participants thought this was about *reading* or *explaining* research articles as opposed to writing them. In the second iteration, we changed this heading to “*Want help writing up research?*”.

On immediate impression, some participants quickly realised the website was about writing, but some did not, or thought it was about methodological guidelines

All participants realised the website was about research. Some researchers realised the website was about writing within 5 seconds:

“From what I have seen, I think probably the website should be about, uh, helping you try to discover...identify the guidelines that you will use for writing your study quickly.” (Midwifery student from Uganda)

“how to go about writing something” (ECR from India)

However, other participants gleaned only vague understandings like “support for doing research” (PhD student from China) or “guidelines of some sorts, I think in relation to research” (Pre-PhD student from Malawi), and one expected the website to be about methodological guidance:

“methodology guidelines one can use or follow when conducting research”
(Researcher from the UK)

Participants with previous experience using reporting checklists realised the website might be about reporting guidelines:

“Well, I hadn’t seen the guide that I’m familiar with (that is the COREQ). But I think the other guides also are like COREQ. So that’s what comes to my mind.

So I think it's a [website] where you are going to find all the checklists or the guides for all the [...] final stages of the research when we are, like, writing the paper, just to [...] double check that everything has been included.”
(ECR from Ecuador)

“That’s the EQUATOR guidelines, so that’s... consensus... expert consensus-developed guidelines for the reporting of different research.” (Researcher from Australia)

“[I read that] you can use reporting guidelines to try and help you [write research articles] more efficiently or quickly. And then I was thinking about, *what the heck are reporting guidelines?* And then I think it might be stuff like STROBE or like those checklists things or PRISMA, if you’re doing a systematic review or something. And that’s all I got” (Researcher from South Africa)

Given a few more seconds to explore the website on their own (during the think aloud task), all participants realised the guidelines were for writing and others gained more insight into what to expect from reporting guidelines.

“after reading this sentence I think you want to give me a framework, a framework about writing. Is that right?”

“I think the guideline would, uh, would clearly state the different sections of the research report or the manuscript and then whatever is required under a section like maybe under methods. Like what are the nitty gritties required under method.” (Midwifery student from Uganda)

Some participants found later, longer descriptions more informative. Referring to content half way down the landing page, one participant said:

“Why couldn’t this be further up? Why can’t that be at the top and then the that stuff here follow? Because then I’d have a better idea of what this is about.[...] I would have liked to have read this at the top and I would have known straight away what this whole website was about.” (Researcher from South Africa)

Deficient component: Include design, features, and language to foster trust

Relevant website features: Professional design. EQUATOR's Logo remains prominent. Citation metrics are presented at the top the reporting guidance. Information about who developed the guidelines, how they developed it, and why the guidance is credible is still provided, and easily findable from the top of the guidance.

Influences addressed: Researchers may not believe stated benefits

When participants talked about trust, they mentioned whether the website came across as professional, credible, and believable. This intervention component is complex because content and design throughout the entire website influenced judgements regarding trust. In particular, participants wanted to know who made the website and why they could be trusted, and identified some design elements that could look more professional.

EQUATOR's introduction could be more prominent

Apart from its logo, EQUATOR was not mentioned at the top of the page. Participants who already knew about EQUATOR said that its brand lent credibility:

“and then I picked up the top my top left hand corner with the EQUATOR logo so it seemed from reputable source.” (Researcher from the UK)

“I already trust the website because I saw that... like this is legit and I see the credentials from EQUATOR network” (ECR from Ecuador)

Participants unfamiliar with EQUATOR expressed wanting to know who developed the website. Whilst looking at the top of the home page, one participant said:

“I really don't get an idea of [...] who's responsible for the website [...] would I trust the developers of the website?” (Pre-PhD student from Malawi)

The home page introduced EQUATOR at the very bottom. Participants recommended moving this introduction (or parts of it) up to the top, using an updated photo, and adding EQUATOR's affiliations and awards.

“So most places put the *about* stuff at the bottom and I would have liked to have seen what [EQUATOR stands for] explained right at the top.” (Researcher from South Africa)

“And then this is the thing on the bottom that I want to look at on every website... because I want to see if they have an actual office. So usually I click this “about us” first” (Research consultant from the Philippines)

The site’s design could be more professional

I tried to create an aesthetic that would make the website appear simple. The first iteration was apparently *too* simple and one participant explained how its simplicity made it less trustworthy:

“it looks kind of like a blog [...] it’s a basic website” (Research consultant from the Philippines)

“I wouldn’t say [it looks] particularly trustworthy, but not particularly suspicious either. Kind of in the middle [...] something that will be more trustworthy will be something which is more sophisticated because I know, ‘OK, this is someone who actually took his time to do a lot of work... put a lot of work in designing it’. Most of the time if it is a fake website, it’s usually much more simple.” (Medical student from Ghana)

One participant viewed both the first and second iterations of the home page (their second interview session occurred after the iteration), and they described the second iteration as better because “It’s like more trustable [...] Scientific. Evidence based. Yeah, of course, legit.” (ECR from Ecuador)

However, one participant still questioned the second iteration’s simplicity and trustworthiness, and drew a comparison with another website that she *did* trust:

“So I’m saying it’s kind of basic, that the format itself is kind of basic [...] [When] I’m looking for information on PUB Med, just the outlet itself gives you the picture that, you know, somehow you can trust it. You know it looks as if there was more work put in it.” (Pre-PhD student from Malawi)

Logos lend credibility and could be more prominent

Participants noticed that the first iteration had no logos:

“I don’t know if this is just me, but I kinda want some logos. So I know who will vouch for [the website] right away. Like, usually [...] there’s some, like, other medical societies that are, like,”We we are on the EQUATOR network”
(Research consultant from the Philippines)

I added logos to the second iteration’s home page to show publishers endorsing reporting guidelines. All participants liked these, but some suggested they could appear at the top of the home page so they are immediately visible.

“Leading publishers....Wow, this is good...Nature. Really? Elsevier, BMJ. Yes, this is good. And this brings some sense of trust and authenticity in the website.” (Midwifery student from Uganda)

“[The publishers’ logos are] encouraging, because these are all publishing houses with mostly reputable journals, probably all reputable journals. [...] You know, if these were higher up, then [...] that would have made me feel a little bit more like ohh, this is good.” (Researcher from South Africa)

Numbers showing reporting guideline endorsements and citations lend credibility, but may not be intuitive

The top of the SRQR guideline page included widgets displaying the number of journal endorsements, and the number of times the reporting guideline had been cited. Some people commented that this information lent credibility:

“I think it is authentic. It’s robust. If it was endorsed by many journals and developed by experienced researchers, if I use it, maybe I’ll get a better quality work.” (Midwifery student from Uganda)

“I think citations here might be some people or some people’s work who has cited this page. So this this button [...] might show people how many other words use this page.” (PhD student from China)”

“...understand [...] that the SRQR guidelines is something that's already widely used.”

“I didn't pay attention before, but I think I like it (the citation information) [...] if it is more cited, I think, like, I will believe it. I will believe it, like, much better, and also like the journal endorsements” (PhD student from China)

However, not everybody understood what these numbers meant.

“Is this [widget] telling me [something], or is it what I am supposed to click on?” (Pre-PhD student from Malawi)

“I was a bit confused there, OK” (Pre-PhD student from Malawi)

“I think the citation tab over here... What is the relevance of it? I mean, why I'm seeing that?” (ECR from India)

Others were not sure whether the citation information pertained to the website or an underlying article.

“I'm not sure if [it is about], you know, the website or connected paper.” (PhD student from China)

“Has it been cited 4000 times? I don't understand that.” (Researcher from South Africa)

Not everybody considered the image at the top of the home page to be trustworthy

For the second iteration, I added an image to the top of the home page comprising of three icons to represent the process of writing a manuscript. One participant described this image as a “bit naff [...] I actually think this [image] reduces [the website's] score on the first impressions of trustworthiness kind of thing. Just because [the icons making up the image] are so, umm, ubiquitous, and, uh cheap?” (Researcher from Australia)

Deficient component: Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications)

Relevant website features: Benefits are prominently and consistently displayed across the home page and guidance pages. Descriptions prioritize personal benefits to the authors above hypothetical benefits to others.

Influences addressed: Researchers may not know what benefits to expect

Benefits are clear, but could be communicated quicker

I wanted website visitors to immediately expect the website to benefit them as researchers and authors. The website headline is one of the first things visitors see, and so was an important feature for communicating benefits.

All participants talked about benefits or help. None talked about the opposite (e.g., rules, requirements, or red tape). In the 5 second test, participants generally talked about *help* in a general sense:

“I see research and writing. So i’m thinking. This is to help me with something with my job.” (Research consultant from the Philippines)

“so it’s going to assist me in research. It’s going to help me somehow. Make things easier for me.” (ECR from Ecuador)

Under the headline, I included a short statement: ‘*reporting guidelines help you describe research quickly, confidently in completely*’. Two participants did not find this brief text meaningful in the five second test:

“I tried to read the sub headline just below the biggest one and it says help you blah blah confidently and blah blah. [...] So I think this information maybe be meaningless for me [...] because it sounds like it didn’t provide some concrete information. It’s just a sentence that tried to cheer me up.”
(PhD student from China)

“And then there’s something to do with ‘*reporting guidelines help you describe research quickly, confidently in completely*’. This information is not really telling me much” (Pre-PhD student from Malawi)

However, they seemed to understand the reported benefits after reading the top of the home page in more detail, and after viewing the section below where benefits are stated more clearly.

“OK, I think this part is very great because when I see something like “*easy writing*”, “*smoother publishing*” I think “*Ohh, that’s great. That’s what I want.*”” (PhD student from China)

“Participant: Now I’m getting a sense of what the website is about. Now looking at the things down here...

Interviewer: OK.

Participant: I’m getting that it might be a useful resource. That actually, umm, because these are, I think, for... early career researchers like me, I’d say I’d be very interested to come in and see this.” (Pre-PhD student from Malawi)

Participants seemed to understand how reporting guidelines might make publishing “smoother”.

“it gives me an impression that maybe this website will help me write my work easily and it will also help me increase the chance of my work getting published [...] Umm, just aligning myself to this standard that already many people use. And hopefully, In doing that, I’ll be up to standard and then I won’t stress myself too much later.” (Midwifery student from Uganda)

Making a distinction between benefits to authors and readers may lead to confusion about the intended user

Further down the home page, the section title ‘*Helping authors and readers*’ made one participant believe the website also hosts resources for readers.

“So this is a bit weird. So is the point here that this one is for the writers. And now it’s saying, OK, but we can also help readers. OK, I suppose that’s interesting” (Researcher from South Africa).

The images depicting benefits could be clearer

One participant said the icons describing writing and impact were appropriate (a blank page and an award, respectively), but the image depicting “smoother publishing” was not intuitive.

“looking at that icon, it doesn’t really tell me anything about smoother [publishing].” (Researcher from Australia)

Deficient component: Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for

Relevant website features: Clear instruction and differentiation of resources

Influences addressed: Researchers may not know what reporting guidelines are; Researchers may not know when reporting guidelines should be used

Tools for drafting and checking were mostly intuitive, but could be more prominent.

Half way down the home page, a section described how to use templates and checklists to draft and check manuscripts. Participants seemed to find this intuitive and appealing:

“I like that: different stages and different tools” (ECR from Ecuador)

“Yeah, writing templates is something I’ve recently come across, and I think that might be useful. I’ve tried it a bit when writing abstracts. And I guess, yeah, that would be something I’d be interested in looking into further.” (Researcher from the UK)

However, describing these tools further up the page might help visitors “get” what RGs are about

“[it] would actually be very good to appear [higher up the page] because then it would now start opening up one’s understanding as to exactly where this kind of guidelines might be applied.” (Pre-PhD student from Malawi)

“If any of these things: *writing research*, *checking manuscripts* and *planning research*, if these can be consolidated on [the top of] your landing page somewhere [...] it might be beneficial because my thought process is that I need to know what I’m doing and only then reporting guidelines can help me, right? So if I know that this website is gonna help me with writing the manuscripts, checking [...] I think then, reporting guidelines can make a logical progression in that particular case?” (ECR from India)

Even though participants could not download templates or checklists (there were buttons, but clicking these did not trigger a download), they had expectations of what these resources might look like.

“The checklist could be a pre-populated document that I can go through..it may have a table that I could go through as a tick box exercise ticking which of the [guideline reporting items] my study includes.” (Researcher from the UK)

“[Regarding templates] I would like to adjust this template by myself. Just like a semi structured interview. I don’t want this template be a structured interview. I want it to be semi structured so I can have the space to adjust it.”
(PhD student from China)

Using a reporting guideline “for planning” was not intuitive

The same section described how to use reporting guidelines when planning or conducting research. The SRQR page had a button to download a “log book” where researchers could document the decisions and data they would later need to report. However, in contrast to the checklist and template, no participants understood what this log book might be:

“it’s not immediately intuitive what a log book might be” (Midwifery student from Uganda)

“OK, what’s a log book? Don’t know.” (Researcher from South Africa)

“nothing has been mentioned about the log book overhead. How [am I] gonna use the log book? Maybe you have mentioned about the template checklist, but uh, maybe we can add [something about] the log book.” (ECR from India)

The word “planning” was not intuitive either. One thought this meant planning a manuscript, and so confused the purpose of the log book with that of the template. Another interpreted it as guidance for writing a research proposal.

“When I’m reading planning research, I think maybe I have already done this when I design my own outline.” (PhD student from China)

“To write can help you plan a study. Yeah, I guess it’s... I I would have thought this might be useful if you’re writing a research grant or a research proposal.” (Researcher from the UK)

I had ordered the tasks and tools as “drafting”, “checking” and then “planning”. I put planning at the end because it is the least conventional way to use a reporting guidelines. Some participants questioned this ordering:

“Why is planning at the end? You have to plan first before you write.” (Researcher from South Africa)

Deficient component: For each item, provide examples of reporting in different contexts

Relevant website features: SRQR already had some examples. No more examples added

Influences addressed: Researchers may not know how to report an item in practice

Many participants said they wanted more, varied examples

Each reporting item in SRQR comes with one or more examples from published literature. All participants stressed the usefulness of these examples. The attention examples received was notable because I did not ask about them; all comments about examples came spontaneously from the participant in the think aloud and plus minus tasks.

“Oh, and you have examples that could be really helpful, yeah.” (ECR from Ecuador)

“the most important thing I would say is the examples” (Midwifery student from Uganda)

“I found [this section] very useful as they give more detailed explanations on each specific section, and particularly the examples.” (Medical student from Ghana)

However, many participants said they wanted more examples, and greater variation in style, length, and conciseness.

“need further explanation and examples” (Midwifery student from Uganda)

“it would have been useful or helpful for me to have more than just one example.” (Pre-PhD student from Malawi)

“illustrations of how to [report] this information in a concise manner, that would be very helpful as well.” (Pre-PhD student from Malawi)

“you may as well put a whole discussion in there, or at least just sort of three or four paragraph discussion” (Researcher from South Africa)

“And here if there could be more examples, because when I [...] started to read through and understand the PRISMA guidelines and use the official explanation file to try to understand exactly what I’m required to write about and the examples particularly helped me a lot.” (Medical student from Ghana)

Many participants wanted examples from their own field, or even entire publications that have used the guideline:

“So I want something more relatable, [because] when I was reading these examples they were not relevant to my work.” (Researcher from South Africa)

“...if you can list some [...] papers who use the SRQR, you can put it here.”
(PhD student from China)

“I’ve searched on pubmed to find an example of a research article that’s used these standards so I could copy or check how they’ve laid it out, which sub-headings they’ve used.” (Researcher from South Africa)

Examples may be less credible if they are old or not referenced

In the original SRQR publication, all examples are referenced. I had not included references for the examples when putting them onto the website because of time constraints. This bothered one participant.

“Well, there’s no references there, so is that a very good example? No, I don’t know the source of these things. [...] You know, I don’t know, where did it come from? Where’s the reference?” (Researcher from South Africa)

When I asked about hypothetically labelling examples as illustrative if they were made up, the participant said “Yeah, I guess that would be alright”.

The same participant also noted that an example was quite old (10 years).

Examples could be more useful if explained or annotated

Because some reporting items contain multiple sub-items, one participant said annotating examples may be helpful. Taking a discussion item about transferability and integrating findings, they suggested “if you could underline or maybe indicate [in the example] that *this* [sentence] is now how they are trying to say the result can be transferable, umm, this is another [sentence] trying to say how they’re trying to integrate.... something like that.” (Midwifery student from Uganda)

10.3.4 Other findings

Many intervention components focussed on structuring guidance to make it appear short, navigable and digestible. All participants liked the structure, and said it made the guidance “more easy to follow” (ECR from Ecuador) compared to the original reporting guideline. Nobody disliked content hidden in collapsible content, but some felt the guidance still appeared very long. One suggested presenting items on separate web pages instead (Researcher from the UK) and another requested a summary (Research consultant from the Philippines). A few participants suggested making the item headings and side-navigation menu more prominent.

All participants realised that the website was aimed at researchers. This is perhaps not surprising as my study advert said I wanted to speak with qualitative researchers “about a new website”. Even though I did not specify who the intended audience was, it was implied. Indeed, some participants had assumed the website was aimed at *qualitative* researchers (I had specifically advertised for qualitative researchers because of our test guideline), suggesting their perception of target audience had been influenced by my recruitment materials. However, these same participants voiced that the website felt “more open” (Pre-PhD student from Malawi), like it was aimed at medical researchers more generally (which it is). Hence although my recruitment materials primed participants to expect the website to be aimed at qualitative researchers, that they correctly identified it to be for medical researchers in general suggests my intervention components were working as intended. This was further evidenced by participants who described the intended audience as “those who are just getting started in their career” (Pre-PhD student from Malawi) or “master research or higher degrees and also for some junior scholars” (PhD student from China)“.

In general, all participants were able to understand and use the interactive website elements. Only a few participants had difficulty with any features. For example, one thought the dotted lines representing a pop-up definitions were “misspelled words” (ECR from India) (because that’s how Microsoft Word highlights errors). A few voiced confusion about the discussion board for each item. Another described feeling frustrated when clicking a footnote caused the page to scroll unexpectedly, and preferred

when notes were placed within each item instead of at the bottom of the page. Nobody was surprised by the drop down expandable boxes or by the search button.

A few components required clarifying the relationships between the website, the EQUATOR Network, the guideline developers, and the original publications, and sometimes this clarification was unsuccessful. For example, a couple of participants were not immediately sure whether the guidance on the website was the same as the guidance within the publication. Some asked whether the citation and journal data (which were supposed to instil trust) were pertaining to the publication or the website, and which one they themselves should cite. When I asked participants where they might look for clarification, all referred to the FAQ and felt reassured after reading how the guidance was developed, but suggested this explanation be summarised and signposted earlier.

A couple of components involved adding quotes to the home page and guideline page. On the whole participants liked or felt neutral about these quotes. For example, one described how the quotes were

“practical from a different point of view. Like why exactly you need this [reporting item]. So now this person [in the quote] is telling from her own perspective how useful it is that you have [the item] described clearly, so it makes it such that if I’m trying to describe [this item], I’ll try to keep that in mind.” (Medical student from Ghana)

Another said

“I like it because each of them tells me why... umm... you know, kind of gives a plain language reason for [...] why it’s a useful thing. [...] That’s gives it, you know, humanity.” (Researcher from Australia)

Regarding quotes from academics who use reporting guidelines, they said

“That makes it relatable to a user, particularly a new user, because we can see that all of these people are, you know, they were first time users once.”
(Researcher from Australia)

However, some participants questioned whether these quotes were from real people.

“Maybe they’re real... maybe it’s legit” (Research consultant from the Philip-pines)

A few others said they “don’t care what people think” (Researcher from South Africa) or “did not pay much attention to [the quotes]” (Pre-PhD student from Malawi).

Six intervention components received no mention. Two of these were purposefully not tested, three others were perhaps too subtle, and one was about *removing* aversive design, so it was good that no participants commented on the presence of ugly or judgemental design. Like all results discussed in this section, these unmentioned components are also in appendix V.

Influences

Participants naturally discussed influences they encountered when applying guidance, either during this study or in their previous experience. These influences were external to the website being tested, and beyond the scope of my intervention components and hence I did not code them as deficiencies. I had identified many of them in my previous work (chapters 3 - 5).

For example, in my thematic synthesis (chapter 3) I described how giving reporting advice at journal submission was a bad time because authors lack the time and motivation to change their writing. When making the website, I acknowledged that most authors would encounter it during journal submission. None of my intervention components seek to alter that initial encounter context directly (doing so would require changing our acquisition channels by, for example, getting more funders to link to our website). Consequently, I expect many authors will arrive at the website in the busy mindset of journal submission and wanting to get things done quickly. One participant articulated this concisely (Researcher from South Africa), when reflecting on their first interview session, when they felt “annoyed by stuff” because they were “working on something at the time” and so instead of “exploring” the website they were “just trying to get to where [they] wanted to be so that [they] could finish the work [they were] doing”. They wanted the reporting guidance as a short checklist and “didn’t want all the additional stuff”, referring to the longer reporting item explanations, the guideline introductory text, and

the persuasive home page content.

However, after using the reporting guideline in their own time their opinion had totally changed by the second session: “when you specifically asked me to look at this and I used it [to write my] discussion, I was embracing it in a different way”. They found the guideline “really helpful” for writing, and then “enjoyed looking at all the different checklists and reporting guidelines”, ultimately deciding that “in a new journal that I’m a deputy editor [of], I’ve just said that in our in our scope and guidance for authors, we have to say that we require the use of reporting guidelines”. This change of heart came after a shift in context: whereas in the first session the participant was looking to get a job done quickly, by the second session, they had given the guideline time and used it in its intended way. The participant attributed this shift in context to being “specifically asked” to use the guideline for writing. Many components seek to achieve such a shift by convincing authors to come back and use the website earlier when writing up their next piece of research (see previous sub sections on describing what reporting guidelines are, when they are best used, and what they are best used for).

Another influence I identified in my thematic synthesis but did not address in this website was how to incorporate reporting guidelines into one’s writing process, or what to do when you don’t have a writing process at all. Although some components communicated that reporting guidelines should be used when drafting or writing manuscripts, none explained *how* (although one component involved directing authors towards training already delivered by EQUATOR). One participant eloquently described the challenge of adopting writing advice into their own practice.

“When I try to look at your guidance on your website, I really want to use it in my own writing, but it is very strange because when I try to, uh, connect the information on the website with my own writing, I found there there might be a great gap because I think everything on your website is very clear (actually they are very specific, those suggestions), but when I try to connect those information with my own writing, I found it just a little bit difficult to generate some specific ideas to start my writing.

So I’m thinking if that’s because the problem of my writing is not the lack

of specific guidance but some other thing like my motivation or, I don't know... it's just a little bit strange.

And I also talk about this with my friends because lots of my friends, they are also PhD students and they are struggling at writing too. So ask them if they have some guidance, uh, if they have looked at some guidance and if [they] have put those guidelines in [their] own writing and their answers were, like, quite similar with me and they all talk about that, '*yes, we look at lots of guidance we try to look at lots of those writing books to teach you how to write, to teach you how to structure your writing. But it's still very hard*'. When you really sit down and start writing, actually you couldn't, uh, call up [the information]." (PhD student from China)

Participants' echoed other influences I had identified in my thematic synthesis when reflecting on their prior writing experiences, including:

- Not having known what reporting guidelines were earlier in their career
- (Previously) finding the checklist, but not the full guidance
- Being limited by journal requirements and word limits
- Struggling to keep writing concise and fluid
- Needing more guidance
- Being unable to report an item because it is their colleague's responsibility, or because they had not done what was being asked when designing their study or collecting data.
- Paywalls
- Not teaching students about reporting guidelines
- Reporting guidelines not existing for funding applications
- Funders not enforcing reporting guidelines

Participants also mentioned influences I had not identified previously. These included:

- When guideline author names appear western, some (non-Western) participants expected the guidance to be less relevant to them.

- The loading speed of websites (thankfully, this was not an issue for the website being tested)
- Not understanding reviewer feedback
- Not wanting to read on a screen
- Not understanding the relationship between the EQUATOR Network and the guidelines or guideline developers.
- Describing a guideline as “version 1.0” might make people feel like the guidance is (too) new, and therefore less trustable. This influence was new to my website as no existing reporting guidelines describe themselves as “1.0”.

Comparisons between the website being tested and the old EQUATOR website & guideline publications.

A few participants ended up exploring the original EQUATOR website and the original SRQR publication during their interviews. Participants instigated these unplanned explorations and comparisons for different reasons. One wanted to retrace their steps to show me the guideline they had used previously. Two others wanted to continue using reporting guidelines in the future and asked me where the original SRQR guidance could be found. Some others spontaneously reflected on their previous experience.

Recounting their experience of seeing the original EQUATOR website for the first time between interview sessions, one participant (ECR from Ecuador) said “Ohh no I didn’t like it. The [new] one is much, much better” because it looked more “trustworthy, more organised” and they preferred the font and colours. Another participant described the original website as “boring”, “outdated” and “text heavy” before recounting their experience of using it:

“Not that long ago I went on to the site because I was looking to complete a reporting checklist and it seemed clear to find the checklist that I wanted. But when I went through the checklist, it wasn’t appropriate. And then I just ended up feeling a bit unsure about what it is, which was the best one to go for.” (Researcher from the UK)

I witnessed another participant (Pre-PhD student from Malawi) experience similar confusion. They wanted to find the original SRQR guidance to continue using it after the study finished. Sharing their screen and thinking aloud, they started on the EQUATOR Network home page and tried to find the original SRQR guidance without my help. Although at first they thought EQUATOR's home page looked "full" and "rich", they were quickly "confused" by both EQUATOR's website and the SRQR publication. After eight minutes and giving up three times, they eventually found the checklist but not the supplement containing the full guidance.

A second participant (Researcher from South Africa) achieved the same outcome a few minutes faster. Because examples only appear in the (not found) supplement they instead looked through "the reference list to see if there was potentially an example paper" and then planned to "go back to PubMed and search for an article that used these guidelines". The participant appeared to have little interest in the article's text, saying they did not "care what [the guideline developers] did to come up with it".

Another participant (ECR from India) echoed this opinion when comparing the original SRQR publication with the redesigned version. They said they "don't need" to know how SRQR was made when they are trying to use it, and they felt the redesigned guidance is "a bit more precise and to the point". When I showed them the original SRQR full guidance (the supplement), they said:

"Participant: That's too heavy on the content.

Interviewer: So if the option was between this this version that you're looking at now [the original supplement] and the website that you saw first, which do you think you prefer to use?

Participant: I think the website is far better than the [supplement]. Yeah, this website is far better."

Another participant reflected on their previous experience of using the PRISMA guidelines and explanation document. They said:

"I rather prefer this form of guidance [the website] than the other one [the publication]. There can be a lot more information presented in this way. [...] That's better because it's more (how can I say?) well presented, well laid out, so that where I need to go deeper, I can go easily. Where I need just surface information or the parts that I'm already familiar with, I can just scroll through [...] So I think I'll prefer something presented in this way than the than the document that I read" (Medical student from Ghana).

10.4 Discussion

The purpose of this study was to identify deficiencies in a website for disseminating reporting guidance. I interviewed 11 researchers and used multiple qualitative methods to identify 53 deficiencies. Most intervention components on the website's home page aim to communicate what reporting guidelines are, that they are best used early in writing, and how they will benefit the author. The results demonstrated most of these components to be somewhat successful, but not yet optimal. For example, some participants needed more than 5 seconds to realise the website was about resources to help them write. Participants often found later, longer content more useful than the short text at the top of the page. In seeking to balance brevity and clarity, perhaps I had been too mean with my word count. If "easier writing" is vague, "faster first drafts" might be concrete. If "complete reporting" isn't intuitive, perhaps "writing up research fully so that everyone can understand, repeat, apply, and synthesise your work" is.

I had sought a similar balance between clarity and brevity when trying to organise the full SRQR guidance (35 pages in its original form) onto a single webpage, in a way that made it appear shorter and less intimidating. Again, the current design was somewhat successful. Participants liked the web features I had used to make the guidance more digestible, like expandable content, navigation menus, subheadings and consistent structure. However, some still felt the guidance looked too long, whilst others wanted to add content that would make it longer still; more examples, more information, more definitions, more signposts to other help. One solution may be to display reporting items on separate pages, as the ARRIVE developers have done on their website [210]. Another

may be to display a summary of the guidance at the very beginning.

Many participants commented on the website's design. Whereas I had been somewhat successful in projecting simplicity, for some participants, this crossed the line to basicness, especially in the first iteration. Many intervention components use design as a way to persuade and communicate with authors: I wanted pictures to depict tools, benefits, and purpose; layout and colours convey a feeling of ease and openness. Sadly neither I nor my colleagues possess expert design skills. Images took a long to create and, unlike text, are hard to iterate. This is a pity, as design often seemed more salient to participants than text, and bad design misled participants and put them off.

Design was also linked to another theme important to this study: credibility. For some participants, the website's basic design eroded its trustworthiness. I mitigated this partially in the second iteration (e.g., by including logos), but future design iterations would ideally include professional design input.

10.4.1 Influences

Credibility rests on more than just design. Participants also wanted assurance that the guidance (text) could be trusted, which necessitated understanding the relationship between EQUATOR, guidelines developers, the original guideline publications, and the content of associated resources. Understanding this relationship was one of six new influences participants mentioned that may affect whether they successfully adhere to reporting guidelines. In chapters 3 and 4 I argued the need for more, in depth qualitative exploration of influences. Although I did not aim to solicit influences in this study, that I found novel influences incidentally suggests I have contributed towards filling that gap. Participants also mentioned eleven influences that I *had* previously identified in my earlier work. Therefore, this study adds credibility to my previous findings whilst also building upon them.

10.4.2 Strengths

Finding novel influences is testament to the strengths of this study. I recruited authors with diverse backgrounds and writing experience. My methods solicited rich information. My thorough analysis used my intervention component table as a framework to draw as much information as possible for the data. In contrast, many studies I reviewed in chapters 3 and 4 recruited homogenous samples, solicited thin description through surveys, and described their analysis techniques poorly. The few studies that elicited rich information focussed on content (e.g., PRISMA 2 [115]) or application (SQUIRE 2 [112]) of a reporting guideline but not the design or the website/publication hosting the guideline. By focussing on diverse recruitment, rich exploration of the guidance text and surrounding platform, and thorough analysis and reporting, I have strengthened my study and addressed limitations seen in others.

10.4.3 Limitations

However, other limitations remain. I will now discuss how 1) this study lacked contextual diversity and 2) not all intervention components were explored.

Context

My web audit (chapter 5) found only half of EQUATOR's current visitors view the home page. Many arrive to the website directly on a reporting guideline page, often as a referral from a journal or a search engine. Because so many visitors never view the home page, many intervention components need to be placed on the home page *and* the reporting guideline page. For instance, naïve visitors should be able to tell what reporting guidelines are whether they arrive on the home page or directly on a guideline page. Some participants noticed this duplication and a few suggested removing or minimising it. However, because all participants viewed the home page first, this study did not capture experiences representative of website visitors that never see the home page. Therefore, future studies should explore the experiences of participants viewing the guideline page without seeing the home page.

Many authors discover reporting guidelines as they are submitting to a journal, whereas authors in this study were not. Because authors described manuscript submission as

an inconvenient moment to intervene (see chapter 3), this may influence how authors experience the website. Once the website is live and journals are directing traffic to it, future studies can explore the experiences of authors using the website in contexts that are more true-to-life, as part of their journal submission journey. Similarly, if funders or ethics boards begin asking applicants to use reporting guidelines this context should be explored too.

Not all intervention components were explored equally:

Some intervention components received little to no discussion. The five second test, think aloud, plus minus test, and writing evaluation all examine *salient* intervention components and will not elicit discussion of un-noticed components.

Sometimes this was useful and expected. For example, one component was to *remove* patronizing language. That nobody spontaneously described the website as patronizing was a success. Similarly, another component was to use terms consistently. This component was only salient when it had not been applied properly, for instance, where I had used the terms “guidelines” and “reporting guidelines” interchangeably. For components like these, a good outcome is to go unmentioned. However, other components still deserve evaluation even though they are purposefully not salient. My semi structured interview questions addressed this limitation to some extent by asking participants directly about less salient features.

Some components could not be fully explored until the website is further developed. For example, although participants recognised the search button, they could not explore the search functionality. Although participants said they liked the links to related guidelines, I could not explore participants’ ability to find and select guidelines because the website only included SRQR. Once other guidelines are uploaded, future studies could use task based protocols [211] to explore how participants find, compare, and select appropriate guidelines.

10.4.4 Future studies

Whereas the limitations above affected my success in reaching my objectives (identifying deficiencies), my objectives were themselves limited and further work is needed to develop the website into a fully functional resource. I will now discuss potential future studies, including 1) prioritising deficiencies; 2) further iterations to address deficiencies; 3) extending the website with other guidelines, checklists, templates, examples, and resources; 4) evaluating components that could not be evaluated in this study; 5) comparing authors' preference between the new and existing website and guidelines; 6) real world evaluations and 7) evaluating reporting guideline content.

Prioritising deficiencies

I made no attempt to prioritise deficiencies. Although some were more commonly raised than others, this was because of saliency and because of the methods I chose. For example, by choosing to use the 5 second test, I encouraged participants to focus on components featured at the top of the landing page. Similarly, my semi structured interview questions drew attention to particular components. Therefore, code frequencies should not dictate deficiencies' priority and I purposefully have not reported them.

Instead, de Jong and Schellens [212] suggest ranking deficiencies according to their likelihood and severity. Likelihood refers to the number of users that may be affected by the deficiency, and severity means the degree to which the deficiency will block the desired behavioural outcome. I made no attempt to estimate these factors systematically in this study. Instead, I judged them instinctively when deciding what I could feasibly change in the first iteration.

More iterations are needed to fix deficiencies

Once prioritised, the remaining deficiencies need addressing and it is my intention, funding permitting, to design and evaluate new iterations after my DPhil. Testing future iterations with an identical protocol would offer continuity. It may be more prudent, however, to adjust the study protocol to target particular components or contexts.

Future evaluations are needed after extending the website

Future evaluations will also be required after the website is extended with more guidelines, search functionality, and with checklists, templates, and links to training and resources. Because different reporting guidelines cater to different research communities, and because these communities may have their own nuances and needs, future evaluations should recruit participants from these communities. For example, CARE may be more commonly used by clinical academics, and ARRIVE authors may come from the life sciences and medical sciences. One reason I chose SRQR was for its diverse user base. As the website grows and its audience expands, recruitment should diversify further.

Once checklists and templates are added, future evaluations should explore participants' experiences of using these resources with and without prior exposure to the website. Just as some authors will bypass the home page and land directly on a guideline page (see Context section within limitations), some authors may receive a checklist or template directly from a colleague or journal without first visiting the website. Therefore, these resources should be evaluated in isolation *and* within the context of the website.

Evaluating components not explored in this study

Some components could not be explored in this study. Optimizing the website for search engines can only be assessed by an audit and by monitoring the website's rankings using a tool like Google's search console [213] once the website is live. Another component involved adding information to items to instruct authors what to do if a particular item was not, or could not be done. This item was more applicable to reporting guidelines for quantitative research, many of which make assumptions about design choices. SRQR is fairly agnostic to design choices, and I only added information to one item (item 5, regarding qualitative approach). In the writing evaluation, I asked participants to describe what part of their manuscript they were working on and I then recommended 2 or 3 reporting relevant reporting items. Item 5 was not relevant to any participants, and so no participants noticed nor commented on the component.

Comparing preferences

This study did not aim to explore whether participants preferred the revised reporting guideline and website over the existing ones. The few participants who made this comparison naturally all expressed preference for the redesign, but future studies could explore preferences in detail. Doing this qualitatively would reveal *reasons* behind preferences. A larger survey could confirm whether authors prefer one version above another.

Real world evaluations

Once the new website and redesigned reporting guidelines are live, real world evaluations should continue to monitor, understand, and improve authors' experiences. This will include using google analytics to monitor how authors use the website, online surveys and other feedback channels, and opportunistic recruitment of authors engaged in their day-to-day work.

Some important metrics include the proportion of authors that return to use the website, the proportion who access resources for drafting vs checking (I would hope to see more authors use the former), and the length of time authors engage with guidance.

Because journals will probably continue to be an important dissemination channel, one possibility would be a mixed methods feasibility study, in collaboration with a journal. Such a study could combine google analytics data with author interviews and writing evaluations of manuscript submissions.

Evaluating guideline content

This study did not attempt to evaluate the SRQR recommendations themselves, but rather the guidelines' presentation. I was interested in what participants thought of the structure, order, and layout of the guidelines, but not of its content. I was trying to look at the guideline on a macro level, and I was not interested in whether participants took issue with particular instructions.

I hope that guideline developers will begin evaluating their content in more detail. They could make use of de Jong and Schellens' advice which, as I mentioned earlier, suggests a range of methods to explore the criteria needed for text to be effective [204].

10.4.5 Conclusions

This study aimed to identify deficiencies in a redesigned version of the SRQR guideline and EQUATOR Network home page. Intervention components were deficient if they could more successfully drive authors towards our target behaviour: to use reporting guidance as early as possible in their research pipeline. In identifying 53 deficiencies, I met my objective, but this success is a double-edged sword. This is the final research chapter of my thesis, and it would have been satisfying to conclude with “I’ve done it! The website is perfect!”, but the results of this study prove otherwise. Unfortunately, the realities of iterative design and limited funding force me to end with unfinished business. Nevertheless, I have suggested further studies to continue and extend the work presented here. In the next chapter, I will discuss my thesis as a whole, directions for future work, and implications for guideline developers and other meta-researchers interested in changing the scholarly system.

10.5 Reflections on this chapter

Writing the methods and results sections of this chapter were high and low points respectively. When writing the methods, I combined what I’d learnt from writing all previous chapters. As in chapter 3 I started by working through the reporting guideline item-by-item but this time I wrote bullet points instead of prose. I then drew on what I had learnt about structure and drafting when writing chapter 9 to reorganise these bullet points into a coherent linear narrative. I then topped and tailed paragraphs with topic and linking sentences, before sandwiching evidence and examples in between. I was happy with the result. The first draft scanned better than other chapters, and I felt confident I’d included guideline content.

The low came when writing my results section. I had intended to report results for each intervention component. If my thesis word count weren’t so limited this is what I would have done. Other departments have higher limits to better accommodate qualitative research with lengthy results. I found it impossible to condense my results further without losing details and nuances I wanted to retain, and so instead I decided to move them to an appendix and to construct a narrative around some highlights. But in selecting

highlights I was giving saliency and importance to a subset of results. Some readers may interpret this as an additional subjective layer of analysis and may criticise me for picking these highlights, but the alternative would have been to reduce my results so much they lost their richness. Stuck between two bad options I chose the former. I have tried to be clear that selecting highlights was not another stage of analysis, and I refer readers to my full results in the appendix. Yet I feel like my word limit has forced me to add another filter to my results that I never wanted to apply.

When conducting these interviews there were moments I was aware of my emotional responses. Sometimes participants would validate my intentions, or confirm my hypotheses. For example, when one spoke about the challenges of incorporating writing guidance into their own practice, this resonated with my experience of writing my previous chapter. Other times participants criticised what I'd made or spoke against my expectations. I believe I dealt with these moments well, in part because of my experience in software development. When collecting feedback on something you've created, it is tempting to dwell on the positive and dismiss the negative. Over the years, I've learnt to pay most attention to the negative, the surprising, the friction felt when someone's experiences are not what you expected. When I first started working in development these moments could feel like a personal attack, uncomfortable or disappointing, but I've learnt that they are often the most valuable. Nowadays I tend to lean into those moments and I felt myself doing that in these interviews. When participants said something unexpected or critical, I would ask more questions to try to fully understand their point of view.

Chapter 11

Discussion

11.1 Chapter Overview

In this chapter I briefly summarise each chapter's findings and my main output; a novel platform for creating and disseminating reporting guideline resources. I discuss how my findings and approach may help other meta-researchers, before considering strengths, limitations, and implications for policy.

11.2 Summary of findings

In this thesis I aimed to identify and address influences affecting whether authors adhere to reporting guidelines. I chose this aim because I believed addressing these influences will increase the proportion of authors adhering to reporting guidelines which will, in turn, make research articles easier to understand, synthesise, replicate, and use, ultimately leading to better patient outcomes.

In chapters 3 - 5 I explored influences through a qualitative evidence synthesis, a review of survey questions, and a service evaluation of the EQUATOR website. I identified 32 influences affecting how authors discover, find, understand, and apply reporting guidelines.

In chapter 7 I described how I used a framework for designing behaviour change interventions called the Behaviour Change Wheel [154] to prioritise intervention options.

Over a series of workshops, reporting guideline experts from the UK EQUATOR Centre and I decided to prioritise education, training, persuasion, modelling (demonstrating using reporting guidelines), and restructuring the environment (both physical and digital environments) as intervention functions. Conversely, we saw restriction and incentivization as inequitable. When considering policy categories, we prioritized communication, guidelines, and service provision.

In chapter 8 I described leading focus groups with stakeholders to generate 128 ideas to address influences, and in chapter 9 I describe how I turned some of these ideas into 46 intervention components and brought them to life by redesigning a reporting guideline and the EQUATOR Network home page.

Finally, in chapter 10 I described how I evaluated the redesigned reporting guideline and home page with a diverse group of authors. The 53 deficiencies I identified can be addressed in future iterations.

11.3 Overview of outputs

11.3.1 A platform for generating and hosting user-friendly reporting guideline resources

The main output of my thesis is the website. Currently this website comprises only a redesigned reporting guideline and EQUATOR Network home page. Although I have only redesigned one reporting guideline, SRQR, my approach will easily scale to others. I have built the website so guideline developers can upload and edit their own content. Developers (or EQUATOR staff, or I) simply need to upload content as plain text files: a file for each reporting item, a file for meta data (like the guideline's scope, authors, publication, version etc.), and a glossary. The website will then automatically generate a fully functional webpage for the guideline with hover definitions, discussion pages, collapsible content etc. Guidelines are citable and versioned, and tracking analytics to monitor authors behaviour is baked in.

Hence although I've described my output as a website, it is more like a platform for generating and disseminating reporting guidelines and resources following an evidence-based blueprint. This blueprint, honed by redesigning SRQR, could be applied to any other reporting guideline to make them easier to understand and use. The website's layout and search engine optimization will change how authors find reporting guidelines, and its content will hopefully nudge them towards using guidelines earlier in their workflow. In my introduction I argued that reporting guidelines are a complex intervention, in part because of the number and variation of resources and how they are used. My website has potential to re-mold quite a lot of this complexity: the guidelines themselves, their associated resources, and how and when authors interact with them.

I believe the website will have a big impact on guideline development groups, as few have sufficient funding nor expertise to design and refine resources or to develop and maintain websites, even when using DIY “no code” tools like Wordpress. One guideline developer I spoke to spent weeks creating a simple website with little functionality. The website I've built is comparatively far more feature-full, and requires zero technical work from guideline developers. What previously took guideline developers weeks is now achievable in a matter of hours. Even guideline groups with no website budgets can use my platform to turn plain text into online resources that mirror the redesigned SRQR guideline.

Two prominent reporting guideline websites (CONSORT and PRISMA) went down during my DPhil. With no software expertise within in the guideline groups, and no budget to hire a developer, the websites stayed offline for many months. Providing a single platform like mine means guideline development groups need not worry about maintaining their own systems. The platform itself uses simple, reliable, globally used gold-standard infrastructure familiar to many DPhil students, so future maintenance will be easy and cheap.

11.3.2 Conferences and publications

I presented chapters at 3 conferences and won 3 awards. At the 2022 World Conference for Research Integrity I presented an overview of chapters 3 - 8, covering my approach, identified influences, and ideas to address them, and I won 1st prize for Excellence in Doctoral Research and 2nd prize for my oral presentation in the Early Career Researcher category. A few months later I presented a poster covering the results of my focus groups (chapter 8) at the Reproducibility, Replicability and Trust in Science 2022 organised by the Wellcome Trust. I won 3rd prize amongst departmental final year DPhil students in 2023 at the Botnar Institute Student Symposium for my presentation showcasing my redesigned reporting guideline (chapter 9), which I then presented again at the 2024 World Conference for Research Integrity.

I intend to publish 5 articles originating from this thesis. These are:

- 1) my qualitative evidence synthesis (chapter 3),
- 2) my review of survey content (chapter 4),
- 3) the workshops and focus groups (chapters 7 and 8),
- 4) intervention refinement (chapter 10 updated after more design iterations), and
- 5) a finalised intervention description (an update of chapter 9 once the design is finalised).

11.4 Contributions and transferability to the meta-research community

Beyond the immediate impact of the website, I believe my work will bring three other benefits to the reporting guideline community and other meta-researchers by 1) providing possible explanations for previous research findings 2) opening new lines of enquiry and funding options and 3) as a model for other grass-root academic movements.

In my introduction chapter I summarised previous research evaluating the impact of reporting guidelines. Few studies employed qualitative methods and process evaluations were scant and shallow. Consequently, although these studies offered a depressing survey of poor reporting standards, they did not explain *why* reporting guidelines had little

effect or *how* they could be improved. Researchers were operating in the dark and could not see how to move forward. My work turns the light on and illuminates hypotheses to explain and address past failures.

I believe EQUATOR staff had a lightbulb moment of their own. My thesis was never meant to be a behaviour change project and my initial supervisory team had no experience in behaviour change nor qualitative methods. They were firmly in the quantitative camp, familiar with statistics, systematic reviews, and randomised trials. They repeatedly warned me against creating software, as they felt it fell firmly within the domain of *development* and not *research*. They typically shoehorned meta-research into clinically-focussed grant applications and maintained their website on a shoestring. I have demonstrated to my colleagues how developing digital interventions requires a great deal of research, just as pharmaceutical, surgical, or physical interventions do, and thus can be packaged as a thesis project and placed centre stage in funding applications. Brainstorming grant ideas in a recent team meeting, I was buoyed to see a colleague pull out a copy of my appendix O — the ideas generated in my workshops and focus groups — covered in highlighter. Framing reporting guidelines as a behavioural intervention means EQUATOR is no longer restricted to medical research funders, and can access new funding sources like the Economic and Social Research Council who have recently announced plans to radically expand UK behavioural research capacity [214]. Framing reporting guidelines as an *online* intervention might release funding to support EQUATOR’s website. Therefore, pursuing my approach will bestow EQUATOR with new research directions, collaborations, and funding options.

These new avenues are open to other meta-researchers too, and I hope my work inspires guideline developers and grass roots movements further afield. Some of my results may be directly applicable to others. For instance, one of my intervention components was to use language to convey confidence and benefits instead of judgement and fear. Sadly, negative language pervades discussions on research integrity and vilifies researchers, accusing them of “waste”, “questionable practices”, “failing”, or “lacking integrity”. This may alienate well-meaning researchers. Shifting the narrative towards “efficiency”, “ease”, or “confidence” might make conversations more welcoming and attractive.

Although my *results* are somewhat transferable, my *approach* is more so. Many of my intervention components are too tethered to reporting guidelines to be of interest to other fields. But the approach that I took and the methods I used could be useful to most meta-researchers seeking to drive change. Although commonly used in medical research, I have not found any meta-research groups using behaviour change frameworks to improve the scholarly system. Although the Reproducibility, Replicability and Trust in Science conference positioned reproducibility and replicability as a behaviour-change problem, I was the only presenter using a behaviour change theory and framework. I think meta-researchers are missing a trick. Without a framework, change-drivers risk getting hung-up on their favourite intervention types (in my experience, most often regulation, training, or education) to the neglect of others. For example, in 2022 I attended a workshop to brainstorm strategies to increase equity and diversity funding applications. One participant gave a rich account of how their research support department re-designed their systems, shared case studies, praised examples of best practice, ran training courses, and held people to account when necessary. The facilitator only noted the word ‘training’. Had he been more familiar with behaviour change, perhaps he would have recognized the participant’s examples of environmental restructuring, persuasion, education, incentivization, and coercion.

11.5 Strengths

Beyond using a behaviour change framework, meta researchers could benefit from integrating the other strengths of my work, including the use of systematic methods, diverse recruitment, and qualitative methods to solicit rich information.

I used systematic methods throughout my thesis. My literature search for chapters 3 and 4 was systematic. The behaviour change wheel and APEASE criteria that I used in chapters 3 - 9 were themselves made systematically, and require users to consider and prioritise options systematically. In all of my data analysis, from identifying influences, ideas, and deficiencies, I sought to code and collate *all* available information: just as systematic search seeks to identify all relevant literature, my coding strategy sought to identify all themes within my data. Similarly, when moving from one stage to another

- from influences to ideas, from ideas to intervention components, from components to deficiencies - I considered and linked items fastidiously, thereby drawing threads through my thesis from start to end.

Another strength is my use of qualitative methods to elicit rich description from participants. I've already extolled the benefit of using a qualitative approach, but within the world of qualitative research one must still be judicious when selecting methods. For example, in our 2019 study with BMJ Open (before my DPhil) we chose poorly. We were seeking to identify deficiencies in a different website, and we thought adding a free text question to an online survey would help. Over 21 months, we contacted 11,000 authors, of whom 93 answered the question "How could we make [the website] more useful?", mostly with very short answers. We only identified 6 themes. In this thesis, by stark comparison, I identified 53 deficiencies in a fraction of the time and by recruiting only 11 participants, by using more appropriate qualitative methods. This tale reveals a warning to guideline development groups: although a qualitative approach may seem accessible, doing it *well* requires expertise. Guideline development groups would be wise to include qualitative experts, preferably those with experience in behaviour change interventions and refining text.

My diverse recruitment was another strength. I wanted diversity because I wanted to understand the perspectives of all stakeholders and because believed it would lead to more discoveries: more influences, more ideas, more deficiencies. I achieved diversity in my stakeholder focus groups (academics, publishers, and guideline developers) and when evaluating the website (authors of varying demographic, disciplines, and experience). Whereas my qualitative synthesis found little geographic diversity amongst participants, my previous chapter included participants from South America, Africa, Asia, Europe, and Australia. I feel proud to have addressed this need, but it was not easy. Twitter proved useless. I relied largely on Penelope.ai - the manuscript checker I created - and I was fortunate to have budget to pay participants. Other development groups may not have such luxuries.

11.6 Limitations

11.6.1 In choosing a framework, I neglected others

Although using a behaviour change framework was a strength, in choosing it I decided *against* using others. In chapter 6 I explained why I chose the Behaviour Change Wheel above two other popular frameworks - the Theoretical Domains Framework and the Person Based Approach. There are plenty of others out there, and behaviouralists may grumble about my decision not to use their preferred framework. I chose not to do a formal comparison as others have already, and shown frameworks vary in their focus, evidence base, and ease of use (e.g., [159,215]). In choosing a framework I did not look for *the best* but rather the one that was *best for me*. It had to be based on evidence - that was a given - but beyond that it had to be the framework-of-least-resistance. I was already leading my (initial three) supervisors down new avenues, so my framework had to be easy to understand and not too far from familiar epistemological ground.

Research groups leaning towards other frameworks may avoid elements of friction I encountered. For example, in applying the Behaviour Change Wheel to redesigning a reporting guideline and creating web pages, some intervention functions and policy categories did not obviously generalise. For instance, I labelled many components as examples of “environmental restructuring”, because I saw the website as a digital environment, and so (for example) adding digital ‘signposts’ to other web content felt the same as installing physical signposts in a hospital. Readers may feel uneasy seeing me compare a digital environment with a physical one, especially if they read the Behaviour Change Wheel’s definition of *environmental restructuring*: *Changing the physical or social context*. They should be reassured, as I was, after reading the example given straight after: *“Providing on-screen prompts for GPs to ask about smoking behaviour”*. The Behaviour Change Wheel developers’ example of environmental restructuring is itself a digital one [150].

Translating the Behaviour Change Wheel to a purely digital environment was one challenge. Another was that in some instances the framework did not go far enough. For instance, although the framework helped me identify that information should be easy to find, or that design should look simple and professional, the Behaviour Change Wheel

does not tell you *how* to do that. A user experience expert might question why I did not use a user experience checklist or information architecture heuristics. They would be justified, and I intend to incorporate these after completing future design iterations. Future work would do well to draw on these domains, although doing so is often harder than one may expect, and is generally outsourced to experts.

11.6.2 Granular components are hard to describe and isolate.

Some intervention designers may be more used to “offline” interventions: conversations, leaflets, physical objects, places, in-person services. Such designers may consider “a leaflet” or “a service” to be a single element implementing a single intervention function. When I had coffee with one such researcher, he suggested I label my redesigned website as a single “enablement” component. I did not follow his advice. Instead, I have tried to describe and justify the redesign as the sum of many smaller components and changes, each linked to an intervention function and to one or more influence. Some may think I stretched the framework too far by applying it with such granularity, or that my interpretation is a distortion of the Behaviour Change Wheel’s intention.

When the behaviour change wheel authors’ applied it to their own digital intervention, Drink Less, they organised their app into modules [216]; in the normative feedback module a widget displays how the user’s drinking compares to others in the UK; the self-monitoring module allows users to track consumption, and so on. Each module employs one or more behaviour change techniques. Organising modules in this way meant the developers could then conduct a factorial screening trial “to identify the individual components, or combinations of components, within the multi-component intervention that affect change and to screen out the ineffective ones” [217]. In contrast, although I consider my breadth of components a strength (and a testament to my diligent approach to identifying influences and ideas), having many small, intermingled components will make it difficult to isolate the efficacy of individual components. Another difference in our approaches is that in developing Drink Less, the creators labelled decisions around design, credibility, navigability, and plain language as “design principles” without considering their impact on behaviour in much detail. Without explicitly stating why, the authors wrote that visual appeal is generally “valued”, credibility

“should be illustrated”, and language should be concise and jargon free. Only two design principles have their impact on behaviour summarised; notifications can “encourage users to perform actions” and gamification can “increase intervention use”. Whereas Drink Less’ developers summarise these design principles in five sentences, I’ve chosen to describe similar principles (and others) as intervention components that pervade throughout all parts of the intervention. I think this was useful, as linking design principles to influences and intervention functions helped guide their implementation and refinement. For example, where DrinkLess’ developers wanted an “appealing” design, I wanted my visual design to communicate simplicity and trust, and feelings of confidence instead of judgement. Because I knew what I wanted the design to do and because I’d linked it to influences I wanted to address, I had a compass to guide my subjective decisions. The same was true for tone of voice, language, and credibility. Because these design principles were included in my table of components, I made sure to explore them in my interviews with authors.

Another drawback of having so many components arose when trying to succinctly describe the intervention. Whereas I can describe the core modules of Drink Less quite easily, my table of intervention components (appendix P is unwieldy. This made writing my chapters on developing and testing the intervention difficult. I had wanted to report my results component-by-component, but this took me way over my word limit and was difficult for my supervisors to digest. Instead I constructed narrative summaries but these inevitably lost detail and nuance.

11.6.3 My logic model is rudimentary

My attachment to my long list of granular components also affects how I think about and communicate my logic model. In their topology of logic models, Mills et al [218] propose a way of moving from a rudimentary list of components like mine, which they define as a *type 1* logic model, towards a model that concisely captures complexity and context. As my intervention matures beyond planning and refinement, I could explore depicting it as relationships between resources, activities, outputs, outcomes, impact, and domains (what Mills et al refer to as *type 2* and *type 3* models).

11.6.4 I neglected the behaviour of other stakeholders

Such a model should also include the behaviour of editors, peer reviewers, and other stakeholders. This was another limitation of my approach: I focussed exclusively on authors' behaviour. Future research could explore influences and solutions faced by others and incorporate them into intervention design and logic models.

In addition to considering the behaviour of all stakeholders, future research should also consider differences between reporting guidelines. In my introduction I described reporting guidelines as variations on a theme. Because they shared so many commonalities, I justified treating them all the same. However, in practice, some variation may matter. For example, guidelines that cater to writing protocols may be best delivered in a different format, by different stakeholders (e.g., funders, registries) and aimed at researchers at early stages of work. Therefore, subsets of reporting guidelines may deserve their own dissemination strategies and logic models.

11.6.5 My context did not reflect the real world

Logic models should also reflect real-world context as far as possible. Although I fought to mimic parts of real life in my interviews with authors (chapter 10) by allowing naïve authors to explore the website similar to how they would in real life, the context was still far from real. Future research should explore how authors use the redesigned reporting guidelines in a real-world context, and how it impacts their writing.

In my introduction I described a real-world study we carried out in collaboration with BMJ Open before my DPhil. I can imagine performing a similar study with the redesigned guidelines. If I were to repeat the BMJ Open study today, my logic model would suggest alternative outcome measures. My evidence synthesis (3) revealed that for many authors, receiving a reporting checklist at the time of journal submission was a bad time to give advice, as authors lacked the time and motivation to act on the guidance. In the workshops (chapter 7), EQUATOR and I began to think differently about the role of journal endorsement. We decided we want authors to use reporting guidelines *as early as possible* in their research journey and so instead of seeing article submission as the moment where authors should be applying reporting guidelines, we realised that

journal endorsement is merely a good way to make authors *aware* of reporting guidelines, and that we should not expect authors to fully apply them there and then, but rather we would hope authors come back to the website to use a reporting guideline *earlier* in their *next* project. Future evaluations should reflect this logic model shift. In our original BMJ Open study we used reporting adherence as our primary outcome, and we tracked manuscripts through journal submission to look for evidence that authors improved their manuscripts after completing a checklist. We found no such evidence. My new logic model would not expect such immediate changes. Instead, I would hope to see the same authors *returning* to the website in the future (after a few weeks or months), and I believe that authors using the redesigned guidelines will be more likely to return than authors using the old version. To test this refined logic model, I would choose *return rate* as a primary outcome measure, and then compare reporting adherence within those returning authors, comparing adherence in their second manuscripts to their first. Tracking authors over time will be possible using web analytics and by following authors that repeatedly cite my new platform.

11.6.6 Lack of quantification

I made little use of quantitative data in this thesis. I sought to understand, identify, and ideate, not to count, measure, or compare. Consequently, whilst my thesis has generated many hypotheses, it tests none of them.

As intervention development moves beyond planning, designing, refinement, and towards evaluation, new questions will require a quantitative approach. Do authors prefer the redesigned guidelines or the original ones? How many authors come back to use guidelines again? Of the remaining influences that are difficult or expensive to address, which occur most frequently? And of course the ultimate question: Which version of the guidelines — the original or the redesigned — leads to better adherence? All of these questions will require a quantitative approach. They all too have an implicit follow-up question - *why?* - and so any quantitative approach should have a qualitative accompaniment.

In summary, my thesis had a number of limitations. In choosing a framework I neglected to consider others which may have shaped my work differently, especially my

approach to digital design. My detailed approach to identifying influences and solutions led to a large number of intervention components, some of which are small or subtle. This may complicate communicating my logic model or identifying the effectiveness of individual components. In focussing on authors' behaviour I have neglected to consider editors, peer reviewers, or other stakeholders. By considering reporting guidelines as a homogenous group I have not accounted for the differences between them or their users. I prioritized qualitative questions above quantitative, and so whilst my thesis raises many hypotheses it tests none of them. The context in which authors gave me feedback did not reflect real life. Future studies addressing these questions should refine my rudimentary logic model, and make it specific to the context being evaluated.

11.7 Recommendations for future research

In addition to the future work to extend the website and to address the limitations permeating my thesis, my work opens up new lines of exploration that could be developed into research strands. My focus groups and author interviews identified a need for training on how to use reporting guidelines and on how to write in general. Focus group participants felt that a network of “reporting champions”, similar to UKRN’s network model, could be a useful way to both advertise reporting guidelines and offer assistance in using them. Focus group participants also felt that getting funders, registries, and institutions to endorse or enforce reporting guidelines would help get authors using them earlier in their research. Future research projects could inform the development of these opportunities, explore how best to deliver them, and evaluate their feasibility and efficacy.

My work also offers new directions for guideline developers. Having identified in my qualitative synthesis that few reporting guidelines have undergone meaningful user testing despite exhibiting many barriers, developers can use my work to justify funding applications to support such work. I hope that future developers will consider my findings and designs when creating their own resources. EQUATOR could facilitate this by updating their existing guidance for guideline developers.

11.8 Implications for policy

My work touches on policy in two ways: reporting guideline policies held by journals and other stakeholders, and funders' policies towards funding meta-research and grass-roots services.

In my introduction I mentioned that guideline developers have long been calling for journals to better enforce reporting guidelines. I argued that pointing the finger solely at journals was unfair and unrealistic. Nevertheless, journal endorsement and enforcement *are* an important piece of the puzzle. In making reporting guidelines easier to use and understand, my work will make such policies easier to enact. The smaller the hurdle, the more likely journal editors will be to lay it in front of their authors, and if editors can better understand reporting guidelines it will be easier for them to check adherence. Similarly, reducing friction will make it easier for funders to begin requiring reporting guidelines, especially if new guidelines are developed for protocols.

Funders looking to support changing the scholarly system should allocate money for meta-research, behaviour change, intervention development and maintenance. They should accommodate software development costs and value the importance of thorough user testing. Once a resource becomes established, they should fund its evaluation, monitoring, and periodic updating, and if necessary they should provision for long term maintenance far beyond the terminus of a traditional grant. Maintenance costs may be a fraction of the initial research costs, but they need to be reliable and persist for years if not decades. For grass-roots movements to successfully change the scholarly system, academics need their digital tools to be adopted by private sector stakeholders, most of whom will value stability and sustainability which can only come from reliable long-term financial support.

11.9 Conclusions

I have demonstrated how I identified and addressed influences affecting whether authors adhere to reporting guidelines. I identified influences through a qualitative evidence synthesis, survey review, and website service evaluation. I identified solutions

through workshops and focus groups with stakeholders, applied a subset of these ideas by redesigning reporting guidelines, and refined the redesign by interviewing authors.

My redesigned reporting guideline and EQUATOR Network home page could be extended to other reporting guidelines. The EQUATOR Network and guideline developers can use my work to inform and justify future funding applications and develop impactful resources.

My work could be extended by adding functionality and reporting guidelines to the web platform I have built. Future research should address the limitations of my work by exploring the utility of alternative frameworks, developing logic models that include the behaviour of other stakeholders, testing these logic models in real world contexts, exploring authors' preferences and evaluating their reporting quality.

I hope other researchers will draw inspiration from my pragmatic approach that made heavy use of qualitative methods and an established behaviour change framework. However, for academic-lead movements to develop digital tools that successfully change the scholarly system, funders will need to reconsider how they fund such endeavours.

References

1. Liberati A. [An unfinished trip through uncertainties](#). BMJ : British Medical Journal. 2004 Feb;328(7438):531.
2. Ziemann S, Paetzolt I, GrüSSer L, Coburn M, Rossaint R, Kowark A. [Poor reporting quality of observational clinical studies comparing treatments of COVID-19 – a retrospective cross-sectional study](#). BMC Medical Research Methodology. 2022 Jan;22(1):23.
3. Glick BS. [Inadequacies in the reporting of clinical drug research](#). The Psychiatric Quarterly. 1963 Apr;37:234–44.
4. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, et al. [Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: Research on research study](#). BMJ. 2017 Jun;357:j2490.
5. Savovi J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. [Association Between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-Epidemiologic Study](#). American Journal of Epidemiology. 2018 May;187(5):1113–22.
6. Glasziou P, Meats E, Heneghan C, Shepperd S. [What is missing from descriptions of treatment in trials and reviews?](#) BMJ. 2008 Jun;336(7659):1472–4.

7. Feinstein AR. **Clinical biostatistics. XXV. A survey of the statistical procedures in general medical journals.** Clinical Pharmacology and Therapeutics. 1974 Jan;15(1):97–107.
8. Davidson SRE, Kamper SJ, Haskins R, Robson E, Gleadhill C, da Silva PV, et al. **Exercise interventions for low back pain are poorly reported: A systematic review.** Journal of Clinical Epidemiology. 2021 Nov;139:279–86.
9. Carp J. **The secret lives of experiments: Methods reporting in the fMRI literature.** NeuroImage. 2012 Oct;63(1):289–300.
10. Kapp P, Esmail L, Ghosn L, Ravaud P, Boutron I. **Transparency and reporting characteristics of COVID-19 randomized controlled trials.** BMC Medicine. 2022 Sep;20(1):363.
11. Pourrazavi S, Fathifar Z, Sharma M, Allahverdipour H. **COVID-19 vaccine hesitancy: A Systematic review of cognitive determinants.** Health Promotion Perspectives. 2023 Apr;13(1):21–35.
12. Yilmaz S, Çolak FÜ, Hökenek NM, Ak R. **Hesitancy Regarding Medical Advice on COVID-19: An Emergency Department Perspective.** Disaster Medicine and Public Health Preparedness. :1–11.
13. Taylor S, Asmundson GJG. **Negative attitudes about facemasks during the COVID-19 pandemic: The dual importance of perceived ineffectiveness and psychological reactance.** PLoS ONE. 2021 Feb;16(2):e0246317.
14. Covid-19: Response and Excess Deaths - Hansard - UK Parliament [Internet]. 2024 [cited 2024 Jun 10]. Available from: <https://hansard.parliament.uk/commons/2024-04-18/debates/9F01F787-D758-43D4-B8D1-4FA357EB3EED/Covid-19ResponseAndExcessDeaths>

15. Santo TD, Rice DB, Amiri LSN, Tasleem A, Li K, Boruff JT, et al. [Methods and results of studies on reporting guideline adherence are poorly reported: A meta-research study](#). Journal of Clinical Epidemiology. 2023 Jul;159:225–34.
16. Andrew E, Anis A, Chalmers T, Cho M, Clarke M, Felson D, et al. [A Proposal for Structured Reporting of Randomized Controlled Trials](#). JAMA. 1994 Dec;272(24):1926–31.
17. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. [Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature](#). Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Annals of Internal Medicine. 1994 Dec;121(11):894–5.
18. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. [Improving the Quality of Reporting of Randomized Controlled Trials: The CONSORT Statement](#). JAMA. 1996 Aug;276(8):637–9.
19. Altman DG. [Better reporting of randomised controlled trials: The CONSORT statement](#). BMJ (Clinical research ed). 1996 Sep;313(7057):570–1.
20. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. [The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration](#). Annals of Internal Medicine. 2001 Apr;134(8):663–94.
21. Hopewell S, Boutron I, Chan AW, Collins GS, de Beyer JA, Hróbjartsson A, et al. [An update to SPIRIT and CONSORT reporting guidelines to enhance transparency in randomized trials](#). Nature Medicine. 2022 Sep;28(9):1740–3.

22. Schlussel MM, Sharp MK, de Beyer JA, Kirtley S, Logullo P, Dhiman P, et al. [Reporting guidelines used varying methodology to develop recommendations](#). Journal of Clinical Epidemiology. 2023 Jul;159:246–56.
23. Schulz KF, Altman DG, Moher D, the CONSORT Group. [CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials](#). BMC Medicine. 2010 Mar;8(1):18.
24. Hopewell S, Boutron I, Altman DG, Barbour G, Moher D, Montori V, et al. [Impact of a web-based tool \(WebCONSORT\) to improve the reporting of randomised trials: Results of a randomised controlled trial](#). BMC Medicine. 2016 Nov;14(1):199.
25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. [The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration](#). BMJ. 2009 Jul;339:b2700.
26. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. [The PRISMA 2020 statement: An updated guideline for reporting systematic reviews](#). BMJ (Clinical research ed). 2021 Mar;372:n71.
27. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. [Preferred reporting items for systematic review and meta-analysis protocols \(PRISMA-P\) 2015 statement](#). Systematic Reviews. 2015 Jan;4(1):1.
28. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. [Improving Biostatistics Research Reporting: The ARRIVE Guidelines for Reporting Animal Research](#). PLOS Biology. 2010 Jun;8(6):e1000412.

29. Sert NP du, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, et al. [The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research](#). PLOS Biology. 2020 Jul;18(7):e3000410.
30. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. [Standards for Reporting Qualitative Research: A Synthesis of Recommendations](#). Academic Medicine. 2014 Sep;89(9):1245–51.
31. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandebroucke JP. [Strengthening the reporting of observational studies in epidemiology \(STROBE\) statement: Guidelines for reporting observational studies](#). BMJ : British Medical Journal. 2007 Oct;335(7624):806–8.
32. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krlea-Jeri K, et al. [SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials](#). Annals of Internal Medicine. 2013 Feb;158(3):200–7.
33. Gagnier JJ, Kienle G, Altman DG, Moher D, Sox H, Riley D, et al. [The CARE guidelines: Consensus-based clinical case reporting guideline development](#). Case Reports. 2013 Oct;2013:bcr2013201554.
34. Learn how to write and publish case reports [Internet]. Scientific Writing in Health and Medicine. [cited 2024 Mar 22]. Available from: <https://www.swihm.com/course>
35. Case report writing simplified [Internet]. CARE-writer. [cited 2024 Mar 22]. Available from: <https://care-writer.com>
36. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. [The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration](#). Annals of Internal Medicine. 2003 Jan;138(1):W1–12.

37. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015 Oct;351:h5527.
38. Brouwers MC, Kerkvliet K, Spithoff K, Consortium ANS. The AGREE Reporting Checklist: A tool to improve reporting of clinical practice guidelines. *BMJ*. 2016 Mar;352:i1152.
39. Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney SE. Publication guidelines for quality improvement studies in health care: Evolution of the SQUIRE project. *BMJ*. 2009 Jan;338:a3152.
40. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): Revised publication guidelines from a detailed consensus process. *BMJ Quality & Safety*. 2016 Dec;25(12):986–92.
41. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMC Medicine*. 2013 Mar;11(1):80.
42. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: Updated reporting guidance for health economic evaluations. *BMC Medicine*. 2022 Jan;20(1):23.
43. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *BMC Medicine*. 2015 Jan;13(1):1.

44. Chen Y, Yang K, Marui A, Qaseem A, Meerpohl JJ, Flottorp S, et al. [A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement](#). Annals of Internal Medicine. 2017 Jan;166(2):128–32.
45. Tong A, Sainsbury P, Craig J. [Consolidated criteria for reporting qualitative research \(COREQ\): A 32-item checklist for interviews and focus groups](#). International Journal for Quality in Health Care. 2007 Dec;19(6):349–57.
46. Caulley L, Cheng W, Catalá-López F, Whelan J, Khoury M, Ferraro J, et al. [Citation impact was highly variable for reporting guidelines of health research: A citation analysis](#). Journal of Clinical Epidemiology. 2020 Nov;127:96–104.
47. ICMJE | Recommendations | Preparing a Manuscript for Submission to a Medical Journal [Internet]. [cited 2021 Feb 8]. Available from: <http://www.icmje.org/recommendations/browse/manuscript-preparation/preparing-for-submission.html#two>
48. Koch M, Riss P, Umek W, Hanzal E. [The explicit mentioning of reporting guidelines in urogynecology journals in 2013: A bibliometric study](#). Neurourology and Urodynamics. 2016 Mar;35(3):412–6.
49. Sharp MK, Tokali R, Gómez G, Wager E, Altman DG, Hren D. A cross-sectional bibliometric study showed suboptimal journal endorsement rates of STROBE and its extensions. Journal of clinical epidemiology. 2019;107:42–50.
50. medRxiv.org - the preprint server for Health Sciences [Internet]. [cited 2023 Oct 17]. Available from: <https://www.medrxiv.org/>

51. Barnes C, Boutron I, Giraudeau B, Porcher R, Altman DG, Ravaud P. **Impact of an online writing aid tool for writing a randomized trial report: The COBWEB (Consort-based WEB tool) randomized controlled trial.** BMC Medicine. 2015 Sep;13(1):221.
52. Hawwash D, Sharp MK, Argaw A, Kolsteren P, Lachat C. **Usefulness of applying research reporting guidelines as Writing Aid software: A crossover randomised controlled trial.** BMJ Open. 2019 Nov;9(11).
53. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. **PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis.** Campbell Systematic Reviews. 2022;18(2):e1230.
54. Struthers C, Harwood J, de Beyer JA, Dhiman P, Logullo P, Schlüssel M. **GoodReports: Developing a website to help health researchers find and use reporting guidelines.** BMC medical research methodology. 2021 Oct;21(1):217.
55. Compliance Questionnaire | ARRIVE Guidelines [Internet]. [cited 2023 Oct 24]. Available from: <https://arriveguidelines.org/resources/compliance-questionnaire>
56. Pouwels KB, Widyakusuma NN, Groenwold RHH, Hak E. **Quality of reporting of confounding remained suboptimal after the STROBE guideline.** Journal of Clinical Epidemiology. 2016 Jan;69:217–24.
57. Sharp MK, Glonti K, Hren D. **Online survey about the STROBE statement highlighted diverging views about its content, purpose, and value.** Journal of clinical epidemiology. 2020;123:100–6.

58. Altman DG, Simera I. **A history of the evolution of guidelines for reporting medical research: The long road to the EQUATOR Network.** Journal of the Royal Society of Medicine. 2016 Feb;109(2):67–77.
59. Kilicoglu H, Jiang L, Hoang L, Mayo-Wilson E, Vinkers CH, Otte WM. **Methodology reporting improved over time in 176,469 randomized controlled trials.** Journal of Clinical Epidemiology. 2023 Aug;
60. de Jong Y, van der Willik EM, Milders J, Voorend CGN, Morton RL, Dekker FW, et al. **A meta-review demonstrates improved reporting quality of qualitative reviews following the publication of COREQ- and ENTREQ-checklists, regardless of modest uptake.** BMC Medical Research Methodology. 2021 Sep;21(1):184.
61. Nedovic D, Panic N, Pastorino R, Ricciardi W, Boccia S. **Evaluation of the Endorsement of the STrengthening the REporting of Genetic Association Studies (STREGA) Statement on the Reporting Quality of Published Genetic Association Studies.** Journal of Epidemiology. 2016 Aug;26(8):399–404.
62. Howell V, Schwartz AE, O’Leary JD, Donnell CM. **The effect of the SQUIRE (Standards of QUality Improvement Reporting Excellence) guidelines on reporting standards in the quality improvement literature: A before-and-after study.** BMJ Quality & Safety. 2015 Jun;24(6):400–6.
63. Bastuji-Garin S, Sbidian E, Gaudy-Marqueste C, Ferrat E, Roujeau JC, Richard MA, et al. **Impact of STROBE statement publication on quality of observational study reporting: Interrupted time series versus before-after analysis.** PloS One. 2013;8(8):e64733.
64. Hopewell S, Ravaud P, Baron G, Boutron I. **Effect of editors’ implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: Interrupted time series analysis.** BMJ. 2012 Jun;344:e4178.

65. Botos J. [Reported use of reporting guidelines among JNCI: Journal of the National Cancer Institute authors, editorial outcomes, and reviewer ratings related to adherence to guidelines and clarity of presentation.](#) Research Integrity and Peer Review. 2018 Sep;3(1):7.
66. Hair K, Macleod MR, Sena ES, Sena ES, Hair K, Macleod MR, et al. [A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines \(IICARus\).](#) Research Integrity and Peer Review. 2019 Jun;4(1):12.
67. Cobo E, Selva-O'Callaghan A, Ribera JM, Cardellach F, Dominguez R, Vilardell M. [Statistical Reviewers Improve Reporting in Biomedical Articles: A Randomized Trial.](#) PLoS ONE. 2007 Mar;2(3):e332.
68. Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. [Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: Masked randomised trial.](#) BMJ. 2011 Nov;343:d6783.
69. Pandis N, Shamseer L, Kokich VG, Fleming PS, Moher D. [Active implementation strategy of CONSORT adherence by a dental specialty journal improved randomized clinical trial reporting.](#) Journal of Clinical Epidemiology. 2014 Sep;67(9):1044–8.
70. Jin Y, Sanger N, Shams I, Luo C, Shahid H, Li G, et al. [Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? – A systematic review of reviews: An update.](#) Journal of Multidisciplinary Healthcare. 2018 Sep;11:495–510.

71. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. [Reducing waste from incomplete or unusable reports of biomedical research](#). Lancet (London, England). 2014 Jan;383(9913):267–76.
72. Percie du Sert N, Ahluwalia A, Alam S, Avey MT, Baker M, Browne WJ, et al. [Reporting animal research: Explanation and elaboration for the ARRIVE guidelines 2.0](#). PLoS biology. 2020;18(7):e3000411.
73. Malterud K. [Qualitative research: Standards, challenges, and guidelines](#). Lancet (London, England). 2001 Aug;358(9280):483–8.
74. Clough P. Narratives and Fictions in Educational Research. Open University Press; 2002.
75. Woolgar S, editor. Knowledge and reflexivity: New frontiers in the sociology of knowledge. Thousand Oaks, CA, US: Sage Publications, Inc; 1988.
76. Wilkinson S. [The role of reflexivity in feminist psychology](#). Women's Studies International Forum. 1988 Jan;11(5):493–502.
77. Anvari F, Lakens D. [The replicability crisis and public trust in psychological science](#). Comprehensive Results in Social Psychology. 2018 Sep;3(3):266–86.
78. Errington TM, Denis A, Perfito N, Iorns E, Nosek BA. [Challenges for assessing replicability in preclinical cancer biology](#). Rodgers P, Franco E, editors. eLife. 2021 Dec;10:e67995.
79. Goldacre B. Bad pharma: How drug companies mislead doctors and harm patients. Fourth Estate; 2012.

80. Ioannidis JPA. [Why Most Published Research Findings Are False](#). PLOS Medicine. 2005 Aug;2(8):e124.
81. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JPA, et al. [Biomedical research: Increasing value, reducing waste](#). The Lancet. 2014 Jan;383(9912):101–4.
82. Penelope.ai automated manuscript checker [Internet]. Penelope.ai. [cited 2020 Feb 14]. Available from: <https://www.penelope.ai>
83. GoodReports.org [Internet]. [cited 2020 Feb 14]. Available from: <https://www.goodreports.org>
84. Albury C, Pope C, Shaw S, Greenhalgh T, Ziebland S, Martin S, et al. [Gender in the consolidated criteria for reporting qualitative research \(COREQ\) checklist](#). International Journal for Quality in Health Care. 2021 Oct;33(4):mzab123.
85. Haynes A, Loblay V. [Rethinking Barriers and Enablers in Qualitative Health Research: Limitations, Alternatives, and Enhancements](#). Qualitative Health Research. 2024 Mar;10497323241230890.
86. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. [PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews](#). Systematic Reviews. 2021 Jan;10(1):39.
87. Tong A, Flemming K, McInnes E, Oliver S, Craig J. [Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ](#). BMC Medical Research Methodology. 2012 Nov;12(1):181.
88. Chinese Biomedical Literature Database [Internet]. [cited 2021 Mar 1]. Available from: <https://www.imicams.ac.cn/>

89. SinoMed [Internet]. [cited 2021 Mar 5]. Available from: <http://www.sinomed.ac.cn/>
90. China National Knowledge Infrastructure [Internet]. [cited 2021 Mar 1]. Available from: <https://www.cnki.net/>
91. Wanfang Data - A Leading Provider of Electronic Resources for China Studies [Internet]. [cited 2021 Mar 1]. Available from: <http://www.wanfangdata.com/>
92. VIP Chinese Medical Journal Database [Internet]. [cited 2021 Mar 1]. Available from: <http://www.cqvip.com/>
93. LILACS [Internet]. Latin America and Caribbean Health Sciences Literature. [cited 2021 Feb 19]. Available from: <https://lilacs.bvsalud.org/en/>
94. Alves B/O/OM. Global Index Medicus [Internet]. [cited 2021 Feb 19]. Available from: <https://www.globalindexmedicus.net/>
95. Health and Medical Articles Database - African Index Medicus [Internet]. [cited 2021 Mar 1]. Available from: <https://indexmedicus.afro.who.int/>
96. Western Pacific Region Index Medicus [Internet]. [cited 2021 Mar 1]. Available from: <http://www.wprim.org/>
97. IMSEAR at SEARO [Internet]. [cited 2021 Mar 1]. Available from: <https://imsear.searo.who.int/>
98. WHO EMRO | IMEMR | Library [Internet]. [cited 2021 Mar 1]. Available from: <http://www.emro.who.int/e-library/imemr/index.html>

99. SciELO [Internet]. [cited 2021 Feb 19]. Available from: <https://scielo.org/en/about-scielo>
100. Projet MiRoR [Internet]. Projet MiRoR. [cited 2021 Feb 6]. Available from: <http://miror-ejd.eu/>
101. Rosumeck S, Wagner M, Wallraf S, Euler U. **A validation study revealed differences in design and performance of search filters for qualitative research in PsycINFO and CINAHL.** J Clin Epidemiol. 2020 Dec;128:101–8.
102. Rogers M, Bethel A, Abbott R. **Locating qualitative studies in dementia on MEDLINE, EMBASE, CINAHL, and PsycINFO: A comparison of search strategies.** Research Synthesis Methods. 2018;9(4):579–86.
103. CASP Qualitative Checklist [Internet]. Critical Appraisal Skills Programme. 2018 [cited 2022 May 5]. Available from: <https://casp-uk.net/casp-tools-checklists/>
104. Thomas J, Harden A. **Methods for the thematic synthesis of qualitative research in systematic reviews.** BMC Medical Research Methodology. 2008 Jul;8(1):45.
105. Bierner M. Markdown Preview Mermaid Support [Internet]. 2022 [cited 2022 Apr 13]. Available from: <https://github.com/mjbvz/vscode-markdown-mermaid>
106. Dewey M, Levine D, Bossuyt PM, Kressel HY. **Impact and perceived value of journal reporting guidelines among Radiology authors and reviewers.** European Radiology. 2019 Aug;29(8):3986–95.
107. Macleod M, Collings AM, Graf C, Kiermer V, Mellor D, Swaminathan S, et al. **The MDAR (Materials Design Analysis Reporting) Framework for transparent reporting in the life sciences.** Proceedings of the National Academy of Sciences. 2021 Apr;118(17):e2103238118.

108. Davies L, Batalden P, Davidoff F, Stevens D, Ogrinc G. [The SQUIRE Guidelines: An evaluation from the field, 5 years post release.](#) BMJ quality & safety. 2015 Dec;24(12):769–75.
109. Fuller T, Pearson M, Peters J, Anderson R. [What affects authors' and editors' use of reporting guidelines? Findings from an online survey and qualitative interviews.](#) PLoS ONE. 2015 Jan;10(4):e0121585.
110. Korevaar DA, Cohen JF, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. [Updating standards for reporting diagnostic accuracy: The development of STARD 2015.](#) Research integrity and peer review. 2016;1(101676020):7.
111. Burford BJ, Welch V, Waters E, Tugwell P, Moher D, O'Neill J, et al. [Testing the PRISMA-Equity 2012 reporting guideline: The perspectives of systematic review authors.](#) PloS one. 2013;8(10):e75122.
112. Davies L, Donnelly KZ, Goodman DJ, Ogrinc G. [Findings from a novel approach to publication guideline revision: User road testing of a draft version of SQUIRE 2.0.](#) BMJ quality & safety. 2016 Apr;25(4):265–72.
113. de Vries RBM, Hooijmans CR, Langendam MW, van Luijk J, Leenaars M, Ritskes-Hoitinga M, et al. [A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies.](#) Evidence-based Preclinical Medicine. 2015;2(1):e00007.
114. Eysenbach G. [CONSORT-EHEALTH: Implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials.](#) Studies in health technology and informatics. 2013;192:657–61.

115. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. [Updating guidance for reporting systematic reviews: Development of the PRISMA 2020 statement](#). Journal of Clinical Epidemiology. 2021 Jun;134:103–12.
116. Prady SL, MacPherson H. [Assessing the utility of the standards for reporting trials of acupuncture \(STRICTA\): A survey of authors](#). The Journal of Alternative and Complementary Medicine. 2007 Nov;13(9):939–43.
117. Prager R, Gagnon L, Bowdridge J, Unni RR, McGrath TA, Cobey K, et al. [Barriers to reporting guideline adherence in point-of-care ultrasound research: A cross-sectional survey of authors and journal editors](#). BMJ Evidence-Based Medicine. 2021 Jan;bmjebm-2020-111604.
118. Rader T, Mann M., Stansfield C., Cooper C., Sampson M. [Methods for documenting systematic review searches: A discussion of common issues](#). Research synthesis methods. 2014;5(2):98–115.
119. Svensøy JN, Nilsson H, Rimstad R. [A qualitative study on researchers' experiences after publishing scientific reports on major incidents, mass-casualty incidents, and disasters](#). Prehospital and Disaster Medicine. 2021 Oct;36(5):536–42.
120. Tam WWS, Tang A, Woo B, Goh SYS. [Perception of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses \(PRISMA\) statement of authors publishing reviews in nursing journals: A cross-sectional online survey](#). BMJ Open. 2019 Apr;9(4):e026271.
121. Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG, et al. [Describing reporting guidelines for health research: A systematic review](#). Journal of Clinical Epidemiology. 2011 Jul;64(7):718–42.

122. Moher D, Schulz KF, Simera I, Altman DG. [Guidance for developers of health research reporting guidelines](#). PLOS Medicine. 2010 Feb;7(2):e1000217.
123. Topor M, Pickering JS, Mendes AB, Bishop DVM, Büttner F, Elsherif MM, et al. [An integrative framework for planning and conducting Non-Intervention, Reproducible, and Open Systematic Reviews \(NIRO-SR\)](#). Meta-Psychology. 2023 Jul;7.
124. Bantry White E, Hurley M, Ó Súilleabháin F. [The Journal Article Reporting Standards for Qualitative Primary, Qualitative Meta-Analytic and Mixed Methods Research: Applying the Standards to Social Work Research](#). Journal of Evidence-Based Social Work. 2019 Sep;16(5):469–77.
125. Fang Z.-P., Leng X., Liu Y.-L., Liu W.-B., Hu W.-J., Zhang Z.-J., et al. [A survey on awareness of the ARRIVE guideline and GSPC in researchers field in animal experiments field in Lanzhou city](#). Chinese Journal of Evidence-Based Medicine. 2015;15(7):797–801.
126. Ma B, Xu J, Wu W, Liu H, Kou C, Liu N, et al. [Survey of basic medical researchers on the awareness of animal experimental designs and reporting standards in China](#). PLoS ONE. 2017 Apr;12(4):e0174530.
127. Giray E, Coskun OK, Karacaatli M, Gunduz OH, Yagci . [Assessment of the knowledge and awareness of a sample of young researcher physicians on reporting guidelines and the EQUATOR network: A single center cross-sectional study](#). Marmara Medical Journal. 2020 Jan;33(1):1–6.

128. Guo S, Qi S, Yang L, Wang X, Zhu Q, Meng X, et al. Recognition status of quality assessment and standards for reporting randomized controlled trials of traditional Chinese medicine researchers. *China Journal of Traditional Chinese Medicine and Pharmacy* [Internet]. 2018 [cited 2022 Jun 20];(3):1077–81. Available from: https://caod.oriprobe.com/articles/53258508/Recognition_status_of_quality_assessment_and_stand.htm
129. McDonough J., O'Dunne A., B. C, Margerum B., Sutton D. Familiarity of non-industry authors with good publication practice and clinical data reporting guidelines. *Current Medical Research and Opinion*. 2011;27:S9.
130. Karada Öncel E, Baaranolu ST, Aykaç K, Kömürlüolu A, Akman AÖ, Kran S. [Knowledge and awareness of optimal use of reporting guidelines in paediatricians: A cross-sectional study](#). *Turk Pediatri Arsivi*. 2018 Sep;53(3):163–8.
131. Phillips A., Lewis L.K., McEvoy M.P., Galipeau J., Glasziou P., Moher D., et al. [Pilot testing of the guideline for reporting of evidence-based practice educational interventions and teaching \(greet\)](#). *Physiotherapy (United Kingdom)*. 2015;101:eS1203–4.
132. Sharp MK, Bertizzolo L, Rius R, Wager E, Gómez G, Hren D. [Using the STROBE statement: Survey findings emphasized the role of journals in enforcing reporting guidelines](#). *Journal of Clinical Epidemiology*. 2019 Dec;116:26–35.
133. The EQUATOR Network | Enhancing the QUAlity and Transparency Of Health Research [Internet]. [cited 2020 Feb 14]. Available from: <https://www.equator-network.org/>
134. [Acquiescence Response Bias](#). In: Encyclopedia of Survey Research Methods. 2455 Teller Road, Thousand Oaks California 91320 United States of America: Sage Publications, Inc.; 2008.

135. STM Association. STM Global Brief 2021 – Economics & Market Size [Internet]. [cited 2023 Jul 31]. Available from: https://www.stm-assoc.org/wp-content/uploads/2022_08_24_STM_White_Report_a4_v15.pdf
136. Allison R, Hayes C, McNulty CAM, Young V. **A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb**. JMIR Formative Research. 2019 Oct;3(4):e14372.
137. Google Analytics [Internet]. Google Marketing Platform. [cited 2023 Jul 31]. Available from: https://marketingplatform.google.com/intl/en_uk/about/analytics/
138. Usage Statistics and Market Share of Traffic Analysis Tools for Websites, July 2023 [Internet]. [cited 2023 Jul 31]. Available from: https://w3techs.com/technologies/overview/traffic_analysis
139. [GA4] Automatically collected events - Firebase Help [Internet]. [cited 2023 Jul 31]. Available from: https://support.google.com/firebase/answer/9234069?sjid=11808073354477924035-EU&visit_id=638263997154270219-1141820337&rd=1
140. [GA4] Predefined user dimensions - Firebase Help [Internet]. [cited 2023 Jul 31]. Available from: https://support.google.com/firebase/answer/9268042?sjid=11808073354477924035-EU&visit_id=638263997154270219-1141820337&rd=1
141. Web & Mobile Tag Management Solutions – Google Tag Manager [Internet]. Google Marketing Platform. [cited 2023 Jul 31]. Available from: https://marketingplatform.google.com/intl/en_uk/about/tag-manager/
142. Popup Builder That Boosts Sales. [Internet]. Popup Smart. [cited 2023 Jul 31]. Available from: <https://popupsmart.com/>

143. Scimago Journal & Country Rank [Internet]. [cited 2020 Nov 6]. Available from: <https://www.scimagojr.com/>
144. Ecommerce Foundation. Ecommerce Benchmark Retail Report 2016 [Internet]. Ecommerce Europe. [cited 2023 Aug 2]. Available from: <https://www.ecommerce-europe.eu/wp-content/uploads/2016/06/Ecommerce-Benchmark-Retail-Report-2016.pdf>
145. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. **A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance.** BMJ. 2021 Sep;n2061.
146. Yardley L, Morrison L, Bradbury K, Muller I. **The Person-Based Approach to Intervention Development: Application to Digital Health-Related Behavior Change Interventions.** Journal of Medical Internet Research. 2015 Jan;17(1):e4055.
147. Cane J, O'Connor D, Michie S. **Validation of the theoretical domains framework for use in behaviour change and implementation research.** Implementation Science. 2012 Apr;7(1):37.
148. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. **Making psychological theory useful for implementing evidence based practice: A consensus approach.** BMJ Quality & Safety. 2005 Feb;14(1):26–33.
149. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. **A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems.** Implementation Science. 2017 Jun;12(1):77.
150. Michie S, van Stralen MM, West R. **The behaviour change wheel: A new method for characterising and designing behaviour change interventions.** Implementation Science. 2011 Apr;6(1):42.

151. Dolan P, Hallsworth M, Halpern D, King D, Vlaev I. MINDSPACE: Influencing behaviour for public policy [Internet]. Institute of Government; 2010 [cited 2024 Oct 11]. Available from: <https://www.instituteforgovernment.org.uk/sites/default/files/publications/MINDSPACE.pdf>
152. Mook DG. Motivation : The organization of action [Internet]. W.W. Norton; 1996 [cited 2023 Jul 14]. Available from: <https://cir.nii.ac.jp/crid/1130282269635015808>
153. Fishbein M, Triandis HC, Kanfer FH, Becker M, Middlestadt SE, Eichler A, et al. Factors influencing behavior and behavior change. *Handbook of health psychology*. 2001;3(1):3–17.
154. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions [Internet]. London: Silverback Publishing; 2014. Available from: www.behaviourchangewheel.com
155. M.Given L. **Constructivism**. In: The SAGE Encyclopedia of Qualitative Research Methods. SAGE Publications, Inc.; 2008. p. 116–20.
156. Lincoln YS, Guba EG. Naturalistic Inquiry. SAGE; 1985.
157. W.Stewart D, N.Shamdasani P, W.Rook D. **Focus Groups**. SAGE Publications, Ltd.; 2007.
158. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. **The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions**. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*. 2013 Aug;46(1):81–95.

159. Carvalho F. [Participatory design for behaviour change: An integrative approach to improving healthcare practice focused on staff participation](#) [Thesis]. Loughborough University; 2020.
160. Parrillo V. [Groupthink](#). In: Encyclopedia of Social Problems. Thousand Oaks: SAGE Publications, Inc.; 2008. p. 421–1.
161. Given L. [Focus Groups](#). In: The SAGE Encyclopedia of Qualitative Research Methods. Thousand Oaks: SAGE Publications, Inc.; 2008. p. 353–4.
162. Malterud K, Siersma VD, Guassora AD. [Sample Size in Qualitative Interview Studies: Guided by Information Power](#). Qualitative Health Research. 2016 Nov;26(13):1753–60.
163. O'Reilly M, Parker N. [“Unsatisfactory Saturation”: A critical exploration of the notion of saturated sample sizes in qualitative research](#). Qualitative Research. 2013 Apr;13(2):190–7.
164. Braun V, Clarke V. [To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales](#). Qualitative Research in Sport, Exercise and Health. 2021 Mar;13(2):201–16.
165. S.Lewis-Beck M, Bryman A, Liao TF. [Social Desirability Bias](#). In: The SAGE Encyclopedia of Social Science Research Methods. Sage Publications, Inc.; 2004.
166. McGrath C, Palmgren PJ, Liljedahl M. [Twelve tips for conducting qualitative research interviews](#). Medical Teacher. 2019 Sep;41(9):1002–6.

167. Allen R, Currie JL. Think/Pair/Share. In: U-Turn Teaching: Strategies to Accelerate Learning and Transform Middle School Achievement [Internet]. Thousand Oaks, UNITED STATES: Corwin Press; 2012 [cited 2024 Oct 11]. p. 136. Available from: <http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=1109180>
168. Bradshaw C, Atkinson S, Doody O. **Employing a Qualitative Description Approach in Health Care Research.** Global Qualitative Nursing Research. 2017 Jan;4:2333393617742282.
169. Kim H, Sefcik JS, Bradway C. **Characteristics of Qualitative Descriptive Studies: A Systematic Review.** Research in nursing & health. 2017 Feb;40(1):23–42.
170. QSR International Pty Ltd. NVivo [Internet]. 2020. Available from: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home>
171. JISC. Online surveys [Internet]. [cited 2024 May 21]. Available from: <https://www.onlinesurveys.ac.uk/>
172. Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, Rao SM. **Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report.** American Psychologist. 2018 Jan;73(1):3–25.
173. TIDieR | Author tool [Internet]. [cited 2024 May 20]. Available from: <https://tidiergeguide.org/#/author-tool>
174. <https://www.skillzone.net> SZL. Definition: Paired depth interview [Internet]. Association for Qualitative Research (AQR). [cited 2024 May 20]. Available from: <https://www.aqr.org.uk/glossary/paired-depth-interview>

175. <https://www.skillzone.net> SZL. Definition: Triad [Internet]. Association for Qualitative Research (AQR). [cited 2024 May 20]. Available from: <https://www.aqr.org.uk/glossary/triad>
176. L.Morgan D. [Focus Groups as Qualitative Research](#). SAGE Publications, Inc.; 1997.
177. [Co-creation, co-design, co-production for public health – a perspective on definitions and distinctions - June 2022, Volume 32, Issue 2 | PHRP.](https://www.phrp.com.au/) <https://www.phrp.com.au/>. 2022.
178. Pastel | Fastest visual website feedback tool for web designers, developers and agencies [Internet]. [cited 2023 Oct 3]. Available from: <https://usepastel.com/>
179. The NHS website [Internet]. 16 Aug 2018, 12:27 a.m. [cited 2023 Oct 3]. Available from: <https://www.nhs.uk/>
180. ISO - International Organization for Standardization [Internet]. ISO. 2023 [cited 2023 Oct 3]. Available from: <https://www.iso.org/home.html>
181. NICE | The National Institute for Health and Care Excellence [Internet]. NICE. NICE; [cited 2023 Oct 3]. Available from: <https://www.nice.org.uk/>
182. The largest icon set in the world. [Internet]. Smashicons. [cited 2023 Oct 3]. Available from: <http://www.smashicons.com/>
183. Download Free Vectors, Images, Stock Photos & Stock Videos [Internet]. Vecteezy. [cited 2023 Oct 3]. Available from: <https://www.vecteezy.com/>

184. Free Icons and Stickers - Millions of images to download [Internet]. Flaticon. [cited 2023 Oct 3]. Available from: <https://www.flaticon.com/>
185. Freepik: Download Free Videos, Vectors, Photos, and PSD [Internet]. Freepik. [cited 2023 Oct 3]. Available from: <https://www.freepik.com>
186. Nielsen Norman Group: UX Training, Consulting, & Research [Internet]. Nielsen Norman Group. [cited 2023 Oct 5]. Available from: <https://www.nngroup.com/>
187. Developing and wording guideline recommendations | The guidelines manual | Guidance | NICE [Internet]. NICE; 2012 [cited 2024 Oct 10]. Available from: <https://www.nice.org.uk/process/pmg6/chapter/developing-and-wording-guideline-recommendations#wording-the-guideline-recommendations>
188. ISO 9241-210:2019(en), Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems [Internet]. [cited 2023 Oct 6]. Available from: <https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-210:ed-2:v1:en>
189. Velsen LS van, Gemert-Pijnen JEWC van, Nijland N, Beaujean D, Steenbergen J van. Personas: The Linking Pin in Holistic Design for eHealth. In: eTELEMED 2012 : The Fourth International Conference on eHealth, Telemedicine, and Social Medicine. 2012. p. 128–33.
190. Personas Make Users Memorable for Product Team Members [Internet]. Nielsen Norman Group. [cited 2023 Oct 6]. Available from: <https://www.nngroup.com/articles/persona/>

191. Personas | Health Education England [Internet]. Health Education England | Digital Transformation. [cited 2023 Oct 6]. Available from: <https://digital-transformation.hee.nhs.uk/building-a-digital-workforce/dart-ed/horizon-scanning/ai-and-digital-healthcare-technologies/applying-the-framework-and-next-steps/personas>
192. Your website, your business, your futureWix.com [Internet]. wix.com. [cited 2023 Aug 4]. Available from: <https://www.wix.com>
193. Website Builder — Create a Website in Minutes [Internet]. Squarespace. [cited 2023 Aug 4]. Available from: <https://www.squarespace.com/>
194. The RStudio Integrated Development Environment (IDE) is the preferred tools for data scientists who develop in R & Python. [Internet]. Posit. [cited 2023 Aug 4]. Available from: <https://posit.co/products/open-source/rstudio/>
195. Quarto: An open source technical publishing system for creating beautiful articles, websites, blogs, books, slides, and more. Supports Python, R, Julia, and JavaScript. [Internet]. Quarto. [cited 2023 Aug 4]. Available from: <https://quarto.org/>
196. Harwood J. EQUATOR Guidelines Website Repository [Internet]. 2023 [cited 2023 Oct 4]. Available from: <https://github.com/jamesrharwood/equator-guidelines-website>
197. GitHub Pages [Internet]. [cited 2023 Aug 4]. Available from: <https://pages.github.com/>
198. Rachael Cayley. Explorations of Style [Internet]. A Blog about Academic Writing. [cited 2024 Mar 21]. Available from: <https://explorationsofstyle.com/>

199. Sword H. Stylish Academic Writing [Internet]. Harvard University Press; 2012 [cited 2024 Mar 21]. Available from: <https://www.jstor.org/stable/j.ctt2jbw8b>
200. Wallach D, Scholz SC. *User-Centered Design: Why and How to Put Users First in Software Development*. In: Maedche A, Botzenhardt A, Neer L, editors. Software for People: Fundamentals, Trends and Best Practices. Berlin, Heidelberg: Springer; 2012. p. 11–38.
201. Nielsen J, Landauer TK. *A mathematical model of the finding of usability problems*. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. New York, NY, USA: Association for Computing Machinery; 1993. p. 206–13. (CHI '93).
202. Paul Doncaster. *The UX Five-Second Rules*. In: The UX Five-Second Rules: Guidelines for User Experience Design's Simplest Testing Technique. Morgan Kaufmann; 2014. p. 19–76.
203. Boren T, Ramey J. *Thinking aloud: Reconciling theory and practice*. IEEE Transactions on Professional Communication. 2000 Sep;43(3):261–78.
204. De Jong M, Schellens PJ. *Reader-Focused Text Evaluation: An Overview of Goals and Methods*. Journal of Business and Technical Communication. 1997 Oct;11(4):402–32.
205. Magaldi D, Berler M. *Semi-structured Interviews*. In: Zeigler-Hill V, Shackelford TK, editors. Encyclopedia of Personality and Individual Differences. Cham: Springer International Publishing; 2020. p. 4825–30.
206. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. *Influences on the Uptake of Health and Well-being Apps and Curated App Portals: Think-Aloud and Interview Study*. JMIR mHealth and uHealth. 2021 Apr;9(4):e27173.

207. Yardley L, Morrison LG, Andreou P, Joseph J, Little P. [Understanding reactions to an internet-delivered health-care intervention: Accommodating user preferences for information provision](#). BMC Medical Informatics and Decision Making. 2010 Sep;10(1):52.
208. Ericsson KA, Simon HA. [Protocol Analysis: Verbal Reports as Data](#). The MIT Press; 1993.
209. Hypothes.is [Internet]. Hypothesis. [cited 2024 May 5]. Available from: <https://web.hypothes.is/>
210. ARRIVE Guidelines [Internet]. [cited 2024 May 5]. Available from: <https://arriveguidelines.org/>
211. Craven J, Nietzio A. [A task-based approach to assessing the accessibility of web sites](#). Performance Measurement and Metrics. 2007 Jan;8(2):98–109.
212. de Jong M, Schellens PJ. [Toward a document evaluation methodology: What does research tell us about the validity and reliability of evaluation methods?](#) IEEE Transactions on Professional Communication. 2000 Sep;43(3):242–60.
213. Google Search Console [Internet]. [cited 2024 May 5]. Available from: <https://search.google.com/search-console/about>
214. ESRC to radically expand UK behavioural research capacity [Internet]. 2023 [cited 2024 Jan 17]. Available from: <https://www.ukri.org/news/esrc-to-radically-expand-uk-behavioural-research-capacity/>
215. Pelly M, Fatehi F, Liew D, Verdejo-Garcia A. [Novel behaviour change frameworks for digital health interventions: A critical review](#). Journal of Health Psychology. 2023 Sep;28(10):970–83.

216. Garnett C, Crane D, West R, Brown J, Michie S. [The development of Drink Less: An alcohol reduction smartphone app for excessive drinkers](#). Translational Behavioral Medicine. 2019 Mar;9(2):296–307.
217. Garnett C, Perski O, Michie S, West R, Field M, Kaner E, et al. [Refining the content and design of an alcohol reduction app, *Drink Less*, to improve its usability and effectiveness: A mixed methods approach](#). F1000Research; 2021.
218. Mills T, Lawton R, Sheard L. [Advancing complexity science in healthcare research: The logic of logic models](#). BMC Medical Research Methodology. 2019 Mar;19(1):55.

Appendix A

PRISMA-S Checklist for chapter 3

TABLE A.1: PRISMA-S Checklist for chapter 3 [86]

No.	Item	Description	Location
1	Database name	Name each individual database searched, stating the platform for each.	See Methods Table 3.1 and Appendix C
2	Multi-database searching	If databases were searched simultaneously on a single platform, state the name of the platform, listing all of the databases searched.	See Methods Table 3.1 and Appendix C
3	Study registries	List any study registries searched.	See Methods Table 3.1 and Appendix C
4	Online resources and browsing	Describe any online or print source purposefully searched or browsed (e.g., tables of contents, print conference proceedings, web sites), and how this was done.	See Methods Table 3.1 and Appendix C

No.	Item	Description	Location
5	Citation searching	Indicate whether cited references or citing references were examined, and describe any methods used for locating cited/citing references (e.g., browsing reference lists, using a citation index, setting up email alerts for references citing included studies).	See Methods, Approach to searching and data sources
6	Contacts	Indicate whether additional studies or data were sought by contacting authors, experts, manufacturers, or others.	See Methods, Approach to searching and data sources
7	Other methods	Describe any additional information sources or search methods used.	No other methods
8	Full search strategies	Include the search strategies for each database and information source, copied and pasted exactly as run.	Appendix C

No.	Item	Description	Location
9	Limits and restrictions	Specify that no limits were used, or describe any limits or restrictions applied to a search (e.g., date or time period, language, study design) and provide justification for their use.	See Methods, Inclusions and exclusion criteria.
10	Search filters	Indicate whether published search filters were used (as originally designed or modified), and if so, cite the filter(s) used.	See Methods, Electronic search strategy
11	Prior work	Indicate when search strategies from other literature reviews were adapted or reused for a substantive part or all of the search, citing the previous review(s).	See Methods, Electronic search strategy
12	Updates	Report the methods used to update the search(es) (e.g., rerunning searches, email alerts).	No update done.
13	Dates of searches	For each search strategy, provide the date when the last search occurred.	See Methods, Table 3.1
14	Peer review	Describe any search peer review process.	See Methods, Electronic Search Strategy and Appendix C

No.	Item	Description	Location
15	Total records	Document the total number of records identified from each database and other information sources.	See Results, Figure 3.1
16	Deduplication	Describe the processes and any software used to deduplicate records from multiple database searches and other information sources.	See Methods Table 3.1

Appendix B

ENTREQ Checklist for chapter 3

TABLE B.1: ENTREQ Checklist [87]

No.	Item	Description	Location
1	Aim	State the research question the synthesis addresses.	See Introduction, final sentence of para. 4
2	Synthesis methodology	Identify the synthesis methodology or theoretical framework which underpins the synthesis, and describe the rationale for choice of methodology (e.g. meta-ethnography, thematic synthesis, critical interpretive synthesis, grounded theory synthesis, realist synthesis, meta-aggregation, meta-study, framework synthesis).	See Methods, Synthesis Methodology para. 1

No.	Item	Description	Location
3	Approach to searching	Indicate whether the search was pre-planned (comprehensive search strategies to seek all available studies) or iterative (to seek all available concepts until they theoretical saturation is achieved).	See Methods, Approach to Searching
4	Inclusion criteria	Specify the inclusion/exclusion criteria (e.g. in terms of population, language, year limits, type of publication, study type).	See Methods, Inclusion and Exclusion criteria

No.	Item	Description	Location
5	Data sources	<p>Describe the information sources used (e.g. electronic databases (MEDLINE, EMBASE, CINAHL, psycINFO, Econlit), grey literature databases (digital thesis, policy reports), relevant organisational websites, experts, information specialists, generic web searches (Google Scholar) hand searching, reference lists) and when the searches conducted; provide the rationale for using the data sources.</p>	See Methods Table 3.1
6	Electronic Search strategy	<p>Describe the literature search (e.g. provide electronic search strategies with population terms, clinical or health topic terms, experiential or social phenomena related terms, filters for qualitative research, and search limits).</p>	Appendix C

No.	Item	Description	Location
7	Study screening methods	Describe the process of study screening and sifting (e.g. title, abstract and full text review, number of independent reviewers who screened studies).	See Methods, Screening
8	Study characteristics	Present the characteristics of the included studies (e.g. year of publication, country, population, number of participants, data collection, methodology, analysis, research questions).	See Results, Table 3.3
9	Study selection results	Identify the number of studies screened and provide reasons for study exclusion (e.g. for comprehensive searching, provide numbers of studies screened and reasons for exclusion indicated in a figure/flowchart; for iterative searching describe reasons for study exclusion and inclusion based on modifications to the research question and/or contribution to theory development).	See Results, Figure 3.1

No.	Item	Description	Location
10	Rationale for appraisal	Describe the rationale and approach used to appraise the included studies or selected findings (e.g. assessment of conduct (validity and robustness), assessment of reporting (transparency), assessment of content and utility of the findings).	See Methods, Describing and appraising records
11	Appraisal items	State the tools, frameworks and criteria used to appraise the studies or selected findings (e.g. Existing tools: CASP, QARI, COREQ, Mays and Pope [25]; reviewer developed tools; describe the domains assessed: research team, study design, data analysis and interpretations, reporting).	See Methods, Describing and appraising records
12	Appraisal process	Indicate whether the appraisal was conducted independently by more than one reviewer and if consensus was required.	See Methods, Describing and appraising records

No.	Item	Description	Location
13	Appraisal results	Present results of the quality assessment and indicate which articles, if any, were weighted/excluded based on the assessment and give the rationale.	See Results, Search and Table 3.3
14	Data extraction	Indicate which sections of the primary studies were analysed and how were the data extracted from the primary studies? (e.g. all text under the headings “results /conclusions” were extracted electronically and entered into a computer software).	See Methods, Synthesis Methodology, para. 2
15	Software	State the computer software used, if any.	See Methods, Synthesis Methodology, para. 2
16	Number of reviewers	Identify who was involved in coding and analysis.	See Methods, Synthesis Methodology, para. 2
17	Coding	Describe the process for coding of data (e.g. line by line coding to search for concepts).	See Methods, Synthesis Methodology, para. 2

No.	Item	Description	Location
18	Study comparison	Describe how were comparisons made within and across studies (e.g. subsequent studies were coded into pre-existing concepts, and new concepts were created when deemed necessary).	See Methods, Synthesis Methodology, para. 2
19	Derivation of themes	Explain whether the process of deriving the themes or constructs was inductive or deductive.	See Methods, Synthesis Methodology, para. 2
20	Quotations	Provide quotations from the primary studies to illustrate themes/constructs, and identify whether the quotations were participant quotations of the author's interpretation.	See Results, Synthesis Findings
21	Synthesis output	Present rich, compelling and useful results that go beyond a summary of the primary studies (e.g. new interpretation, models of evidence, conceptual models, analytical framework, development of a new theory or construct).	See Results, Synthesis Findings

Appendix C

Search strategies

I did not seek any external peer review of my search. I did not record the database versions at the time of searching due to an oversight. I performed forward and backwards citation searching but found no additional records. I did not set up email alerts.

C.0.1 Ovid search strategy

Databases: Medline, Embase, AMED, PsycINFO.

Search date: 08/12/2021

I used a federated search. The *kw* field does not exist in PsycINFO or AMED and so was ignored by these databases. The *tw* field does not exist in AMED either and was mapped to *af* instead.

1. ((reporting or writ\$ or author\$) adj2 (checklist\$ or statement\$ or guid\$ or template\$ or standard\$ or recommendation\$)).ti,kw.
2. ((consort\$ or strobe\$ or stard\$ or prisma\$ or moose\$ or squire\$ or arrive\$ or remark\$ or tripod\$ or cheers\$ or spirit\$ or srqr\$ or coreq\$) adj3 (guid\$ or statement\$ or checklist\$)).ti,kw.
3. (experience\$ or interview\$ or survey\$ or questionnaire\$ or "focus group\$" or facilitat\$ or barrier\$).af.

4. qualitative.tw.

5. 1 or 2

6. 3 or 4

7. 5 and 6

Platform-specific filter applied: 1996 – current year

C.0.2 Global Index Medicus & SciELO search strategy

Databases: Latin American and Caribbean Health Sciences Literature, African Index Medicus, Western Pacific Region Index Medicus, Index Medicus for South-East Asia Region, and Index Medicus for the Eastern Mediterranean Region, searched using Global Index Medicus (<https://www.globalindexmedicus.net/>); Scientific Electronic Library Online (<https://scielo.org/en/>).

Search date: 08/12/2021

ti:(((reporting OR writ* OR author*) AND (checklist* OR statement* OR guid* OR template* OR standard* OR recommendation* OR experience* OR interview* OR survey* OR questionnaire* OR "focus group*" OR facilitat* OR barrier* OR qualitative*)))

Platform-specific filter applied: 1996 – current year

C.0.3 Chinese Biomedical Literature Database

Database URL: <https://www.imicams.ac.cn/>

Search Date: 25/10/2021

1. OR OR

2. OR OR OR OR OR OR OR

3. CONSORT OR PRISMA OR STROBE OR SPIRIT OR STARD OR SRQR OR ARRIVE OR SQUIRE OR CHEERS OR TRIPOD OR COREQ

4. OR OR OR OR OR OR
 5. OR
 6. OR
 7. (#2) AND (#1)
 8. (#7) OR (#3)
 9. (#6) OR (#5) OR (#4)
 10. (#9) AND (#8)
 11. ((#9) AND (#8)) AND ("[] OR "")[])

C.0.4 China National Knowledge Infrastructure

Database URL: <https://www.cnki.net/>

Search Date: 25/10/2021

(((TI = " OR TI = " OR TI = ") AND (TI = " OR TI = ")) OR (TI = 'CONSORT' OR TI = 'STROBE' OR TI = 'PRISMA' OR TI = 'SPIRIT' OR TI = 'STARD' OR TI = 'SRQR' OR TI = 'ARRIVE' OR TI = 'SQUIRE' OR TI = 'CHEERS' OR TI = 'TRIPOD' OR TI = 'COREQ')) AND (TI = " OR TI = ")

C.0.5 Wanfang Data

Database URL: <http://www.wanfangdata.com/>

C.0.6 VIP Chinese Medical Journal Database

Database URL: <http://www.cqvip.com/>

((M=(OR OR)) AND (M=(OR OR OR OR OR OR OR OR))) OR (M=(CONSORT
OR PRISMA OR STROBE OR SPIRIT OR STARD OR SRQR OR ARRIVE OR SQUIRE OR
CHEERS OR TRIPOD OR COREQ))) AND (M=(OR OR OR OR OR OR OR))

C.0.7 OSF

URL: <https://osf.io/>

Search Date: 15/12/2021

title:(((reporting OR writ* OR author*) AND (checklist* OR statement* OR guid* OR tem-
plate* OR standard* OR recommendation* OR experience* OR interview* OR survey*
OR questionnaire* OR "focus group*" OR facilitat* OR barrier* OR qualitative*)))

C.0.8 Methods in Research on Research

URL: <http://miror-ejd.eu/publications/>

Search Date: 14/12/2021

I manually searched the list of publications.

Appendix D

SRQR Checklist for chapter 7

TABLE D.1: SRQR Checklist [30]

No.	Item	Description	Location
1	Title	Concise description of the nature and topic of the study. Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended.	Title of chapter

No.	Item	Description	Location
2	Abstract	Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	N/A
3	Problem Formation	Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	Introduction and previous chapters
4	Purpose or research question	Purpose of the study and specific objectives or questions	Introduction para. 1 & 2
5	Qualitative approach and research paradigm	Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/interpretivist) is also recommended; rationale*	Methods para. 1 & 2 and chapter 1, Aims and Objectives, where I discuss pragmatism.

No.	Item	Description	Location
6	Researcher characteristics and reflexivity	Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	Methods para. 2 and see chapter 2
7	Context	Setting/site and salient contextual factors; rationale*	Methods para. 6
8	Sampling strategy	How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale*	Methods para. 1

No.	Item	Description	Location
9	Ethical issues pertaining to human subjects	Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	Methods, Ethics
10	Data collection methods	Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale*	Overall methods are reported in Methods. Stage-specific methods are reported in the Results subsections.
11	Data collection instruments and technologies	Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	See appendices E - K

No.	Item	Description	Location
12	Units of study	Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	Results subsections
13	Data processing	Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/deidentification of excerpts	Few workshops stages required processing or analysis. Stage specific processing steps are reported in Results subsections
14	Data analysis	Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale*	Stage specific analysis steps are reporting in Results subsections

No.	Item	Description	Location
15	Techniques to enhance trustworthiness	Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale*	Methods para. 4 and @tbl-trust-workshops
16	Synthesis and interpretation	Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	Results
17	Links to empirical data	Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	There were no analytic findings, but raw data (in the form of our co-edited worksheets) are in appendices E - K

No.	Item	Description	Location
18	Integration with prior work, implications, transferability, and contribution(s) to the field	Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/ generalizability; identification of unique contribution(s) to scholarship in a discipline or field	Discussion para. 1-4
19	Limitations	Trustworthiness and limitations of findings	Discussion para. 5 & 6
20	Conflicts of interest	Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	See chapter 2
21	Funding	Sources of funding and other support; role of funders in data collection, interpretation, and reporting	See Acknowledgements

Appendix E

BCW Step 1 – Define the problem in behavioural terms

Instructions [154]:

Define the problem in behavioural terms , including:

- *What behaviour? e.g., Improve hygiene practices in all opportunities identified by national or local guidelines*
- *Where does the behaviour occur? e.g., Hospital wards*
- *Who is involved in performing the behaviour? e.g., Hospital nursing staff*

JdB, PL: To report completely in a journal article. When people are writing a journal article we want them to add in important details that all users will need to judge and understand the work. Occurs when writing articles (e.g., universities, RIs, hospitals, editing firms). Involves anybody writing about biomedical research.

CS, SK: Writing up research. Happens at work places/homes. Solitary activity.

JH: Report research in line with guidelines. Work computer. Biomedical researchers.

Appendix F

BCW Step 2 – Select the Target Behaviour

Instructions [154]:

Select the target behaviour

Generate a long list of candidate target behaviours that could bring about the desired outcome. This list can include the behaviour of other people too e.g., editors, guideline developers.

E.g.:

- *Keep sterile items in packaging until use*
- *Wash patients as appropriate*
- *Clean wards, toilets, offices, theatres*
- *Clean hands with soap*
- *Clean hands with alcohol gel*

SK, CS

- Supervisors must encourage and endorse use of guideline
- Authors could ask for help when writing
- RG developers could respond to feedback / help authors

(JdB, PL)

- Find appropriate guidance/guidelines
- Read guidance (especially E&E) in full
- Use guidance when planning
- Use guidance when drafting (subheadings/comments)
- Peer reviewers check articles against guidelines, tell authors what is missing, insisting missing items are added
- Editors & journal staff check articles against guidelines, tell authors what is missing, insisting missing items are added. Additionally, check and support peer reviewer behaviours

JH

- Study guidance (in an abstract way)
- Considering guidance when planning
- Drafting - Reporting guidance content in funding proposals/drafting
- Editing - Checking articles against guidance before submitting to a journal
- Reviewing – checking articles written by someone else against guidance (editor, peer reviewer , co-author)

Retrospective thoughts:

- JH: Could have included the behaviour of funders e.g., requesting guidance adherence
- SK: Could have specified using guidance for writing protocols, but not everyone writes a protocol
- JH: We could distinguish between using an E&E and using a checklist. E.g., “Authors should use E&E document when drafting” v.s. “Authors should use checklist when drafting”. Or maybe distinguishing between checklist & E&E can come in step 4.

Instructions: Prioritise the behaviours.

- *Impact: How much of an impact changing the behaviour will have on the desired outcome*
- *Likelihood: How likely it is that the behaviour can be changed (think about the capability, opportunity, and motivation to change of those performing the behaviour)*
- *Spill-over: How likely it is that the behaviour (or group of behaviours) will have a positive or negative impact on other, related behaviours*
- *Measurement: How easy it will be to measure the behaviour*

Key: Unacceptable (1), unpromising but worth considering (2), promising (3), very promising (4)

TABLE F.1: Behaviours scored for impact, likelihood, spill-over, and measurement

Potential target behaviour	Impact	Likelihood	Spill-over	Measurement
Supervisors encourage use of guidance	3,4,4,4,2	2,2,3,2,2	4,4,4,3,4	4,1,2,2,1
Authors ask for help when writing	3,3,3,4	2,3,3,3	3,4,3,3	4,2,2,2
Find appropriate guidance	4,3,4,3	3,3,3,2	3,2,4,2	4,3,3,2
Read/study guidance in full (esp E&E)	4,3,3,2	2,2,3,2	4,2,3,2	1,1,1,2
Consider guidance when planning	3,3,3,3	2,2,2,3	3,3,3,3	1,2,1,2
Report guidance when writing funding applications	3,3,3,3	3,2,3,3	3,3,3,3	4,3,4,3
Report guidance when drafting (subheadings/comments)	4,3,4,4	3,3,3,2	3,1,3,2	4,2,2,3
Check submitted articles against guidance & insist on missing items being added (peer reviewer)	4,4,4,3	3,2,4,2	4,1,4,3	4,4,3,3
Check submitted articles against guidance & insist on missing items being added (editor)	4,4,4,3	3,2,4,2	4,1,4,3	2,1,2,3
Check article drafts against guidance during preparation and insist missing items are added (co-author)	4,4,4,2	2,2,3,2	4,1,3,2	1,1,2,2

Target behaviour selected:

No clear winner so we decided to take a few different target behaviours on to step 3.

Appendix G

BCW Step 3 – Specify the Target Behaviour

Instructions [154]:

Specify the target behaviour

- Who needs to perform the behaviour?
- What does the person need to do differently in order to achieve the desired change?
- When will they do it?
- Where will they do it? (Research institute, office etc).

- How often will they do it?
- With whom will they do it?

CS

Lead researcher use RGs (and other EQUATOR resources) to plan research, when planning and again throughout the project, for each piece of research, with collaborators.

MS

First authors will use RG finder tools to find appropriate RG when they begin drafting a manuscript, for each paper they write, on their own with collaborators.

Lead researcher use RG to write funding applications, when applying for funding, for each applications, with collaborators.

ECRs seek external help to support their writing, when drafting papers (until they feel confident enough to not need help), every paper, on their own, with coauthors.

JdB

First authors need to use RGs to plan a manuscript before starting to write, for every paper, with co-authors support and approval

Peer reviewers must request missing items, for each manuscript they review, with support from editors.

PL

Reviewers must use RGs when they check every manuscript for content, they must report about RGs use (or non-use, misuse) in their review reports, the place they do that is journals and liaise with editor, with journals requiring it first.

Co-authors should check RGs content items in all manuscripts they participate, during drafting/editing, on their own in context of collaborators.

All authors, editors, reviewers should read the E&E at least once in their life time

SK

Authors should read E&E (not just the checklist), as early as possible in the research pipeline (ideally before writing), for each study and at multiple points within each study, on their own but in collaborative context

JH

First authors should fully report each item when writing, as early as possible in the writing process, for each piece of research they write.

Co-authors should add missing items when reviewing research, as early as possible in editing process, for every piece of research they are involved in.

Peer reviewers should request missing items when reviewing research, for every article they review.

Research teams should consider items when planning their research reporting, for each piece of research they plan.

General themes

There were some themes that came out of this step & the previous one.

- Authoring – the earlier the better. Use E&E over and above the checklist
- Co-authoring – use to check work you are involved with
- Reviewing – use to review reporting of work you aren't involved with
- Planning reporting – use to plan how you will report something
- Asking for help – seeking help and advice with writing

Retrospective thoughts:

Are we too quick to talk about checklists & E&Es. I know they exist, but perhaps there is a better format to deliver guidance. I used the abbreviation RG for convenience (which we all probably read as Reporting Guideline) but perhaps I should have used “reporting guidance” instead to clarify that we are still talking about the conceptual guidance, not a particular document.

Is “report each item” actually too strict? Some guidelines specify that authors need not report every item (e.g. GRIPP2) but still want authors to consider each item. I suppose other guideline groups are more strict and want every item to be reported, even if it didn’t happen or isn’t very important to that particular study. Maybe we re-word to “adhere to guidance” or “report all necessary items”.

How early can RGs be used, in reality? Is design stage too early? Is there a sweet spot?

Proposed revision:

Problem behaviour: We want research teams to adhere to guidance when reporting their work.

Target behaviour:

Researchers should use reporting guidance as early as possible in their research pipeline. They will do this for every piece of research, at their place of work, on their own but in the context of collaboration.

Justification:

We've distinguished between considering guidance and actually writing. I know these are two separate behaviours but they are so coupled that I think we can combine them.

"As early as possible" – keeping this vague. It might differ between guidelines. The main point is that this needs to happen way before editing/submission.

We've removed mention of E&Es and checklists. Because maybe in our blue sky thinking we want to do away with these documents as neither do a great job of communicating guidance.

Appendix H

BCW Step 4 – Identifying what needs to change

Instructions [154] :

When it comes to an individual/group/population doing/not doing X, use evidence and theory to form a judgement about which of the following needs to change and in each case whether it should be targeted directly or through one of the other elements of the COM-B system. Note that it might be appropriate to apply this analysis to one or more supportive or competing behaviours.

TABLE H.1: Workshop participants' ideas around what needs to change for authors to perform our target behaviour. Based on the COM-B questionnaire [154]

AUTHORS WOULD HAVE TO...	
Capability	

AUTHORS WOULD HAVE TO...	
Know more about why it was important	e.g., better understand the benefits of stopping smoking <ul style="list-style-type: none"> • Know the importance of research waste PL • Understand the benefits of clearly reporting (to themselves and the community). These might include smoother review, citations, makes their work more useful to the community. JdB, SK <ul style="list-style-type: none"> • Authors need to know why items are important and who they are important to JH • Understand how a RG can help them with their personal goals
Know more about how to do it	e.g., better understand effective ways to lose weight <ul style="list-style-type: none"> • Know that guidance exists PL, JH • Better understanding of how to find guidance JdB,SK • Better understanding of how to use guidance JdB • Better understanding of how to use items to improve writing SK • Need to know what to actually report, how to report it, and how to keep writing fluid and concise SK, JH • How to report “imperfect” things • Authors need to know where they can report items JH • Know the process of writing (maybe not requirement) JdB • Understand how a team can use RGs in practice CS • Authors need to know what reporting guidance is, and how it is different to design guidance or design requirements • Authors need to know which guideline applies to their work, JH • Authors need to know which guidelines exist, JH
Have better physical skills	e.g., learn how to operate machinery

AUTHORS WOULD HAVE TO...	
Have better mental skills	e.g., learn how to reason more effectively <ul style="list-style-type: none"> • Be able to understand the guidance (language) PL • Be able to understand the guidance (conceptually) PL, SK, JH • Have a good writing process (but maybe this isn't a requirement) (JdB, JH) including time management CS • Know how to write/communicate clearly JdB, CS (could be delivered through training, games) • Learning how to search for reporting guidance SK • Better knowledge of designs PL
Have more physical strength	e.g., build up muscle for demanding physical work
Have more mental strength	e.g., develop stronger resilience against cravings <ul style="list-style-type: none"> • Knowing how to stand up to objections from co-authors, editors and peer reviewers JdB, CS • Authors need to be open to new ways of working, or changes to the way they normally write/do their research SK
Overcome physical limitations	e.g., get around problems of stature or disability
Overcome mental obstacles	e.g., reduce unwanted urges or feelings <ul style="list-style-type: none"> • Overcome the feeling that good reporting is too hard or not something I need to do. Not for me. JdB • Overcome the feeling that RGs are an extra burden SK
Have more physical stamina	e.g., develop greater capacity to maintain physical effort

AUTHORS WOULD HAVE TO...	
Have more mental stamina	e.g., develop greater capacity to maintain mental effort <ul style="list-style-type: none"> • Persisting with good reporting even if it feels like it takes longer, or in the face of objection from others. JdB, CS • Being dogged (persisting, keep on keeping on) • Authors need to know that it's never too late (in your career) to make improvements in how to write. SK
Opportunity	
Have more time to do it	e.g., create dedicated time during the day <ul style="list-style-type: none"> • Time management • Authors encounter guidance at the very end of the writing process which is when they have least time • Authors need to ask for time for writing CS
Have more money	e.g., be given or earn funds to support the behaviour <ul style="list-style-type: none"> • Budget for writing time and writing resources/help e.g., a publications coordinator, CS
Have the necessary materials	e.g., acquire better tools for the job <ul style="list-style-type: none"> • Translations PL
Have it more easily accessible	e.g., provide easier access to facilities <ul style="list-style-type: none"> • Fixing broken links PL • Guidance is split between documents that aren't linked • Examples are often separate to the checklists and hard to discover • Guidance behind paywall • Navigability of EQUATOR site, JH • Guidance can be buried within development or E&E documents or in figures JH, JdB • Guidance should be in a usable format JdB, JH • Authors must have access to all the information about their own study JdB

AUTHORS WOULD HAVE TO...

Have more people around them doing it	e.g., be part of a 'crowd' who are doing it <ul style="list-style-type: none"> • For it to become a norm in their community JdB • Be aware that others are using guidelines – belonging to a crowd JH, PL
Have more triggers to prompt them	e.g., have more reminders at strategic times <ul style="list-style-type: none"> • Being reminded in the early stages of writing (e.g., text books, social, journal pages, funders, automated personal reminders after interacting with other services) JH, SK, JdB, CS
Have more support from others	e.g., have one's family or friends behind one <ul style="list-style-type: none"> • Sometimes others can discourage use of guidelines JH • Having others (supervisors, librarians, editors, peer reviewers) support the behaviour JdB • Someone to give support for each guideline CS

Motivation

Feel that they want to do it enough	e.g., feel more of a sense of pleasure or satisfaction from doing it <ul style="list-style-type: none"> • Felt like cumbersome reports JH • Felt patronized JH • Felt like RGs are 'design straitjackets' JH • Felt afraid of reporting study limitations JH • Authors should believe that RGs will make their writing easier JdB, JH • Believe that RGs will help with career progression, SK
Feel that they need to do it enough	e.g., care more about the negative consequences of not doing it <ul style="list-style-type: none"> • Felt like it's the responsibility of the editor or reviewer JH • They should feel like it is their responsibility JdB
Believe that it would be a good thing to do	e.g., have a stronger sense that one should do it <ul style="list-style-type: none"> • Even if an author understands why a RG is important, they still need care (could be addressed through modelling) JH • Give recognition or reward to authors for using RGs PL • Authors don't think editors/reviewers use checklist info JH

AUTHORS WOULD HAVE TO...

Develop better plans for doing it	e.g., have clearer and better developed plans for achieving it • Have a plan for how and when they will use guidance JdB, JH
Develop a habit of doing it	e.g., get into a pattern of doing it without having to think • Need to have a habit of using guidance PL • Develop a habit that the first step of writing is to check for guidance

Other

(JH added):	• Guidance should be easier to understand JH, PL
Have it easier to do	• Items are not prioritised. Often the first items are generic background items that seem patronising. JH • Making websites easier to use PL • Simplification PL

Appendix I

BCW Step 4 – Consolidated list of influences needing to change for the target behaviour to occur

The following list of influences was generated from my thematic synthesis, review of survey questions, and my evaluation of EQUATOR’s website. I have labelled each influence with ideas to address them, generated in my workshops with reporting guideline experts, and my focus groups with publishers, guideline developers, and advocates.

I.1 Researchers may not know what reporting guidelines are

Researchers may have never heard the term “reporting guideline” or may misunderstand it. Researchers may more commonly use terms like “writing” or “writing up” and the word “reporting” may get interpreted as a formal task (such as reporting progress to a funder). The word “guideline” may be interpreted by some as rules (as per journal “author guidelines”) and others as recommendations. Some researchers may perceive reporting guidelines as a set of design requirements, especially if they only use checklists, which typically lack the instructions and nuances included in the full guidance.

Ideas to address this influence:

Describe reporting guidelines where they are encountered

Keep reporting guidelines agnostic to design choices

Promote reporting guidelines

Install reporting champions

I.2 Researchers may not know what reporting guidelines exist

Researchers may not be aware of which reporting guidelines exist. Most guidelines on the EQUATOR site are hardly ever accessed

Ideas to address this influence:

Show and encourage citations

Avoid confusing authors with too many reporting guidelines

Make resources easy to discover and find

Endorse and enforce reporting guidelines

Promote reporting guidelines

Describe each reporting guideline fully

I.3 Researchers may not know whether a reporting guideline applies to them

If the scope of a reporting guideline is undefined or unclear, then researchers won't know whether the guidance applies to them. Researchers may not understand study designs, making it difficult for them to identify which guidance applies.

Ideas to address this influence:

Describe reporting items fully

[Describe each reporting guideline fully](#)

I.4 Researchers may not know what reporting guideline is their best fit

Researchers may not know when more specific guidance exists. An author's "perfect fit" guideline may not exist, in which case they may not know when to stop searching, and they may try to use an "imperfect fit" guideline without understanding which items are applicable.

Ideas to address this influence:

[Avoid confusing authors with too many reporting guidelines](#)

[Make resources easy to discover and find](#)

[Describe each reporting guideline fully](#)

I.5 Researchers may not know what resources exist for a reporting guideline

Resources include the guidance itself, checklists, E&E files, templates, and web tools (e.g., PRISMA flow chart maker). Not all resources exist for each reporting guideline and researchers may be unaware of the ones that do. Many researchers may only use the checklist. Sometimes this is purposeful, but other times it may be because researchers don't know that full guidance and examples exist.

Ideas to address this influence:

[Make resources easy to discover and find](#)

I.6 Researchers may not know when reporting guidelines should be used

Researchers may not know when they should use reporting guidelines in their research workflow. Guideline developers may want researchers to use guidance as early as possible, but this is may not be obvious to researchers who may only ever receive instruction to complete a checklist as part of journal submission and may [never discover the full guidance](#). Consequently, researchers may assume that reporting guidelines are supposed to be used by single authors as pre-submission checklists to demonstrate adherence. It may not occur to them that reporting guidelines can be used earlier, or by teams. Some researchers, having come to this realisation themselves, report wanting to be told to use reporting guidelines earlier in their research.

Ideas to address this influence:

[Describe reporting guidelines where they are encountered](#)

[Create ways to catch authors earlier](#)

[Install reporting champions](#)

[Describe each reporting guideline fully](#)

I.7 Researchers may misunderstand

Researchers may not understand concepts, terms or words within the guidance, or they may understand them differently to how the developers intended. Some items (or entire guidelines) might be new concepts. e.g., SQUIRE guidelines written at a time where Quality Improvement was still a new concept to many people, and some items (e.g., Context, Study of the intervention) were less familiar than others. Researchers may have nowhere to turn for help should they not understand something.

Ideas to address this influence:

[Make reporting guidelines easy to understand](#)

Create discussion spaces

Install reporting champions

Provide additional teaching

Make updating guidelines easier

I.8 Researchers may not know what benefits to expect

Researchers may not know what benefits to expect from using a reporting guideline.

These benefits may include:

- improved completeness of reporting which helps readers use research and reduces research waste.
- improved flow and less “waffle” in writing
- facilitated discussions between collaborators, especially at the design or protocol stage
- publishing and passing peer review more efficiently
- increased publisher acceptance rates
- efficient, confident writing
- increased impact of manuscript, as the article is easier to search for and information within the article is easier to find.

Ideas to address this influence:

Describe reporting guidelines where they are encountered

Install reporting champions

Provide testimonials

I.9 Researchers may not know why items are important

Researchers may not know why an item is important, or who it is important to.

Ideas to address this influence:

Describe reporting guidelines where they are encountered

Describe reporting items fully

Install reporting champions

Provide additional teaching

I.10 Researchers may not know how to do an item

Researchers might not know how to do something (e.g., a sample size calculation)

Ideas to address this influence:

Create discussion spaces

Describe reporting items fully

I.11 Researchers may not know how to report an item in practice

Researchers may not understand how to report a particular item in practice

Ideas to address this influence:

Create discussion spaces

Describe reporting items fully

Provide additional teaching

I.12 Researchers may not know what to write when they cannot report an item

Researchers may not know how to report an item that they did not do (deliberately or as an oversight), or an item that they are unable to report for external reasons (e.g., IP, or data was missing from primary studies).

Ideas to address this influence:

[Describe reporting items fully](#)

I.13 Researchers have limited time

Guidelines take time to find, read, understand, and apply. Sometimes they may require time and work from multiple co-authors. Researchers & guideline developers may underestimate the time required for writing, and time is often most limited at the point of submission as grant funding may have run out.

Checklists take time to complete, and completing them with page numbers or pasted content can be annoying if future edits necessitate updating the checklist too. Checklists also generate work for editors and peer-reviewers who must cross check page numbers or pasted content with manuscript content.

Ideas to address this influence:

[Make resources ready-to-use](#)

[Budget for reporting](#)

[Create ways to catch authors earlier](#)

[Make information digestible](#)

[Describe reporting items fully](#)

[Describe each reporting guideline fully](#)

[Keep guidance short](#)

I.14 Researchers may not encounter reporting guidelines early enough to act on them

Some reporting guideline items require work that has to be done within a certain time windows such as:

- during planning or designing
- before or during data collection
- when other colleagues are available
- during the duration of a grant

Ideas to address this influence:

[Create reporting guidance for early stages of research](#)

[Create ways to catch authors earlier](#)

[Create additional tools](#)

[Provide additional teaching](#)

I.15 Researchers may not understand the language

Researchers may not understand the language guidance is written in. A lot of research comes from countries where English is not the first language, as do a lot of EQUATOR website visitors. Even if a researcher speaks English as a second language, language may be an additional barrier.

Ideas to address this influence:

[Make reporting guidelines easy to understand](#)

I.16 Researchers may struggle to keep writing concise

Following a guideline can result in lengthy, bloated reports which are unpleasant to read and breach journals' word limits. Researchers may not know how to keep writing fluid and concise or where they can report an item (e.g., what section, in the text or in a table or figure, in the manuscript or in supplementary material).

Ideas to address this influence:

[Avoid prescribing structure](#)

[Describe reporting items fully](#)

[Describe reporting items fully](#)

I.17 Researchers may not have tools for the job at hand

Researchers use reporting guidelines for different tasks and want tools to make that job easier. Researchers report using reporting guidelines for:

- Planning research
- Designing research
 - Researchers report wanting items presented in the order in which decisions need to be made
 - Researchers report wanting links to resources
- Whilst collecting data
 - Researchers report wanting items ordered in the order they are done
 - Researchers report wanting items embedded into data collection tools
- Drafting manuscript
 - Researchers report wanting templates
- Checking manuscripts
- Demonstrating compliance
 - Researchers report wanting checklists embedded into submission workflows
- Reviewing the reporting of other people's manuscripts
- Appraising the quality of other people's manuscripts

Ideas to address this influence:

[Create reporting guidance for early stages of research](#)

[Create additional tools](#)

I.18 Reporting guidelines can become outdated

Guidelines can become out of date compared to other guidance or compared to current research standards.

Ideas to address this influence:

[Make updating guidelines easier](#)

I.19 Researchers may struggle to reconcile multiple sets of guidance

Researchers must adhere to journal guidelines, multiple reporting guidelines (e.g., PRISMA + PRISMA-Abstracts + PRISMA-S) and other best practice guidelines (like NIH principles). Using multiple guidelines increases complexity and costs, and guidelines can contradict each other.

Ideas to address this influence:

[Avoid prescribing structure](#)

[Avoid confusing authors with too many reporting guidelines](#)

I.20 Researchers may be asked to remove reporting guideline content

Researchers may be asked to remove guideline content by co-researchers, editors or reviewers.

Ideas to address this influence:

[Describe reporting items fully](#)

I.21 Researchers may forget to use reporting guidelines at earlier research stages

Having been told to complete a checklist upon journal submission, researchers may forget to use a reporting guideline earlier next time.

NB forgetting is different to [not realising](#) that reporting guidelines can be used early.

Ideas to address this influence:

Create ways to catch authors earlier

I.22 Guidance may be difficult to find

Researcher should be able to easily find guidance and resources that they believe to exist. However:

- search functions can be hard to find or use,
- researchers may not know which search terms to use,
- websites may be hard to navigate,
- guidance can be buried within articles,
- resources may not be optimised for search engines,
- and resources may not be in the same place.

Ideas to address this influence:

Make resources easy to discover and find

Make information digestible

I.23 Reporting guidelines may be difficult to access

Researchers may be unable to access guidance published in subscription journals. Journal websites can feature broken links.

Ideas to address this influence:

Make resources accessible

I.24 Reporting guideline resources may not be in usable formats

Resources differ in how easy or readily usable they are. For example, some checklists are published as PDF tables that cannot be filled or copied. Some guidance can be dense,

unstructured text that is hard to digest or navigate; whereas some researchers will read the guidance sequentially, others may dip in and out whilst writing, and unstructured text can make information harder to find.

Ideas to address this influence:

[Make resources ready-to-use](#)

I.25 Researchers may feel afraid to report transparently

Researchers may feel afraid or uncertain when trying to report something that they didn't (or couldn't) do.

Ideas to address this influence:

[Keep reporting guidelines agnostic to design choices](#)

[Use persuasive language and design](#)

[Provide testimonials](#)

I.26 Researchers may feel restricted if reporting guidelines prescribe design

Advice or assumptions about design choices narrow the scope of the guidance and can make checklists appear prescriptive. Sometimes design assumptions can be implicit. For example, in requiring authors to report the method used to assess risk of bias, PRISMA is implying that authors should have designed their review to assess risk of bias.

Ideas to address this influence:

[Keep reporting guidelines agnostic to design choices](#)

I.27 Researchers may feel patronized

Researchers can feel patronized by checklists.

Ideas to address this influence:

Create discussion spaces

Use persuasive language and design

Describe each reporting guideline fully

I.28 Researchers may not believe stated benefits

Researchers may not believe that using a reporting guideline will affect their acceptance rate or publication speed, that using a reporting guideline will help them write, or improve the quality of their manuscript.

Ideas to address this influence:

Show and encourage citations

Create discussion spaces

Use persuasive language and design

Evidence the benefits

Make reporting guidelines appear as a priority

Provide testimonials

I.29 Researchers may not care about the benefits of using a reporting guideline

Researchers may understand that reporting guidelines aim to reduce poor reporting, but may not feel that poor reporting matters. Instead of hypothetical benefits or benefits to others, researchers report caring more about personal, immediate benefits like

feeling confident, efficiency, and job performance.

Ideas to address this influence:

Create rewards

Describe reporting items fully

Make reporting guidelines appear as a priority

Provide testimonials

Provide additional teaching

I.30 Researchers may expect the costs to outweigh benefits

Researchers may feel that the costs of using a reporting guideline - the time and work required and the added manuscript length - outweigh the benefits.

Ideas to address this influence:

Keep reporting guidelines agnostic to design choices

Endorse and enforce reporting guidelines

Make information digestible

Keep guidance short

Provide testimonials

I.31 Researchers may feel that checking reporting is someone else's job.

Researchers report feeling that completing a reporting checklist should be the job of the editor or peer reviewer, not the author. Editors and reviewers may also disagree about whose role it is.

(NB. researchers, editors and reviewers could *all* check for reporting quality, but this research focusses only on researchers).

Ideas to address this influence:

Describe reporting guidelines where they are encountered

Use persuasive language and design

Describe each reporting guideline fully

I.32 Researchers may not consider writing as reporting

Researchers may need to change their approach to writing or what they consider writing to be. Researchers differ in their writing process. Authors that follow a structured approach to writing may find it easier to incorporate reporting guidelines into their workflow. Some experienced researchers may be used to a way of working and reluctant to change, and some inexperienced researchers may be unaware of alternative writing processes.

Ideas to address this influence:

Budget for reporting

Provide additional teaching

Appendix J

BCW Step 5 – Identifying intervention functions

Instructions [154]

Use the APEASE criteria to identify appropriate intervention functions based on the behavioural diagnosis arrived at in Step 4

- *Affordability*
- *Practicability*
- *Effectiveness and cost-effectiveness*
- *Acceptability*
- *Side-effects/safety*
- *Equity*

We identified education, training, persuasion, modelling, and structural changes as more practical, acceptable, and affordable. We felt that coercion, restriction, and incentivisation have too many equity and side effect issues, as well as being unacceptable to many stakeholders. We think that it might be a case of right now we need the education, modelling, etc. Once the structural issues around access and knowledge are removed, then we can think about restriction etc..

TABLE J.1: Workshop participants' considerations on the affordability, practicability, effectiveness and cost-effectiveness, acceptability, side-effects/safety, and equitability of intervention functions. Intervention functions and their definitions come from [154].

INTERVENTION		
FUNCTION	EXAMPLES (from step 4)	APEASE
Education	<p>e.g., Providing information to promote healthy eating</p> <p>Telling someone what a reporting guideline is.</p> <p>Telling someone how to use an RG.</p> <p>Telling someone when to use an RG.</p> <p>Seniors prompting juniors to use RGs within collaboration.</p> <p>Show the difference between good and poor reporting. (modelling)</p>	<p>Affordability: needs to be spread to all universities; affordable</p> <p>Practicability: practical, possible, feasible.</p> <p>(Cost)-Effectiveness: potentially cost-effective, because one student well educated may produce lots of good papers; one website can lead to lots of good reports based on RGs. (Needs to be good education, tested, and disseminated. Education that isn't effective isn't cost effective)</p> <p>Equity: adapting education pieces to different publics (for language, for experience with publishing etc.)</p> <p>Side effect: We can't see any risk on being educated! People leaving their research group because they don't use RG?s :-)</p>

INTERVENTION

FUNCTION	EXAMPLES (from step 4)	APEASE
Persuasion	<p>e.g., Using imagery to motivate increases in physical activity</p> <p>Using stories about when good reporting helped someone and when poor reporting caused problems.</p> <p>Using stories about different times when RGs were used and what the results were for publication / collaboration.</p> <p>Seniors prompting juniors to use RGs in collaboration.</p> <p>Using persuasive language/branding in any education or marketing.</p> <p>Using language to make people feel understood and respected, while promoting RGs (emotional response).</p> <p>Using design/layout to make the website more friendly and convey the idea that it is simple/easy to use</p>	<p>Affordability: affordable, because we can publish stories in our website for free.</p> <p>Practicability: practical, although maybe requiring some work hours and different experts. Would look to solicit stories from lots of different people, from different situations.</p> <p>(Cost)-Effectiveness: Can be effective if a structure is used to create stories</p> <p>Acceptability: We expect that these stories can be accepted. However, persuasive language can be seen as patronising (seniors can react poorly to the idea of being told what to do). So we would need to adapt for different audiences.</p> <p>Side-effects/safety: usually safe, but there is a risk that those against a patronising language might resist to use RGs. Small risk of trolling</p> <p>Equity: necessary to tailor the language to different audiences (seniors x ECR; people experienced with the use of RG x not experienced)</p>

INTERVENTION	FUNCTION	EXAMPLES (from step 4)	APEASE
	Incentivisation	<p>e.g., Using prize draws to induce attempts to stop smoking</p> <p>Journals reward good reporting with reduced APCs / fast-track publication.</p> <p>Funders reward good reporting with prioritised funding.</p> <p>Universities use reporting quality for selection committees, promotions, awards.</p> <p>Creating incentives by making products embedding guidance that make writing an easier/quicker task</p> <p>Creating products that give immediate gratification when guidance is followed (gamification).</p> <p>Creating badging systems that well-reported articles can be rewarded with by journals (getting a badge is a shiny incentive).</p> <p>Creating badging systems that well-reported articles can be rewarded with by EQUATOR (getting a badge is a shiny incentive).</p>	<p>Affordability: Ok because already budgeted (no way to calculate cost). Some things very expensive: journals reducing APCs are taking a financial hit. Creating a product with gamification will cost money.</p> <p>Practicability: not difficult to implement. Many of these rely on some measure of ‘good’ reporting. Who is the creator of this standard? Do people agree with it? Getting all universities / funders/ journals to use reporting quality would be very challenging (lots of time to contact so many people and persuade).</p> <p>(Cost)-Effectiveness: cost-effective, will bring results (will get people to do what needs to be done) and these things cost zero or very little. Big variability here in cost effectiveness, depending on costs.</p> <p>Acceptability: OK, no ethics or practical problem in adopting these measures. Likely to be huge push-back by universities / journals / funders, or by the researchers who are then experiencing these things.</p> <p>Side-effects/safety: Gaming, potentially undermining research integrity (reporting things as done when they were not)</p> <p>Equity: problematic because some people might not have access to support for good reporting. (Paraolympics?)</p>

INTERVENTION		
FUNCTION	EXAMPLES (from step 4)	APEASE
Coercion	<p>e.g., Raising the financial cost to reduce excessive alcohol consumption</p> <p>Funders blacklist people who routinely report poorly.</p> <p>Universities refuse promotion for poor reporters.</p> <p>REF does not allow articles that don't follow reporting guidance.</p> <p>Media refuse to speak with authors who are not badged as good reporters.</p> <p>Researchers refuse to collaborate with routine poor reporters.</p> <p>Name and shame poor reporters (opposite of a badge)</p> <p>"If you don't use RGs, peer reviewers will give you a hard time"</p> <p>"If you report poorly, then you'll be viewed as a sloppy researcher and people won't want to work with you"</p>	<p>Affordability: totally, nothing of this has a high cost. Checking by funders etc. would be time consuming and expensive.</p> <p>Practicability: not practical. Difficult to do.</p> <p>(Cost)-Effectiveness: Difficult to evaluate.</p> <p>Acceptability: Low.</p> <p>Side-effects/safety: Bad impact on long-term reputation. How to remove people from blacklists (clean cache!)? How to remove the black list from people? If you keep telling people that bad things will happen and then they never do, then the threats are empty and your other statements become less powerful.</p> <p>Equity: Same as above. You need to support to learn how to report well so that you are not blacklisted. These things will negatively affect ECRs and people without access to education</p>

INTERVENTION		
FUNCTION	EXAMPLES (from step 4)	APEASE
Training	<p>e.g., Advanced driver training to increase safe driving</p> <p>Give people opportunities to practice using RGs.</p> <p>Opportunities to practice finding the right RG for your work.</p> <p>Practice integrating RGs into workflow/ identifying when RGs can be used.</p>	<p>Affordability: we've shown that this can be done in affordable ways, online reaching many people</p> <p>Practicability: We've shown that this is feasible! In person teaching has all the travel/venue/admin practicalities come up. Would need to think about models for scaling: if small group training, train the trainer and ambassador model. If large groups, develop MOOCs (which is time consuming, but has been done before, eg Coursera examples)</p> <p>(Cost)-Effectiveness: Depends on the model, some very cost effective (when reaching many people with little effort, or when reaching fewer people but large impact). As with education, needs to be effective training.</p> <p>Acceptability: Historically, ECRs and students are very receptive to training initiatives. Seniors are less interested for themselves, unless planning to teach their own courses.</p> <p>Side-effects/safety: If dealing with research integrity issues, people can feel judged. When talking about past experiences, people can have difficult feelings come up.</p> <p>Equity: Language (could use existing translators – although not systematic choices of language), time zones (for online courses), location (for physical courses), cost for paid courses (need free versions, freemium models, sponsored models)</p>

INTERVENTION	FUNCTION	EXAMPLES (from step 4)	APEASE
Restriction		<p>Prohibiting/blocking the submission of manuscripts that are not adherent to RGs by introducing editorial staff prechecks.</p> <p>(Journals not publish articles that don't follow RGs.)</p> <p>Seniors/PIs refuse to sign off on manuscripts that don't follow RGs.</p> <p>Funders refuse to fund unless researchers commit to using reporting guidance.</p>	<p>Affordability: cheap for us... Expensive for the ones enforcing (eg journals)</p> <p>Practicability: ERQUATOR doesn't have that much clout/weight/influence!</p> <p>(Cost)-Effectiveness</p> <p>Acceptability: unlikely to be accepted</p> <p>Side-effects/safety: guidance adherence can be subjective, so this could open up folks to being blacklisted for no good reason / underlying bias.</p> <p>Equity: could really negatively affect lower resources folks; what about people not being able to afford to read RGs.</p>

INTERVENTION	FUNCTION	EXAMPLES (from step 4)	APEASE
Environmental restructuring		<p>Instruction to authors pages</p> <ul style="list-style-type: none"> • Changing order of instructions (placing references to reporting guidelines at the top of instructions pages) • Highlighting important info (changing colour, boldening fonts of info about RGs) • Including indications/links to RGs in authors' instructions pages that don't have it already <p>Submission systems</p> <ul style="list-style-type: none"> • Including reminders (in Editorial manager's pages, e.g., "are you sure you've included the relevant RG alongside your manuscript?" prompt messages) <p>Linking different resources together within RG resources – eg, within RG website, or within RG itself, link to the E&E.</p> <p>Make sure all guidance has usable/editable checklists</p> <p>Making sure all guidance is accessible (copyright)</p> <p>Alternative ways of interacting with guidance</p>	<p>Affordability: our time for campaigning, but little cost for whoever owns websites etc (particularly if we supply draft text). Possible cost for things like permissions.</p> <p>Practicability: potentially time heavy, but all doable! Limited by willingness of journals in some things. Making checklists accessible and editable is trickier, as we don't have a lot of control there, but could influence developers during updates and campaign to have journals make changes.</p> <p>(Cost)-Effectiveness: potentially very sensible – small changes that can then be accessed by many people</p> <p>Acceptability: high, little extra work for researchers. RG developers might be more open to us changing the format of guidelines than changing language.</p> <p>Side-effects/safety: journals might get annoyed with us, affect our relationship with them. But no side effects for researchers.</p> <p>Equity: no issues we can think of.</p>

INTERVENTION

FUNCTION	EXAMPLES (from step 4)	APEASE
Modelling	<p>Publishing model papers (with complete reporting) and citing them; or pointing to exemplary papers published (with a badge)</p> <p>Modelling ways to use RGs (case studies / stories showing how others use RGs)</p> <p>Model examples of how to report each item (in E&Es)</p> <p>Share models with medical writers etc, not just for use in education and training</p>	<p>Affordability: cheap, just our time in collating examples. Badging system might require some investment.</p> <p>Practicability: Need to find perfect papers, which can be tricky. Badging system might be tricky, if we want something high tech that journals can apply or that is systematic. But can do very low tech version of EQUATOR just locating a few good examples and pointing to them.</p> <p>(Cost)-Effectiveness: potentially very useful</p> <p>Acceptability: People are always looking for good examples!</p> <p>Side-effects/safety: Need to make sure good examples don't have bad habits within them that might get propagated, or include a disclaimer of what exactly is good.</p> <p>Equity: need to make sure the model papers/examples are open access, language issues</p>

INTERVENTION	FUNCTION	EXAMPLES (from step 4)	APEASE
	Enablement	<p>GoodReports templates (under testing in the GRReAT trial);</p> <p>Space for people to ask experts questions about guidance Eg</p> <ul style="list-style-type: none"> • Medical writing online support (chat) • Ask an RG developer • Forums <p>FAQs on guidance websites, EQUATOR website</p> <p>Making guidance easier to read and follow</p> <ul style="list-style-type: none"> • Replacing difficult language • Using clever formatting, spacing, bullets 	<p>Affordability: all relatively affordable, except where someone is doing live chat – that could be expensive.</p> <p>Practicability: things we're doing or can do, but would need software, outside involvement (experts). All feasible. Changing existing guidance is trickier, as we're not in control of the RGs themselves. Could influence updates. Could also make these changes through eg GoodReports, but would still need consensus from a wider group.</p> <p>(Cost)-Effectiveness: FAQs: small changes for lasting effect. Live chat might not be that cost-effective: only helping one person at a time. Updating guidance: big time investment, potentially big reward</p> <p>Acceptability: RG developers might not like us changing guidance. Might be easier to get buy-in if we have evidence that lots of people struggle with something. Authors would welcome these changes.</p> <p>Side-effects/safety: if questions move beyond experts' field of knowledge, how do they react? Do they give misleading advice?</p> <p>Equity: time zones for Q&As? Language that chats/forums are run in / FAQs are written in. But making guidance easier to understand could increase equity</p>

INTERVENTION**FUNCTION****EXAMPLES (from step 4)****APEASE**

Appendix K

BCW Step 6 – Identifying policy categories

Instructions [154]:

Use the APEASE criteria to identify appropriate policy categories based on the intervention functions identified in Step 5:

- *Affordability*
- *Practicability*
- *Effectiveness and cost-effectiveness*
- *Acceptability*
- *Side-effects/safety*
- *Equity*

Favourite policy routes:

CS: Environmental

JH: Environmental first. Once environmental is better, then can look at guidelines – we're swimming in guidelines now. Service provision after guidelines and comms, due to scalability

JdB: Environmental first; once it is done, then we can share with communication and guidelines and service provision

PL: Environmental first and then guidelines and communication – they need each other to work

SK: Environmental and communication more easily changed by us

Consensus: legislation and fiscal would be more acceptable once environmental, communication, and guidelines are done 'right', not fair here yet

TABLE K.1: Workshop participants' considerations on the affordability, practicability, effectiveness and cost-effectiveness, acceptability, side-effects/safety, and equitability of policy categories. Policy categories and their definitions come from [154].

POLICY		
CATEGORY	EXAMPLE	APEASE
Communication / marketing	Eg: Conducting mass media campaigns Newsletters Adverts/posts on social media Newspaper articles? Blog posts People sharing their good practice Editorials in journals Articles in professional publications Webinars Hotline/chat function?	Affordability: some are and some aren't affordable Practicability: all practical and not difficult to do, except the chat (Cost)-Effectiveness: it will be cost-effective if the right marketing professionals are available to help planning Acceptability: there won't be problems of acceptability Side-effects/safety: overload? (which can be prevented by proper planning) Equity: equity can be good if targeting is adequately planned

POLICY	CATEGORY	EXAMPLE	APEASE
Guidelines		Eg: Producing and disseminating treatment protocols. Recommendations for curricula (from accreditation orgs? To unis?) Research organisations telling their staff what they recommend Funders recommending good reporting practice Guidance for publishers on how to use RGs Updated instructions for authors (from publishers/journals to authors) Guidance for authors Guidance for guideline developers	Affordability: development of sets of rules required work hours; implementation is not a problem for cost (lawyers can cost a lot!) - it depends on who works Practicability: not difficult to write new rules/guidance, especially when using existing frameworks (Cost)-Effectiveness: cost-effectiveness depends on whether the guidance is actually used, which is likely to require some form of enforcement (someone to check). There is also a cost to folks using guidance, as takes time – is that time commitment giving them something? Acceptability: possibly low, if things are rules that researchers HAVE to follow – might feel dictated to. Acceptability possibly dependent on who produces the guidance – some researchers are looking for help and might appreciate Side-effects/safety: researchers exodus, when talking about dictated rules? Equity: independent researchers (who don't get institutional support) will have problems. Language that guidance is available in. Accessibility of guidance.

POLICY		
CATEGORY	EXAMPLE	APEASE
Fiscal measures	<p>Eg: Increasing duty or increasing anti-smuggling activities</p> <p>Changing APCs for good/bad reporting</p> <p>Funders adding incentives (more money for good reporting, less money for poor reporting, funding being conditional on good reporting)</p>	<p>Affordability: Would cost publishers / funders a great deal. But the measures proposed increase the affordability of good reporting for authors.</p> <p>Practicability: difficult because it requires additional checks. Publisher/funder has to make a judgement of good reporting, which might be challenging, particularly when guidance isn't clear.</p> <p>(Cost)-Effectiveness:</p> <p>Acceptability: Publishers might not like having to do more work. Authors already struggle with reporting – can see push back particularly from those who are judged negatively.</p> <p>Side-effects/safety: Bias when publishers/funders judge reporting quality</p> <p>Equity: Issues around who can access and use the guidance, who gets support and education. to report well.</p>

POLICY	CATEGORY	EXAMPLE	APEASE
	Legislation	Eg: Prohibiting sale or use Publishers/journals refusing to publish poorly reported articles Publishers require that their journals must enforce RG use Legislation requiring journals/publishers to use reporting standards (research police) Mandating that funders use reporting quality of protocols as a condition for funding (eg, UKRI could make this a condition of the way MRC etc gives out funding; could be done via wider orgs, such as the way the EU has mandated OA)	Affordability: Publishers' policies: cheap for us, some campaigning needed. Could be less affordable for publishers, if they lose authors or if they need to enforce. 'Research police': expensive: lots of campaigning, development, etc. Practicability: 'Research police': impractical, a lot of work, difficult to do. Publishers: we have existing relationships to lean on, difficult but plausible. A lot of time, money, knowledge for publishers to check RG use – high barrier? If authors find RGs difficult to use, why would publishers be any different? (Cost)-Effectiveness: Unsure, but feeling is that it would be a lot of work for potentially little gain, potentially much gain, depending on whether things worked and whether enforcement was done or not. Acceptability: Researcher pushback? Publisher pushback? Side-effects/safety: Authors might avoid publishers/funders that mandate RGs. Equity: Authors from lower resource areas could be negatively affected – it isn't fair to be penalised if they haven't had education / unable to access guidance (due to copyright).

POLICY

CATEGORY	EXAMPLE	APEASE
Environmental / social planning	Adding Q&As and forums to the EQUATOR website / RG websites Improving the design of the EQUATOR website / RG websites to make them easier to use Changing journal submission systems Posters/reminders in physical environments where researchers are (e.g., universities, hospitals)	Affordability: Some are affordable, and are scalable (can make a few small changes at a time). Journal submission system changes could be expensive for publishers. Practicability: Many of these are practical for us and RG developers. Journal submission systems are trickier though. (Cost)-Effectiveness: small change, big impact for many of these. Reporting clinics could be less cost effective, as reaching fewer people. Acceptability: RG developers likely to be happy with us making these changes or recommendations, as close relationships. Researchers likely to really like these options, as often looking for help and guidance Side-effects/safety: Potentially overwhelming for us Equity: Time zones, languages,

POLICY	CATEGORY	EXAMPLE	APEASE
	Service provision	<p>Training, e.g. Pubschool, in person, MOOC,</p> <p>Interactive chat</p> <p>'Reporting clinics' where people can get help</p> <p>University's own 'reporting guideline help space'</p> <p>Writing centres, publication officers</p> <p>Third-party services, like a reporting quality check from a company</p> <p>Funders offering help to their recipients.</p> <p>Sponsors offering help to their recipients' research</p> <p>Journals offering help to their authors</p> <p>Universities offering help to their members</p>	<p>Affordability: Depends on the service, but many would be less affordable for EQUATOR: lots of time to reach only a small number of people.</p> <p>Could be something that we provide commercially.</p> <p>Could be an affordable way for funders / journals / universities to provide help, meeting their research integrity requirements</p> <p>Practicability: For EQUATOR, a lot of time required. Possibly not practical unless could get specific funding for it. Services offered by funders / journals / universities: could be sensible at scale, but specific support might be lost in the generic support that everyone would need.</p> <p>(Cost)-Effectiveness: Depends on the amount of benefit to providing organisation. If automated, could be a very cost-effective option.</p> <p>Acceptability: Researchers appreciate offered help, but they need to know it exists and that it is helpful. Danger that it could be seen as another red tape thing to get through when designing research. Depends on the format of the help – is it timeous, actionable, how much work does it make for the researchers?</p> <p>Side-effects/safety: Would need to be effective help / to a quality standard to ensure researchers aren't negatively affected by the help. Eventually the only people using it would be junior / new researchers. Would need to be clear that the help doesn't mean a guarantee of publication / don't mislead</p> <p>Equity: If a paid service, then not equitable. If by funder / sponsor, then only people connected to more wealthy funders / sponsors would be</p>

POLICY**CATEGORY****EXAMPLE****APEASE**

Appendix L

SRQR Checklist for chapter 8

TABLE L.1: SRQR Checklist [30]

No.	Item	Description	Location
1	Title	Concise description of the nature and topic of the study. Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended.	Chapter Title

No.	Item	Description	Location
2	Abstract	Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	N/A
3	Problem Formation	Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	Introduction para. 2; Previous chapters
4	Purpose or research question	Purpose of the study and specific objectives or questions	Introduction para. 2; Methods para. 1
5	Qualitative approach and research paradigm	Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/interpretivist) is also recommended; rationale*	Methods para. 1–6; Reflexivity & Trust para. 1

No.	Item	Description	Location
6	Researcher characteristics and reflexivity	Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	Reflexivity and Trust; Reflections on This Chapter; Chapter 2
7	Context	Setting/site and salient contextual factors; rationale*	Methods, Focus Group Sessions para. 1
8	Sampling strategy	How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale*	Methods, Sampling

No.	Item	Description	Location
9	Ethical issues pertaining to human subjects	Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	Methods, Ethics and Data Management
10	Data collection methods	Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale*	Methods, Focus Group Sessions
11	Data collection instruments and technologies	Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	Methods, Materials

No.	Item	Description	Location
12	Units of study	Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	Results, Units of Study
13	Data processing	Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/deidentification of excerpts	Methods, Processing and analysis
14	Data analysis	Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale*	Methods, Processing and analysis

No.	Item	Description	Location
15	Techniques to enhance trustworthiness	Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale*	Methods, Reflexivity and Trust para. 2 and Table 8.1
16	Synthesis and interpretation	Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	Results, Synthesis and Summary
17	Links to empirical data	Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	Results, Synthesis and Summary

No.	Item	Description	Location
18	Integration with prior work, implications, transferability, and contribution(s) to the field	Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/ generalizability; identification of unique contribution(s) to scholarship in a discipline or field	Discussion, Summary of Results & Implications
19	Limitations	Trustworthiness and limitations of findings	Discussion, Limitations
20	Conflicts of interest	Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	See chapter 2
21	Funding	Sources of funding and other support; role of funders in data collection, interpretation, and reporting	See Acknowledgements

Appendix M

Focus Groups - emails

M.0.1 Recruitment Email

Dear [Name],

I would like to introduce myself and invite you to participate in a study that I hope you may find interesting - *Improving the dissemination of reporting guidance using behaviour change theory: focus groups with guideline developers and editors.*

I am a PhD student at the UK EQUATOR Centre, University of Oxford, supervised by Professor Gary Collins, and I'm interested in making reporting guidelines easier to use.

In December 2021 we held a series of internal workshops at the UK EQUATOR Centre where we used the Behaviour Change Wheel to identify possible improvements to how reporting guidance is disseminated. The Behaviour Change Wheel is a method for designing behaviour change interventions, built on top of the COM-B model of behaviour, and is very accessible to non-behavioural experts.

We are now seeking input from the broader community of guideline developers and publishing professionals.

I believe the study's output will be useful to guideline developers, and will include:

- Ideas and justification for future funding applications

- Advice on how to design user-friendly guidance, checklists, and E&E documents
- Ideas to improve guideline websites, the EQUATOR Network website, and a complete redesign of goodreports.org with user feedback channels ‘baked in’ so guideline development groups can collect feedback automatically
- The basis of a subsequent feasibility study and randomised evaluation of some of the identified improvements

If you decide to take part you will be invited to attend a 2-hour, online focus group where we will edit a document together. You will be asked to spend some time before the focus group thinking about reasons that authors may fail or succeed to successfully use reporting guidelines. During the workshop we will discuss these reasons and consider possible ways to improve how reporting guidance is disseminated. You can read the full information for participants here: https://osf.io/esn4j/?view_only=1dcddadfdcbe4479a38d79f8f0f4dd52.

This is an open invitation so please do share it with your colleagues. If you are interested, then please let me know some dates and times that would suit you for a focus group. I am in London, UK and will try my best to accommodate time zones. We can organise a focus group for you and your colleagues, or I can invite you to a focus group with other participants.

Should you be interested but unable to attend a focus group, I would be happy to share the outputs for you to give feedback on at your leisure.

Yours sincerely,

James Harwood

M.0.2 Confirmation Email

Dear [Name],

Thank you for agreeing to join a focus group for the study, *Improving the dissemination of reporting guidance using behaviour change theory: focus groups with guideline developers and editors.*

I'm looking forward to seeing you on [DATE AND TIME OF MEETING]. You can use this link to attend the meeting: [LINK].

Before we meet, please make sure to read through the [participant information sheet](#) and **provide your consent by completing this online form**. If you have time before we meet, please take 10 minutes to **think about facilitators and barriers that may affect whether authors use reporting guidelines successfully and note these down in [THIS SHARED FILE]**. These barriers may come from evidence, theory, or personal experience.

Yours sincerely,

James Harwood

Appendix N

Focus Groups - participant information

James Harwood

DPhil Candidate in Musculoskeletal Sciences

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences

Botnar Research Centre, University of Oxford

University email: james.harwood@stx.ox.ac.uk

Improving the dissemination of reporting guidance using behaviour change theory: focus groups with guideline developers and editors

Participant information sheet

Central University Research Ethics Committee Approval Reference: R80414/RE001

What is the purpose of this study?

Reporting guidance is disseminated through a system that includes the guidelines themselves, websites (e.g., journal instructions, the EQUATOR Network website), tools (e.g., checklists), and other people (e.g., editors, reviewers). The aim of this study is

to identify barriers in any part of this system that may stop authors from successfully using reporting guidelines and possible ways to overcome those barriers.

Who is conducting the study?

Focus groups will be led by the project lead, James Harwood, a DPhil student at the UK EQUATOR Centre supported by a grant from Cancer Research UK supervised by Professor Gary Collins (Professor of Medical Statistics, University of Oxford).

Why are we approaching you?

We are inviting individuals who have insight into how reporting guidance is disseminated. The current dissemination system includes resources created and maintained by different groups: developers of reporting guideline and advocates, publishers, and the EQUATOR Network.

What will happen to me if I take part?

Pre meeting activity: We will ask you to think about facilitators and barriers that may affect whether authors use reporting guidance successfully, and to make note of these in a shared online document. All members of your group will access and edit the same file.

Online focus group: The focus group will last 2 hours. We will begin by discussing the facilitators and barriers that your group members identified before the focus group. We will then compare these with the facilitators and barriers identified by previous participants. You will then be asked to come up with ways to address these facilitators and barriers, before considering the thoughts of previous participants.

You will be asked to collaboratively edit an online document. Each focus group will build upon the output of the previous focus groups. Your writing will be anonymous but will be viewable to other participants.

If you would like more time you will have the option of extending the focus group or arranging a second date.

After the focus group: You will receive a copy of the document produced in your focus group and invited to add comments to it should you have any thoughts after the focus group ends. When all focus groups have taken place, you will be sent a copy of the final document and invited to give feedback on it.

Do I have to take part?

Participation is completely voluntary. If you change your mind about participating, you can withdraw at any point by letting the project lead know (james.harwood@stx.ox.ac.uk) including during the focus group. If you choose to leave we will retain any research data that you have contributed.

How will my data be used?

We will collect your name and email so that we can share the outputs of your focus group and of the study and invite you to comment on them. We will record the audio of the focus group so that the lead researcher can refer to it if needed to avoid recall bias. Your thoughts about barriers and solutions will influence subsequent research projects.

Who will have access to the files co-produced in my focus group?

At the end of your focus group we will share a copy of the file to all participants in that focus group. We will remove all tracked changes, comments, or other identifiable information before sharing it. Once all focus groups have been held, we will share a version of the final, de-identified output.

What happens to the data provided?

The information we acquire during the study is the research data. Any research data from which you can be identified, such as your contact details, consent record and audio recording, is known as personal data. All data will be stored on secure, backed-up servers at the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford and will be held by and accessible to the research team. Audio recordings and contact details will be deleted immediately after the results are written up and the data is no longer needed. Other research data (including consent

records) will be stored for at least 3 years after publication or public release of the work of the research. Responsible members of the University of Oxford may be given access to data for monitoring and/or audit of the research.

We would like your permission to use short excerpts from your focus group, against a pseudonym, in our publications and research presentations. All personal information that could identify you will be removed or changed before doing so.

Will the research be published?

The research may be published in a student thesis, academic publications and on the EQUATOR Network site and shared in academic presentations in venues such as conferences. A copy of my thesis will be deposited both in print and online in the [Oxford University Research Archive](#) where it will be publicly available to facilitate its use in future research.

Who is funding the research?

This research is funded by Cancer Research UK.

Who has reviewed this study?

This study has been reviewed by, and received ethics clearance through, a subcommittee of the University of Oxford Central University Research Ethics Committee (R80414/RE001).

Who do I contact if I have a concern about the study or if I wish to complain?

If you have a concern about any aspect of this study, please contact James Harwood (james.harwood@stx.ox.ac.uk) and I will do my best to answer your query. I will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible. Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB.

Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study.

The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest.

Further information about your rights with respect to your personal data is available from <https://compliance.admin.ox.ac.uk/individual-rights>.

Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

James Harwood

DPhil Candidate in Musculoskeletal Sciences

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences,
Botnar Research Centre, University of Oxford

University email: james.harwood@stx.ox.ac.uk

Appendix O

Ideas generated from workshops and focus groups

The following list of ideas were generated from my workshops with EQUATOR staff and focus groups with guideline developers, journal staff, and academics (chapters 7 and 8). Each bullet point is a code, and codes were grouped inductively into broader ideas that felt cohesive. Each group is labelled with the influence(s) the idea addresses (from Appendix I) and the stakeholders involved.

In chapter 9 I describe how I used this list to define intervention components, and that during this process additional ideas came to me. I've denoted these ideas with ^{JH}.

Before developing a reporting guideline

O.1 Create reporting guidance for early stages of research

Consider creating reporting guidance to help authors write protocols, funding applications, and ethics applications.

Who could do this: Guideline developers

Influences addressed:

Researchers may not have tools for the job at hand

Researchers may not encounter reporting guidelines early enough to act on them

O.2 Avoid confusing authors with too many reporting guidelines

- To avoid duplicating resources, before commencing a new reporting guideline:
 - consult EQUATOR's register of reporting guidelines under development;
 - * EQUATOR could make this register easier to find and search;
 - contact the developers of related reporting guidelines;
 - journals could ask reporting guideline developers to prove that they have registered their guideline with EQUATOR (like they do for clinical trials).
- When a new reporting guideline is justified, build upon existing reporting guidelines instead of starting from scratch. This could mean extending or replacing subsets of items instead of publishing a totally new reporting guideline.
- Consider making modular guidance. For instance, the Journal Article Reporting Standards (JARS) are a set of reporting guidelines for psychology. JARS has a main guideline for quantitative studies. The *Design* item can be extended by one of three modules depending on whether the study involved experimental manipulation or was conducted on a single individual. the experimental manipulation module can be further extended by modules for random assignment, non-random assignment, and clinical trials. Other extension modules exist for longitudinal and replication studies.

Who could do this: Guideline developers, EQUATOR Network, Publishers

Influences addressed:

Researchers may not know what reporting guidelines exist

Researchers may struggle to reconcile multiple sets of guidance

Researchers may not know what reporting guideline is their best fit

When developing a reporting guideline

O.3 Avoid prescribing structure

- Avoid prescribing structure of a journal article as it may clash with journal requirements or other reporting guidelines.
- Instead, give options for where items can be reported.
- Include options beyond the article body where authors can report information, like tables, figures, or appendices be.

Who could do this: Guideline developers

Influences addressed:

Researchers may struggle to reconcile multiple sets of guidance

Researchers may struggle to keep writing concise

O.4 Keep reporting guidelines agnostic to design choices

Ask authors to describe methods transparently without making assumptions about, or prescribing, methods or design choices. For example, an instruction to “describe how you determined your sample size” may be more helpful than “report your sample size calculation” for authors who encounter checklists at submission and did not perform a sample size calculation before collecting data.

Avoid recommending or admonishing design choices within the reporting guidance because:

- doing so may make authors feel nervous or ashamed, and therefore *less* likely to report transparently;
- design advice elongates reporting guidelines;
- including design advice may give the impression that the reporting guideline is for designing or appraising design.

Consider linking to external design or appraisal tools instead.

Who could do this: Guideline developers

Influences addressed:

Researchers may feel restricted if reporting guidelines prescribe design

Researchers may feel afraid to report transparently

Researchers may expect the costs to outweigh benefits

Researchers may not know what reporting guidelines are

O.5 Describe reporting items fully

For each item, authors may need to know the following:

- What needs to be reported – a brief description could go in all resources (checklists, templates etc.) with a longer description in the full guideline document.
- Why the information is important, and to whom
- Any circumstances where the item is not applicable and what to write
- Indicate priority, and any circumstances that modify importance
- Where the item can be reported, including beyond the main article body (e.g., section, table, figure, appendix)
- What to write if an item wasn't, or couldn't be done
- What to write if an item cannot be reported for external reasons. For example, if items cannot be reported because of intellectual property restrictions.
- Examples, which could be real or generated, including:
 - examples of good *and* bad reporting with explanations.
 - examples of concise or word-count-friendly reporting, perhaps in alternative formats like tables and figures.^{JH}
 - examples of well reported “imperfect” items (items that were not done)
 - examples from different research contexts
- Links to external design or appraisal advice

Who could do this: Guideline developers

Influences addressed:

Researchers may not know whether a reporting guideline applies to them

Researchers may not know how to report an item in practice

Researchers may not know how to do an item

Researchers may not know what to write when they cannot report an item

Researchers may struggle to keep writing concise

Researchers may not know why items are important

Researchers may not care about the benefits of using a reporting guideline

Researchers may be asked to remove reporting guideline content

Researchers have limited time

Researchers may struggle to keep writing concise

O.6 Describe each reporting guideline fully

For each reporting guideline, authors may need the following information:

- A clear definition of the reporting guideline's intended scope in plain language.
- If-then rules to direct authors to other, more appropriate reporting guidelines. For example, CONSORT could point authors writing protocols to SPIRIT.
- If no better guidance exists then indicate which items do/do not apply. For example, no guideline exists for authors writing protocols for observational epidemiology. Their best option currently is to use STROBE, but only some items will be required in a protocol.
- What tasks the reporting guideline can and cannot be used for
- How long the resource will take to use
- Why the guidance should be trusted and link to how it was developed

Who could do this: Guideline developers, EQUATOR Network

Influences addressed:

Researchers may not know whether a reporting guideline applies to them

Researchers may not know what reporting guidelines exist

Researchers may not know what reporting guideline is their best fit

Researchers may not know when reporting guidelines should be used

Researchers may feel that checking reporting is someone else's job.

Researchers have limited time

Researchers may feel patronized

O.7 Keep guidance short

Keep guidance as a short as possible:

- Be concise but clear.
- Be realistic about what to expect from authors as each additional item increases the chances an author will be put off
- Link to other guidance elsewhere if desired.
- Consider splitting broad guidance that tries to cater for different options into shorter, modular guidance (modularity avoids duplication).

Who could do this: Guideline developers

Influences addressed:

Researchers have limited time

Researchers may expect the costs to outweigh benefits

When writing guidance down and creating resources

O.8 Make resources ready-to-use

Ensure resources are ready-to-use e.g., checklists as Word files, not as tables within published articles.

Who could do this: Guideline developers, EQUATOR Network

Influences addressed:

[Reporting guideline resources may not be in usable formats](#)

[Researchers have limited time](#)

O.9 Make reporting guidelines easy to understand

- Use plain language.
- Define key terms.
- Use consistent terms across related resources.
- Provide translations.
- Update guidance in response to user feedback.

Who could do this: Guideline developers

Influences addressed:

[Researchers may misunderstand](#)

[Researchers may not understand the language](#)

O.10 Use persuasive language and design

- Use language and design to communicate confidence and simplicity as opposed to judgement and complexity.
- Encourage explanation even when choices were unusual or sub-optimal.

- Reassure authors that most research has limitations that can be addressed in Discussion sections.
- Reassure authors that reporting guidelines are *just* guidelines.
- Avoid patronizing authors.
- Consider wording instructions directly at the intended user.^{JH}

Who could do this: Guideline developers, EQUATOR Network, Publishers, Funders, Ethics committees, Institutions, Conference organisers, Registries, Preprint servers

Influences addressed:

Researchers may feel afraid to report transparently

Researchers may feel patronized

Researchers may not believe stated benefits

Researchers may feel that checking reporting is someone else's job.

O.11 Create additional tools

Create tools for different tasks:

- discussion points for planning research in the order decisions are made
- to-do lists for conducting research in the order data is collected
- templates for drafting
- writing assistance tools (e.g., COBWEB)
- checklists for checking manuscripts that are easy to fill out, update, and cross-check
- tools for co-researchers to check each other's work
- tools for generating tables and figures
- resources for peer reviewers who wish to review reporting quality including:
 - guidance specifically for peer reviewers.
 - commonly-used words that reviewers can search for to quickly find relevant text.^{JH}

- suggested text that peer reviewers can copy to request information
- tools to generate feedback reports
- journal articles where reporting guideline items are annotated/highlighted

Who could do this: Guideline developers, EQUATOR Network, Funders, Ethics committees, Publishers

Influences addressed:

Researchers may not have tools for the job at hand

Researchers may not encounter reporting guidelines early enough to act on them

O.12 Make resources easy to discover and find

Link resources:

- Ensure all resources link to each other. For example, checklists should link to explanation and elaboration documents and vice versa.
- Related reporting guidelines should link to each other.
- Reporting guidelines and resources should link to translations
- Links should be permanent (e.g., DOIs) where possible and old links should be maintained or redirected. Broken links should be replaced.

Make searching easy:

- Host resources somewhere consistent, like the EQUATOR Network website and database.
- Provide easy-to-use website search functions
- Web pages should be optimized for search engines^{JH}
- Create curated collections for study types
- Create decision tools for identifying reporting guidelines

Names reporting guidelines to make them easy to discover and find:

- Reporting guideline names could be descriptive, as acronyms may be meaningless to novice users.
- Related reporting guidelines should use consistent names to show relationships (e.g., PRISMA and PRISMA-P appear more related than CONSORT and SPRIT).

Who could do this: Guideline developers, EQUATOR Network

Influences addressed:

Researchers may not know what reporting guidelines exist

Researchers may not know what resources exist for a reporting guideline

Guidance may be difficult to find

Researchers may not know what reporting guideline is their best fit

O.13 Make information digestible

Organise information so it is easy to navigate and not overwhelming.

- Cater to users that read from start to finish, and those that dip in and out.
- Structure text with headings.
- Use section URLs to send authors directly to relevant parts of guidance.
- Consider hyperlinking related resources
- Consider embedding reporting guidelines that “fit together”, like PRISMA and PRISMA-Abstracts
- For information presented online, consider showing/hiding information as required. For example, if PRISMA-Abstracts were embedded into PRISMA, users could choose to expand or collapse it. Or you could show/hide guidance depending on whether the author is writing a funding application, protocol, manuscript.

Who could do this: Guideline developers, EQUATOR Network

Influences addressed:

Researchers have limited time

Researchers may expect the costs to outweigh benefits

Guidance may be difficult to find

When disseminating resources

O.14 Describe reporting guidelines where they are encountered

- When authors first encounter reporting guidelines they may need to know:
 - what reporting guidelines are
 - how and when to use them
 - why authors should use them, including:
 - * what personal benefits to expect
 - * the importance to others.
- Descriptions could be succinct (e.g., on journal instruction pages) or long (e.g., in publications)^{JH}
- A generalised description can go where authors first encounter reporting guidelines e.g., journal author guidelines, EQUATOR's home page.
- A reporting guideline-specific description could go at the top of guidance documents, checklists, and templates.
 - Consider specifying whether the reporting guideline is also a design guideline.
 - Specify whether the reporting guidelines are *just* guidelines, or whether they are intended to be requirements. Name the resource appropriately - words like *guideline*, *standards*, *criteria*, *recommended*, *preferred*, and *templates*, have different meanings.

Who could do this: Publishers, EQUATOR Network, Guideline developers, Funders, Ethics committees, Institutions, Registries, Preprint servers, Conference organisers

Influences addressed:

Researchers may not know what reporting guidelines are

Researchers may not know when reporting guidelines should be used

Researchers may not know what benefits to expect

Researchers may not know why items are important

Researchers may feel that checking reporting is someone else's job.

O.15 Make resources accessible

Ensure resources are open access. This allows access to authors without journal subscriptions and allows others to build upon the guidance.

Who could do this: Guideline developers, EQUATOR Network

Influences addressed:

Reporting guidelines may be difficult to access

O.16 Show and encourage citations

- Display usage data (like citations or downloads) alongside the guidelines as a form of social proof.
- Encourage authors to cite the reporting guideline so readers discover it.

Who could do this: Guideline developers, EQUATOR Network, Publishers, Conference organisers, Preprint servers

Influences addressed:

Researchers may not know what reporting guidelines exist

Researchers may not believe stated benefits

O.17 Provide testimonials

Testimonials can be short quotes or longer case studies. They could come from:

- researchers who have had positive experiences using reporting guidelines, including researchers that were nervous about transparency,
- decision makers (e.g., editors/grant managers) that value good reporting and/or check for reporting as part of their evaluation,
- peer reviewers that use reporting guidelines to check for good reporting,
- patients who are affected by research waste,
- and researchers who need to understand, synthesise, or apply research articles.

Who could do this: Guideline developers, EQUATOR Network

Influences addressed:

Researchers may not know what benefits to expect

Researchers may not believe stated benefits

Researchers may expect the costs to outweigh benefits

Researchers may not care about the benefits of using a reporting guideline

Researchers may feel afraid to report transparently

On an ongoing basis

O.18 Budget for reporting

Funders and research supervisors could encourage researchers to allocate sufficient time and money for documenting and reporting results of their research.

Who could do this: Funders, Institutions

Influences addressed:

Researchers have limited time

Researchers may not consider writing as reporting

O.19 Create rewards

Stakeholders could create new rewards:

- journals could fast-track submissions or review for papers that followed a reporting guideline,
- journals could offer discounts on article processing charges for papers that followed a reporting guideline,
- journals, preprint servers, or peer review platforms could badge well reported articles,
- EQUATOR could offer a certification service,
- funders could reward good reporting financially,
- institutions could offer prizes for good reporting.

Who could do this: Guideline developers, EQUATOR Network, Publishers, Funders, Institutions, Preprint servers, Registries

Influences addressed:

Researchers may not care about the benefits of using a reporting guideline

O.20 Create discussion spaces

Create spaces for authors to discuss reporting and reporting guidelines. These could be:

- online (forums, social media, email),
- or offline (meet-ups, clubs).

Try to make spaces accessible to researchers from all nationalities, professional disciplines and other demographics. Spaces will allow authors to:

- solicit help,
- share experiences,

- provide feedback to guideline developers,
- and cultivate a feeling of inclusivity and community ownership.

Who could do this: EQUATOR Network, Guideline developers, Institutions

Influences addressed:

Researchers may misunderstand

Researchers may feel patronized

Researchers may not believe stated benefits

Researchers may not know how to report an item in practice

Researchers may not know how to do an item

O.21 Create ways to catch authors earlier

- Consider creating email campaigns to prompt researchers at early stages.
- The EQUATOR website could encourage visitors to use reporting guidelines for planning and drafting research.
- Websites could be optimised for search terms like “how to write [study type]”, “protocol”, “research plan” or “funding application”. For example, reporting guideline pages on EQUATOR’s website rank highly in Google searches for “STROBE checklist” but not “How to write an observational epidemiology study”.^{JH}
- Writing clubs and writing training could flag reporting guidelines.

Who could do this: EQUATOR Network, Guideline developers, Publishers, Funders, Ethics committees, Institutions, Conference organisers, Preprint servers, Registries

Influences addressed:

Researchers may forget to use reporting guidelines at earlier research stages

Researchers may not encounter reporting guidelines early enough to act on them

Researchers have limited time

Researchers may not know when reporting guidelines should be used

O.22 Endorse and enforce reporting guidelines

Stakeholders could:

- endorse reporting guidelines
- enforce their use by mandating checklists or (preferably) checking adherence to items.
- Funders could ask about reporting guidelines or checklists when collecting updates from grant recipients.

Who could do this: Publishers, Institutions, Ethics committees, Funders, Registries, Conference organisers, Preprint servers

Influences addressed:

Researchers may not know what reporting guidelines exist

Researchers may expect the costs to outweigh benefits

O.23 Evidence the benefits

Evidence any stated benefits:

- Quantifiable benefits could be evidenced with data (e.g., acceptance rates, publishing speed, writing speed).
- Experiential benefits could be evidenced by collecting case studies from authors who find that reporting guidelines help them feel confident and write more easily, and from readers who value well-reported research.

Who could do this: Guideline developers, EQUATOR Network, Publishers

Influences addressed:

Researchers may not believe stated benefits

O.24 Make reporting guidelines appear as a priority

Journals, funders and ethics committees could make reporting guidelines appear as a priority:

- Make them prominent in author instructions.
- Placing checklists earlier in the PDFs that are automatically created by journal submission systems.
- Publicize when reporting guidelines are used by reviewers.

Who could do this: Publishers, Funders, Ethics committees, Institutions, Preprint servers, Conference organisers, Registries

Influences addressed:

Researchers may not believe stated benefits

Researchers may not care about the benefits of using a reporting guideline

O.25 Promote reporting guidelines

- Promote reporting guidelines on and offline.
 - Online may include websites, email campaigns, social media, and blogs.
 - Offline may include appearing at conferences, seminars, and workshops.
- Institutions could promote reporting guidelines in their curricula, learning materials, or through [reporting champions](#). Reporting guideline developers or EQUATOR could push for reporting guidelines to be included in text books.
- Promotion can begin before a reporting guideline has been published so that researchers know about guidelines being developed.

NB. Promotion is different to endorsement; a journal could run an email campaign to promote reporting guidelines without having an endorsement policy.

Who could do this: Institutions, Publishers, Guideline developers, EQUATOR Network, Ethics committees, Funders, Societies, Registries, Conference organisers, Preprint servers

Influences addressed:

Researchers may not know what reporting guidelines are

Researchers may not know what reporting guidelines exist

O.26 Install reporting champions

All stakeholders could have members to promote and facilitate the usage of reporting guidelines.

- This could follow a local network model with EQUATOR as the central organiser.
- Could make use of existing networks, like regional reproducibility networks.

Who could do this: EQUATOR Network, Guideline developers, Institutions, Funders, Ethics committees, Publishers, Conference organisers, Preprint servers, Registries

Influences addressed:

Researchers may not know what reporting guidelines are

Researchers may not know what benefits to expect

Researchers may misunderstand

Researchers may not know when reporting guidelines should be used

Researchers may not know why items are important

O.27 Provide additional teaching

Provide education or training (e.g., courses, videos) specific to particular reporting guidelines.

More generally, students could:

- learn about writing as a process and workflows for documenting and communicating research,

- learn about research waste from poor reporting,^{JH}
- attempt a replication to learn about the importance of complete reporting,
- and use a reporting guideline as part of their studies.

Who could do this: Guideline developers, EQUATOR Network, Publishers, Institutions, Funders, Ethics committees

Influences addressed:

Researchers may not consider writing as reporting

Researchers may misunderstand

Researchers may not know why items are important

Researchers may not know how to report an item in practice

Researchers may not encounter reporting guidelines early enough to act on them

Researchers may not care about the benefits of using a reporting guideline

O.28 Make updating guidelines easier

Update guidance in response to user feedback or changes in the field. This would be easier if:

- reporting guideline developers could easily collect feedback from authors.
- small updates or refinements could be made without publishing a new article.
- reporting guideline developers had funding to evaluate, refine, and update their resources.^{JH}

Who could do this: EQUATOR Network, Funders

Influences addressed:

Reporting guidelines can become outdated

Researchers may misunderstand

Appendix P

Intervention Components Table

In chapter 9 I described how I combined the outputs of previous chapters to create the intervention components listed below.

TABLE P.1: Intervention Components Table. Intervention components labelled with the behaviour change techniques and intervention functions they employ, and grouped according to the key behaviours, barriers, and behavioural drivers that they aim to target. Where possible, examples demonstrate how components were (or were not) used originally (Before) and, how they are included within the redesigned intervention (Now).

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Key Behaviour: Engage with (read) appropriate reporting guidance as early as possible				
Targeted barrier: Researchers may not know what reporting guidelines are				
Behavioural driver: Capability				
Describe what reporting guidelines are where they are first encountered	Instruction on how to perform the behaviour	Education	No prominent description of what reporting guidelines are on EQUATOR home page or in reporting guidelines resources.	Prominent definition on home page and guideline page. Example: See Figure 9.4, Figure 9.7 Example: See Figure 9.3, Figure 9.6

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for	Instruction on how to perform the behaviour	Education	No clear instruction on what tasks reporting guidelines or their resources can and cannot be used for. Example: See Figure 9.3, Figure 9.5, Figure 9.6	Clear instruction and differentiation of resources Example: See Figure 9.4, Figure 9.7
Targeted barrier: Researchers may not know what reporting guidelines exist				
Behavioural driver: Capability				
Instruct authors to cite reporting guidelines so readers may learn about them	Instruction on how to perform the behaviour	Education	No consistent instruction to cite reporting guidelines	Consistent instruction to cite reporting guidelines
Decision tools for discovering appropriate resources*	Instruction on how to perform the behaviour	Enablement	We previously made a “reporting guideline wizard” but it was difficult to find. Example: see Figure 9.3	Not included yet

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Collections of related reporting guidelines*	Adding objects to the environment	Environmental Restructuring	Collections exist on EQUATOR site but are difficult to find.	Not included yet
Links between related guidelines	Restructuring the physical environment	Environmental Restructuring	Guideline publications may cite guidelines published previously, but these can be buried in text and are not updated. EQUATOR website guideline pages feature links to extensions, but these may be hard to find. Checklists do not link to related resources. Example: See Figure 9.5, Figure 9.6	Guidelines prominently link to other relevant guidelines and explain when they should be used. Example: See Figure 9.7

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Embed reporting guidelines that “fit together”*	Instruction on how to perform the behaviour	Enablement	Checklists and their extensions are published separately. The best example of modular guidance is perhaps the JARS guidelines, but even these are published as separate documents.	No change
Targeted barrier: Guidance may be difficult to find				
Behavioural driver:				
Opportunity				
Centralised hosting	Restructuring the physical environment	Enablement	EQUATOR maintains a database of reporting guideline meta-data, but the guidance and checklists were published and hosted in different locations and in different ways.	A core set of frequently accessed guidelines are now presented on a single website.

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Search function on website	Restructuring the physical environment	Enablement	EQUATOR's search function was difficult to find. Example: See Figure 9.3	Search function is easier to find as a recognizable icon in the navigation bar of every page. The home page includes additional ways to access search functionality. Example: See Figure 9.4

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Search Engine Optimization	Restructuring the physical environment	Enablement	EQUATOR's website did not make use of some commonly used search optimization heuristics. It ranked well for guideline acronyms (like STROBE) but not for general terms that naive authors may use, like "observational epidemiology" or "how to write-up research". The site wasn't optimized for viewing on mobile devices, which will also harm google search rankings. Example: (Not visible)	The site has additional meta-data. Each reporting guideline page has its own meta-data. The site is optimized for mobiles. Example: (Not visible)

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Permanent document object identifiers (DOIs)*	Restructuring the physical environment	Enablement	Although guideline publications have DOIs, tools (commonly hosted on guideline developer's websites) do not. EQUATOR's website does not use document object identifiers. If resources move (e.g., a website is reorganised or deprecated) then links can "die".	No change
Targeted barrier: Reporting guidelines may be difficult to access				
Behavioural driver:				
Opportunity				
Ensure guidelines and tools are open access*	Restructuring the physical environment	Enablement	Some guidelines are published behind paywalls	No change

INTERVENTION INGREDIENT	BCT	INTERVENTION		NOW
		FUNCTION	BEFORE	
Targeted barrier: Researchers may not know whether a reporting guideline applies to them				
Behavioural driver: Capability				

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Describe the scope of a reporting guideline at the top of every resource	Instruction on how to perform the behaviour	Education	Some reporting guidelines may describe their scope within a publication. Others might not, or may only describe their scope broadly without fully explaining the design assumptions within the guidance. Guideline publications rarely explain circumstances where the guidance should not be used. Checklists rarely define intended scope beyond the title of the guideline.	The intended scope of a guideline is clearly & prominently described. This definition includes contexts in which the guidance should not be used. Example: See Figure 9.7

Targeted barrier: Researchers
may not know what reporting
guideline is their best fit

Behavioural driver: Capability

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Use if-then rules to direct authors to more appropriate and up-to-date guidance when available	Instruction on how to perform the behaviour	Education	Reporting guidelines do not consistently point authors towards related resources that might be better fits. Guidelines are not updated as-and-when other guidelines become available.	Reporting guidelines clearly and consistently point authors to more appropriate guidance when appropriate, using if-then rules. These links can be updated any time. Example: See Figure 9.7
Explicitly state when no better guidance exists for a use case	Instruction on how to perform the behaviour	Education	Reporting guidelines rarely explain what to do when no better guidance exists for a use case.	Reporting guidelines warn authors when no better guidance exists for a use case, and how the current guidance can be adapted instead Example: See Figure 9.7

INTERVENTION INGREDIENT	BCT	INTERVENTION		NOW
		FUNCTION	BEFORE	
Targeted barrier: Researchers may not understand the language				
Behavioural driver: Opportunity				

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Provide translations	Instruction on how to perform the behaviour	Enablement	<p>Some guidelines have been translated, but many haven't.</p> <p>Links to translations are present on reporting guideline database pages but these links may not be easy to find. The EQUATOR website has an automatic translation tool which will translate content on web pages, but this doesn't cover the guidance itself.</p> <p>Example: SRQR was available in French but the translation wasn't advertised on the main guidance or database page. See Figure 9.6 and Figure 9.5</p>	<p>Translations are prominently listed above the guidance</p> <p>Example: See Figure 9.7</p>

INTERVENTION INGREDIENT	BCT	INTERVENTION			
		FUNCTION	BEFORE	NOW	
Targeted barrier: Researchers may expect the costs to outweigh benefits					
Behavioural driver: Motivation					
Make guidance appear shorter by removing superfluous information, hiding optional content, splitting long guidelines, using concise language, and separating design advice	Restructuring the physical environment	Environmental restructuring	reporting guideline publications may include lengthy explanation of development, verbose language, and can be bloated by design advice Example: See Figure 9.8	SRQR has been edited. The only text presented immediately is instruction on what the author needs to describe. Additional information is hidden at first and can be expanded. Text is shortened through editing and by using active voice. In the case of SRQR, this reduced the text length by 60%. Example: See Figure 9.9	

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Cater to different kinds of user (readers vs dippers) by structuring guidance with headings, itemisation, hyperlinking to particular sections, and with optional content	Restructuring the physical environment	Environmental restructuring	Besides being split into items, reporting guidance is largely unstructured and different items can be organised in different ways. Checklist items do not link to items within an elaboration document.	SRQR items are structured consistently, making information easier to find. Itemisation is used consistently, content is hyperlinked when useful. Example: See Figure 9.9
Include testimonials from researchers who were nervous about being punished for reporting transparently	Demonstration of the behaviour	Persuasion	No such testimonials exist	Quotes included alongside guideline Example: see Figure 9.9

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Decrease fear of judgement by making reporting guidelines design agnostic	Remove aversive stimulus	Coercion (Removal of)	Reporting guidelines may conflate reporting advice with design advice or design assumptions. The justification for why an item is important to describe is frequently presented in terms of good and bad design.	SRQR explicitly states that it makes no assumptions about design. Inadvertent design assumptions were edited.

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Remove branding and messaging that may invoke feelings of judgement, complexity, or administrative red-tape	Remove aversive stimulus	Coercion (Removal of)	EQUATOR's website looked cluttered and visually unappealing. Guidance published in articles can look unappealing and dense. When justifying why authors should use reporting guidelines, guideline developers frequently referenced research waste, (lack of) transparency, bias, and poor design. Example: See Figure 9.3	A clean, simple interface for the home page and guidance pages. Text makes less use of judgemental phrases and fewer references to the negative consequences of poor reporting. Example: See Figure 9.4
Reassure that all research has limitations to encourage explanation over perfect design	Social support (unspecified)	Persuasion	Few guidelines would include this kind of reassurance Example: See Figure 9.6	This reassurance appears on the home page and all guidance pages Example: See Figure 9.7

INTERVENTION INGREDIENT	BCT	INTERVENTION			
		FUNCTION	BEFORE	NOW	
Targeted barrier: Researchers may feel that checking reporting is someone else's job.					
Behavioural driver: Motivation					
Address communications to authors	Instruction on how to perform a behaviour	Persuasion	It wasn't always clear whether reporting guidelines and the EQUATOR Network website were aimed at authors, editors, reviewers, or all.	All resources and website copy are directed predominantly at authors.	
Communicate why reporting is primarily the responsibility of the author*	Instruction on how to perform a behaviour	Education	Because it wasn't clear how reporting guidelines and checklists should be used, they (especially checklists) could appear as administrative tasks that should be the responsibility of the editor or reviewer.	Clear explanation of why guidelines and tools should be used by authors primarily, although can also be used by others.	

INTERVENTION INGREDIENT	BCT	INTERVENTION			NOW
		FUNCTION	BEFORE		
Targeted barrier: Researchers may not consider writing as reporting					
Behavioural driver: Motivation					
Educate authors about writing as a process	Instruction on how to perform a behaviour	Education	Many researchers don't get trained on writing as a process, they just do it. EQUATOR provided education about how to write but this wasn't advertised on guidelines.	Some SRQR items now link to relevant EQUATOR materials and courses. Example: See Figure 9.9	
Key Behaviour: Apply reporting guidance to writing					
Targeted barrier: Researchers may not know what resources exist for a reporting guideline					
Behavioural driver: Capability					

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Link all resources to each other	Restructuring the physical environment	Environmental restructuring	Reporting guideline development articles, explanation and elaboration articles, and checklist files may not link to each other.	Guidance links to all tools and development article Example: See Figure 9.7
Targeted barrier: Researchers may not know what benefits to expect				
Behavioural driver: Capability				

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications)	Information about emotional consequences, Information about others' approval, Information about social and environmental consequences	Education	Benefits are not prominently described on EQUATOR's home page, nor within the guideline publications. Benefits that are described may be hard to find, and often focus on hypothetical benefits to the research community, but not personal benefits to the author.	Benefits are prominently and consistently displayed across the home page and guidance pages. Descriptions prioritize personal benefits to the authors above hypothetical benefits to others. Example: See Figure 9.3

Targeted barrier: Researchers may not believe stated benefits
Behavioural driver: Motivation

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Gather and communicate evidence for benefits	Information about emotional consequences, Information about others' approval, Information about social and environmental consequences	Persuasion	Benefits often presented without evidence (if at all)	Dummy quotes provides evidence for experienced benefits. Example: See Figure 9.9

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Include design, features, and language to foster trust	Credible source, Social comparison	Persuasion	<p>Website design looked amateur.</p> <p>Citation metrics are available for guideline publications, but are not displayed on the EQUATOR website or within the guidance publications or checklists themselves. Guidelines were often preceded by lengthy explanations of development.</p> <p>Example: See Figure 9.6</p>	<p>Professional design. EQUATOR's logo remains prominent.</p> <p>Citation metrics are presented at the top of the reporting guidance. Information about who developed the guidelines, how they developed it, and why the guidance is credible is still provided, and easily findable from the top of the guidance.</p> <p>Example: See Figure 9.7</p>
Create spaces for authors to discuss reporting guidelines with others	Social comparison, Credible source, Adding objects to the environment	Persuasion	<p>There were no official on- or offline spaces for authors to discuss guidelines.</p>	<p>Each reporting item has its own discussion board.</p> <p>Example: See Figure 9.10</p>

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Use tone of voice and design to communicate personal benefits; confidence and simplicity	Framing/reframing Persuasion		Guidance text made little use of a reassuring tone or words. The EQUATOR website and guideline articles looked dense and complex. Example: See Figure 9.3 and Figure 9.6	A clean, simple interface for the home page and guidance pages. Text uses phrases like “confidence”, “quick”, “maximum impact”. Example: See Figure 9.4 and Figure 9.7
Targeted barrier: Researchers may not care about the benefits of using a reporting guideline				
Behavioural driver: Motivation				
Include testimonials from research users who benefit from complete reporting	Salience of consequence	Persuasion	Testimonials not included in reporting guidelines.	SRQR includes dummy testimonials and quotes from research users Example: See Figure 9.9

INTERVENTION INGREDIENT	BCT	INTERVENTION			
		FUNCTION	BEFORE	NOW	
Targeted barrier: Researchers may misunderstand					
Behavioural driver: Capability					
Use plain language	Instruction on how to perform the behaviour	Enablement	Although developers aspire to write clearly, authors may misinterpret guidance or fail to understand it completely.	SRQR is edited to use plainer language.	
Define key terms	Instruction on how to perform the behaviour	Education	Few guidelines came with a glossary. Some key terms may be defined within the guideline text. Including definitions this way makes them hard to find and elongates the guidance.	SRQR now has a glossary, and text is marked-up with definitions that appear upon click. Example: See Figure 9.9	

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Use consistent terms	Instruction on how to perform the behaviour	Enablement	Guidelines may use different terms to refer to the same thing (or the converse - use the same term to refer to different things). A single guideline may do this too, as can websites.	The website uses terms consistently
Provide translations	None	See above		
Create spaces for authors to discuss reporting guidelines with others (see above)	None			
Targeted barrier: Researchers may not know why items are important				
Behavioural driver: Capability				

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
For each item, explain why the <i>information</i> is important and to whom (not just what constitutes “good” design)	Information about social and environmental consequences	Education	Sometimes there was no explanation as to why an item should be reported. Other times the justification would be about why a particular design choice was important. Example: See Figure 9.8	Information added when necessary Example: See Figure 9.9
Explain importance of complete reporting to the scientific community	Information about social and environmental consequences	Education	EQUATOR and most reporting guidelines do this already	Continue to do this

Targeted barrier: Researchers

may not know how to do an item

Behavioural driver: Capability

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Provide links to other resources that explain how an item can be done	Instruction on how to perform the behaviour	Education	Some reporting guideline publications (or elaboration articles) include instruction in text but many don't. SRQR did not.	Links included when relevant. Example: See Figure 9.9
Targeted barrier: Researchers may not know how to report an item in practice				
Behavioural driver: Capability				
For each item, provide clear instruction of what needs to be described	Instruction on how to perform the behaviour	Education	Writing instructions are often mixed in with other explanation and context. Example: See Figure 9.8	Writing instruction occurs first for each item. Example: See Figure 9.9

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
For each item, provide examples of reporting in different contexts	Demonstration of the behaviour	Modelling	Not all reporting guidelines provide examples. Examples may not cover different contexts.	SRQR already had some examples. No more examples added Example: See Figure 9.9
Create spaces for authors to discuss reporting guidelines with others (see above)		None		Example: See Figure 9.8
Targeted barrier: Researchers may not know what to write when they cannot report an item				
Behavioural driver: Capability				
Provide clear instruction of what needs to be described when an item was not done, could not be done, or does not apply	Instruction on how to perform the behaviour	Education	Rarely instructed Example: See Figure 9.8	Instructed where relevant Example: See Figure 9.9

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Provide examples of reporting “imperfect” items*	Demonstration of the behaviour	Modelling	Examples not provided	No changes made
Targeted barrier: Researchers				
have limited time				
Behavioural driver:				
Opportunity				
Ensure all resources and tools (e.g., checklists and templates) are in ready-to-use formats*	Adding objects to the environment	Enablement	Some checklists not in immediately usable formats e.g., PDFs	No changes made
Structure guideline items to make them quicker to digest	Restructuring the physical environment	Enablement	E&E documents not structured below the item level Example: See Figure 9.8	Items have consistent structure and use bullet points consistently Example: See Figure 9.9
Tell authors how long the guidance will take to read	Instruction on how to perform the behaviour	Education	Estimated reading time not given	Estimated reading time given Example: See Figure 9.7

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Tell authors how long guidance will take to apply*	Instruction on how to perform the behaviour	Education	No advice given	No changes made Example: No changes made
Targeted barrier: Researchers may not know when reporting guidelines should be used				
Behavioural driver: Capability				
Tell authors when to use reporting guidelines, or that reporting guidelines are best used as early as possible	Instruction on how to perform the behaviour	Education	Rarely stated prominently	Stated prominently Example: See Figure 9.7
Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for (see above)		None		

INTERVENTION INGREDIENT	BCT	INTERVENTION		
		FUNCTION	BEFORE	NOW
Targeted barrier: Researchers may not encounter reporting guidelines early enough to act on them				
Behavioural driver:				
Opportunity				
Optimize websites for search terms aimed at early use like “how to write”, or “funding application”. (See <i>Search Engine Optimization</i> above)		None		
Create prompts / communication campaigns to target authors early in their research*	Prompts/cues	Enablement	EQUATOR had no way to do this	No changes made

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Create tools to be used for early writing tasks*	Adding objects to the environment	Enablement	Most reporting guidelines come with a checklist but none come with a template, or tools/guidance specific to protocols or funding applications.	No changes made
Targeted barrier: Researchers may struggle to keep writing concise				
Behavioural driver:				
Opportunity				
Provide instruction as to how and where information can be reported without breaching word count limits or making articles bloated.	Instruction on how to perform the behaviour	Education	Few reporting guidelines include this information	Added instruction at top of reporting guideline and in some items where most useful Example: See Figure 9.7

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Provide examples of concise reporting*	Demonstration of the behaviour	Modelling	No examples specifically to display concise reporting	No changes made
Targeted barrier: Researchers may not have tools for the job at hand				
Behavioural driver: Opportunity				
Create guidance for planning research, or for writing protocols/funding applications (see <i>Create tools to be used for early writing tasks</i>)*				

		INTERVENTION		
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Create to-do lists in the order research is conducted, to help authors collect information they will need to report (see <i>Create tools to be used for early writing tasks</i>)*				
Create templates for drafting (see <i>Create tools to be used for early writing tasks</i>)*				
Create tools to facilitate checklist completion*	Adding objects to the environment	Enablement	Updating page numbers in a checklist is time consuming. It takes editors time to double check page numbers and content. Checklists may not include instructions of how to complete them.	No changes made

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Create tools to facilitate particular reporting items*	Adding objects to the environment	Enablement	Some tools exist (e.g., PRISMA flow chart diagram maker, COBWEB)	No changes made
Create tools to help collaborators check each other's work*	Adding objects to the environment	Enablement	Checklists exist but aren't specifically designed for collaborators	No changes made
Create tools to help peer reviewers check reporting and request missing information*	Adding objects to the environment	Enablement	Checklists are reporting guidelines are not specifically aimed at peer reviewers	No changes made
Targeted barrier: Reporting guidelines can become outdated				
Behavioural driver:				
Opportunity				

INTERVENTION INGREDIENT	BCT	INTERVENTION		
		FUNCTION	BEFORE	NOW
Provide feedback channels to help developers keep guidance updated (see <i>Create spaces for authors to discuss reporting guidelines with others</i>)		None		
Targeted barrier: Researchers may struggle to reconcile multiple sets of guidance				
Behavioural driver:				
Opportunity				
Explain when reporting guidelines do not intended to prescribe structure	Instruction on how to perform the behaviour	Education	Not always stated. Not always prominent	Explained at top of guidance Example: see Figure 9.7
			Example: see Figure 9.6	

INTERVENTION INGREDIENT	BCT	INTERVENTION		
		FUNCTION	BEFORE	NOW
Provide instruction as to how and where information can be reported without breaching word count limits or making articles bloated (see above)		None		
Embed reporting guidelines that “fit together” (see above)*		None		
Targeted barrier: Researchers may be asked to remove reporting guideline content				
Behavioural driver:				
Opportunity				
Provide advice regarding how to respond if asked to remove reporting guideline content by a colleague, editor, or reviewer	Problem solving	Education	No advice given	Advice given in FAQ Example: See Figure 9.7

INTERVENTION INGREDIENT	BCT	INTERVENTION		
		FUNCTION	BEFORE	NOW
Targeted barrier: Reporting guideline resources may not be in usable formats				
Behavioural driver:				
Opportunity				
Ensure all resources and tools (e.g., checklists and templates) are in ready-to-use formats (see above)*		None		
Targeted barrier: Researchers may feel afraid to report transparently				
Behavioural driver: Motivation				
Present design advice separately to reporting advice	Restructuring the physical environment	Coercion (removal of)	Some reporting guideline E&Es include design advice Example: SRQR did not include design advice	No changes made

INTERVENTION INGREDIENT	BCT	INTERVENTION			NOW
		FUNCTION	BEFORE		
Make reporting guidelines agnostic to design choices (see <i>Decrease fear of judgement by making reporting guidelines design agnostic</i>)		None			
Encourage explanation even when choices are unusual or not optimal	Instruction on how to perform the behaviour	Education	Not always present Example: See Figure 9.8	Added to items Example: See Figure 9.9	
Reassure authors that all research has limitations (see <i>Reassure that all research has limitations to encourage explanation over perfect design</i>)		None			

INTERVENTION INGREDIENT	BCT	INTERVENTION		
		FUNCTION	BEFORE	NOW
Include testimonials from researchers who were nervous about being punished for reporting transparently (see <i>Include testimonials from researchers who were nervous about being punished for reporting transparently</i>)		None		
Targeted barrier: Researchers may feel restricted if reporting guidelines prescribe design				
Behavioural driver: Motivation				
Present design advice separately and remain design agnostic (see <i>Decrease fear of judgement by making reporting guidelines design agnostic</i>)		None		

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Reassure when guidelines are <i>just</i> guidelines	Social support	Persuasion	Not always present or prominent Example: See Figure 9.6	Prominently displayed at top of reporting guideline Example: See Figure 9.7
Targeted barrier: Researchers may feel patronized				
Behavioural driver: Motivation				
Create spaces for authors to discuss reporting guidelines with others (see above)		None		
Avoid patronizing language	Remove aversive stimulus	Persuasion	Although authors may feel patronized when asked to adhere to a reporting guideline, reporting guidelines themselves rarely use patronizing language	Continue to avoid using patronizing language

INTERVENTION				
INTERVENTION INGREDIENT	BCT	FUNCTION	BEFORE	NOW
Explain how the guidance was developed and why it can be trusted	Credible source	Education	Most reporting guidelines explain this in a published article. Checklists do not Example: See Figure 9.6	Brief description included on home page and at top of reporting guideline, links to full to development information Example: See Figure 9.7
Key Behaviour: Repeat engagement with reporting guidelines for subsequent studies				
Targeted barrier: Researchers may forget to use reporting guidelines at earlier research stages				
Behavioural driver: Opportunity				

INTERVENTION INGREDIENT	BCT	INTERVENTION		
		FUNCTION	BEFORE	NOW
Create prompts / communication campaigns to target authors early in their research (see above)*		None		

Appendix Q

SRQR Checklist for chapter 10

TABLE Q.1: SRQR Checklist [30]

No.	Item	Description	Location
1	Title	Concise description of the nature and topic of the study. Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended.	Title of chapter

No.	Item	Description	Location
2	Abstract	Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	N/A
3	Problem Formation	Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	Introduction and previous chapters
4	Purpose or research question	Purpose of the study and specific objectives or questions	Methods, para. 1
5	Qualitative approach and research paradigm	Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/interpretivist) is also recommended; rationale*	Methods, Reflexivity and Trust

No.	Item	Description	Location
6	Researcher characteristics and reflexivity	Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	See Reflections on This Chapter and chapter 2
7	Context	Setting/site and salient contextual factors; rationale*	Methods, Procedures para. 1
8	Sampling strategy	How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale*	Methods, Sampling Strategy

No.	Item	Description	Location
9	Ethical issues pertaining to human subjects	Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	Methods, Ethics
10	Data collection methods	Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale*	Methods, Procedures
11	Data collection instruments and technologies	Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	Methods, Procedures; Appendix U

No.	Item	Description	Location
12	Units of study	Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	Results, Recruitment
13	Data processing	Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/deidentification of excerpts	Methods, Data Processing and Analysis
14	Data analysis	Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale*	Methods, Data Processing and Analysis

No.	Item	Description	Location
15	Techniques to enhance trustworthiness	Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale*	Methods, Reflexivity and Trust
16	Synthesis and interpretation	Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	Results, Main findings & Other findings
17	Links to empirical data	Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	Results, Main findings & Other findings

No.	Item	Description	Location
18	Integration with prior work, implications, transferability, and contribution(s) to the field	Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/ generalizability; identification of unique contribution(s) to scholarship in a discipline or field	Discussion para. 1-6
19	Limitations	Trustworthiness and limitations of findings	Discussion, Limitations
20	Conflicts of interest	Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	See chapter 2
21	Funding	Sources of funding and other support; role of funders in data collection, interpretation, and reporting	See Acknowledgements

Appendix R

Author interview recruitment image



Image posted to Twitter / X to recruit authors

Appendix S

Participant information sheet for author interviews

James Harwood DPhil Candidate in Musculoskeletal Sciences Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences Botnar Research Centre, University of Oxford Email: james.harwood@stx.ox.ac.uk

Seeking feedback on a website for researchers

S.0.1 Participant information sheet

What is the purpose of this study and why are we inviting you?

We are looking for feedback on a website from people that perform qualitative research.

Who is conducting the study?

This research is being conducted by researchers at the University of Oxford, and lead by James Harwood, a PhD student, supported by a grant from Cancer Research UK and supervised by Professor Gary Collins.

What will happen to me if I take part?

You will be invited to attend 2 online interviews within a month of each other, with a short reading task in between. Each interview will last around 45 minutes and you

will be given a \$50 (USD) voucher after the second interview. This will be an Amazon voucher for participants from Europe and the US. Participants from elsewhere will be able to suggest another appropriate voucher provider.

In the first interview we will ask about your opinions of the website and ask you to navigate it to find information.

You will then have a month to read a small section of the website and annotate any parts that you like or dislike with “+” and “-“ signs. You will be asked to use the website whilst writing a piece of research and we will ask you to share an excerpt of your writing.

In a second interview, we will ask about your writing, annotations and experiences using the website.

Once we have completed all interviews, you will have an opportunity to comment on our findings.

Do I have to take part?

Participation is voluntary and you can withdraw at any point, including during the interview, by notifying the project lead (james.harwood@stx.ox.ac.uk). If you choose to leave before completing both interviews and the writing task, you will not receive a voucher.

How will my data be used?

We will collect your name and email to communicate with you through the duration of the study. We will record the video of your interview to refer to during data analysis. We will ask for examples of your work to observe how you have interpreted the website content.

What happens to the data provided?

The information we acquire during the study is the research data. Any research data from which you can be identified, such as your contact details, consent record and video recording, is known as personal data. All data will be stored on secure servers at the University of Oxford and will be held by and accessible to the research team.

Personal data: Your contact details will be deleted as soon as the research is finished. We use Zoom to record interview sessions and a secure third-party software called Trint to turn video recordings into anonymised transcripts. Once these transcripts are made, all data will be deleted from Trint and Zoom within 6 months.

Research data and consent records will be stored for at least 3 years after publication or public release of the work of the research. Transcripts and writing samples will be anonymised. Responsible members of the University of Oxford may be given access to data for monitoring and/or audit of the research.

We would like your permission to use short, anonymised quotes from your interview in our publications and research presentations. All personal information that could identify you will be removed or changed before doing so. We will not use quotes from any writing samples you provide.

Will the research be published?

The research may be published in a student thesis, academic publications, and shared in academic presentations. The student thesis will be deposited both in print and online in the Oxford University Research Archive where it will be publicly available to facilitate its use in future research.

Who is funding the research?

This research is funded by Cancer Research UK.

Who has reviewed this study?

This study has been reviewed by the University of Oxford Central University Research Ethics Committee who considered it to be a service evaluation. Therefore, ethical approval was not required.

Who do I contact if I have a concern about the study or if I wish to complain? If you have a concern about this study, please email james.harwood@stx.ox.ac.uk who will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please

contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible.

Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB.

Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study.

The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest.

Further information about your rights with respect to your personal data is available from <https://compliance.admin.ox.ac.uk/individual-rights>.

Further Information and Contact Details

If you would like to discuss this study with someone beforehand (or if you have questions afterwards), please contact:

James Harwood DPhil Candidate in Musculoskeletal Sciences Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford University email: james.harwood@stx.ox.ac.uk

Appendix T

Emails for author interviews

T.0.1 Email 1: recruitment

Title: Seeking feedback on a website for researchers

Dear [NAME]

I am James Harwood, a researcher at the University of Oxford.

We are looking for volunteers to give feedback on a website for people doing research to help us understand what works and what might be improved.

We would love to hear from you if: You are currently doing qualitative research as part of your job or studies You are conducting, or writing-up a study

You will be given \$50 for attending two, 45-minute online interviews and completing a short reading and writing task in between. Participants from the UK will be given an Amazon voucher instead worth 50 GBP.

Please contact james.harwood@stx.ox.ac.uk to take part or ask questions. You can read the full information for participants [here](#).

With many thanks,

James

T.0.2 Email 2: consent

Title: Feedback on a website for researchers - consent

Dear [NAME],

Thank you for your interest in taking part in this study. We would love to hear your views.

Please could you:

1. Read the information for participants
2. Complete this informed consent form
3. Suggest some days and times for a 45 minute online interview within the next month. Please include your time zone.
4. Suggest some days and times for a second 45-minute online interview, within a month after the first. You will need enough time between interviews to spend some time working on your own research project.

Many thanks,

James

T.0.3 Email 3: scheduling

Title: Feedback on a website for researchers - confirmation

Dear [NAME],

Thank you for filling in the consent form and for sending those dates. I have sent you two calendar invites with the links for joining the Zoom meeting.

Let me know if you need to reschedule.

Many thanks,

James

T.0.4 Email 4: confirmation of second interview

Title: Feedback on a website for researchers - confirmation of second interview

Dear [NAME],

Thank you for sharing your opinions with me today. As discussed, items [insert item numbers] apply to the work you plan on doing before we next meet. As you read the guidance and write your work, please annotate the website with “+” or “-“ signs whenever you feel positive or negative towards it .

Please share your writing with me before we next speak. It can be a draft or work in progress. Nobody else will see it.

We will talk about your writing, annotations, and experiences using the website at our next meeting at [insert time and date].

Many thanks,

James

Appendix U

Author interview schedule

Below is a list of topics/questions to be discussed in this study. Following best practice, the qualitative work will remain flexible with respect to participants' agendas but we will cover the broad topics/questions noted. It is common in qualitative work to iteratively develop topics and questions as new ideas emerge from early data collection. Therefore, we may add new topics as the interviews progress and data collection continues. However, the key topics will remain the same.

U.0.1 Interview 1

Introduction

Thank you for agreeing to participate today. I am interviewing you to get feedback on a website we are building for researchers.

Your participation is completely voluntary and you may decline to answer any question, stop the interview or withdraw from the study at any time and for any reason.

Today's interview will take around 45 minutes. I'll then ask you to read something in your own time, as you get on with your normal, day-to-day research work, before inviting you back to a second interview, sometime within the next month, which will also take around 45 minutes. If you complete both interviews and the short task at home, we'll send you a 50 GBP (Amazon) voucher / \$50 dollars.

There are no right or wrong answers to any of my questions. I'm interested in your opinions and experiences, and I'd love it if you spoke your mind freely – the more you say, the more we can improve the website. So don't hold back, and please be honest. The site you'll see was made by someone else, so you won't offend me.

Do you have any questions at this stage?

With your permission, I'd like to record the interview. I'll record the audio because I don't want to miss any of your ideas. The audio recordings will be transcribed and anonymised. These anonymous transcripts will be stored as data, but you won't be identifiable from them.

I'm also going to record the video so that I can remember what you were looking at whilst speaking. The video recordings will remain confidential – only I will see them, and I'll delete them once I have finished analysing the data.

At times, I'm going to ask you to share your screen as you interact with the website so that I can understand what you are seeing. That screen share will also be recorded in the video, so can I please ask you to close any windows that you do not want to be on that recording.

[wait for participant]

Do you have any questions about what I have just explained?

May I start the recording?

Building rapport and demographic questions

Before we begin it would be nice if you could tell me a little bit about yourself.

- What are your research interests?
- Where do you work?
- How many years have you done research?
- Is English your first language?

Can you tell me a little about what you are working on?

- What stage are you at?

5 Second Test

I'm going to share my screen now. Hold on a moment....

[Confirm participant can see website]

I'd like you to have a look at this website.

[Wait 5 seconds before stopping sharing].

I only gave you a few seconds there. Can you tell me what you think the website is about?

- What do you think reporting guidelines are?
- What do you think reporting guidelines can be used for?
- How would you expect the website to impact your work?
- How would you describe the look and feel of the website?

If the participant has no clue, give them another 10 seconds.

Once the participant understands that the website is about reporting guidelines, ask the following:

- What experience do you have with writing-up research?
- What aspects of writing do you find most difficult?
- What parts of research do you find most difficult?
- Have you ever heard of reporting guidelines before, or perhaps names like STROBE, COREQ, SRQR?
- Have you ever heard of the EQUATOR Network before?
- Have you ever used a reporting guideline?
- Have you ever filled out a reporting checklist?
- If yes to either of above 2 questions: what did you think of it?

In a moment, I'm going to ask you to explore the website on your own computer, and I'd like you to say everything that is going through your mind as you do. I'd like you to tell

me what you notice, what you like, don't like, how it makes you feel, what you expect to happen if you click on things. Everything. I'll demonstrate what I mean quickly now [demonstrate think aloud].

[go back to home page]. OK so now you try. I'm going to show you the home page again and I'll leave it up on the screen for longer this time. Just tell me what you notice, how you feel, and what you expect. Verbalise everything that goes through your head.

[allow participant to warm up into think aloud].

[give feedback]

OK, I'm going to stop sharing my screen now and I'd like you to share yours instead. Remember to close anything that you don't want recorded.

[wait for screen share]

I'll paste the link into the chat [paste into chat]. Go ahead and open it.

Thank you.

You've already told me how you feel about the top of this page, so I'd like you to go ahead and start to scroll down, whilst continuing to tell me everything that you are thinking as it happens.

User Protocols

1. Find guidance for qualitative research

OK wonderful. Please scroll back to the top of the page. I'd like you to find a reporting guideline that you think would apply to your work. Please can you tell me which guideline you would click on?

[wait]

Why have you chosen this guideline?

- What kinds of research does this guideline apply to?

I'd like you to click on to SRQR guideline for qualitative research.

I'd like you to continue verbalising everything you notice and think on this page too.

[Once they get to the guidance]

Let's pause there for now. What do you expect from this guideline?

How trustworthy do you find this guideline? Why?

2. Finding item information

Ok great. Now I know the guidance is quite long and I don't expect you to read all of it now. But let's try finding some specific information instead. According to the guidance, what information do other people need to know about the researchers involved in the study?

- According to the guideline, why do readers need to know this?

[Consider doing another user protocol test if time permits]

- Have a look at the word "approach" at the top of the item (item 6). Why do you think it is underlined?
- Have a look at the number 4 at the end of the first sentence. What do you think it is?
- Now look at the title of the item - "Researcher Characteristics and reflexivity". Look at the icon to the right, what do you think that is?

We are nearing the end, so I'd like to talk about our next interview and your little bit of homework. Thinking about your own research, what do you expect to be working on over the next few weeks?

- Do you expect to write anything / have you already written anything?

[Tailor to stage of work]

I think you will find items [specify items] are relevant to what you're planning to work on. Can I ask you read through those items before we next talk? You can read through

the rest of the guidance too if you like, but I expect those items to be most relevant to your work.

I'd like you to annotate any parts that you like – for any reason – with a plus sign. And any parts that you don't like with a “-“. Let's practice that now. Have a read through [specify] and try highlighting parts that you like or don't like.

[give feedback].

Would you be happy to send me [portion of text they are working on] before we next speak so that we can talk about it? You can send it by email two days before we talk, so that I have enough time to read and understand it. Again, remember that all information will be kept confidential, and we won't quote anything that you have written without your permission. Does that sound OK to you?

[Schedule a date for the next interview and an expected deadline for submitting writing sample].

Thank you so much for your time today, and for everything that you shared with me. Feel free to email me with any questions, and I'm looking forward to speaking with you again soon.

U.0.2 Interview 2

Introduction

Hello again. Thank you for joining this second interview. I'm looking forward to hearing how you got on.

As a reminder, there are no right or wrong answers. With your permission I'm going to record today's session like I did last time. Am I ok to start recording now?

We will look at your plus and minus markings in a moment. Before we do, can you tell me, in your own words, how did you find [describe task that was set]?

Plus minus test

Let's take a look at your plus and minus markings.

[For each marking] Can you tell me why you put a [plus/minus] here?

Writing sample

Now I'd like to talk about the writing that you sent me.

[For items that were reported well]

- Can you tell me how you wrote this section?
- Can you tell me how the guidance influenced your work, if at all?
- Did you use any other resources? If so, what were they?

[For items not reported fully]

- Can you tell me how you wrote this section?
- What does this item mean to you? (referring to guidance)
- [If they misunderstood] Can you tell me where you have described this item?
- [If they understood correctly] The item recommends that you describe how you did [insert]. Why did you decide not to include that?

Overall, how do you feel about what you've written?

- Do you feel confident that it's complete?
- Do you like it?

Would you continue using the guidance for your work? Why?

Closing thoughts

Before we finish, I'd like to talk again about the website more generally. How would you describe it to a colleague?

Do you have some other thoughts about the website?

Thank you so much for your time. Would you like to see my findings once my research is finished? If you have time, you are welcome to comment on them to make sure I've represented your views fairly.

Please feel free to contact me at any time if you have any questions.

Appendix V

Deficiencies

Intervention components found to be deficient by one or more authors during qualitative interviews, as described in chapter 10.

V.1 Include design, features, and language to foster trust

Relevant website features: Professional design. EQUATOR's Logo remains prominent. Citation metrics are presented at the top the reporting guidance. Information about who developed the guidelines, how they developed it, and why the guidance is credible is still provided, and easily findable from the top of the guidance.

Barriers addressed: Researchers may not believe stated benefits

See text

V.2 Describe what reporting guidelines are where they are first encountered

Relevant website features: Prominent definition on home page and guideline page.

Barriers addressed: Researchers may not know what reporting guidelines are

See text

V.3 Use tone of voice and design to communicate personal benefits; confidence and simplicity

Relevant website features: A clean, simple interface for the home page and guidance pages. Text uses phrases like “confidence”, “quick”, “maximum impact”.

Barriers addressed: Researchers may not believe stated benefits

See text

V.4 Describe personal benefits and benefits to others where reporting guidelines are introduced (home page, on resources, in communications)

Relevant website features: Benefits are prominently and consistently displayed across the home page and guidance pages. Descriptions prioritize personal benefits to the authors above hypothetical benefits to others.

Barriers addressed: Researchers may not know what benefits to expect

See text

V.5 Clarify what tasks (e.g., writing, designing, or appraising research) guidelines and resources are designed for

Relevant website features: Clear instruction and differentiation of resources

Barriers addressed: Researchers may not know what reporting guidelines are; Researchers may not know when reporting guidelines should be used

See text

V.6 For each item, provide examples of reporting in different contexts

Relevant website features: SRQR already had some examples. No more examples added

Barriers addressed: Researchers may not know how to report an item in practice

See text

V.7 Define key terms

Relevant website features: SRQR now has a glossary, and text is marked-up with definitions that appear upon click.

Barriers addressed: Researchers may misunderstand

Participants wanted more definitions within the guidance *and* on the rest of the website

Some complex words in the redesigned SRQR guideline had blue dotted lines underneath. When participants clicked them, a definition would pop up. Participants stressed usefulness of this feature, and everybody liked that the definitions appeared in a popup, and not on a new page.

“I got great help there for me as a writer.” (ECR from Ecuador)

“One feature that I really found to be very useful was where you [described] different kind of kinds of terms there [...] When I’m not quite sure about the term there was. If I clicked on it, there was a description and that was very clearly written out and it made it quite easy to use the information that was there.” (Pre-PhD student from Malawi)

Participants wanted more definitions within the guidance (some struggled with words like “transferability” or “generalizability”) and other areas of the website. For example, when reading about the scope of SRQR and related guidelines, participants struggled to understand terms like “qualitative evidence synthesis”

“A qualitative evidence synthesis. What’s that? Is that like a review?” (Researcher from South Africa)

When reading a list of guidelines for different study types on the home page, one participant explained how not understanding terms like “cohort study” led her to feel anxious and worried.

“Actually, I feel a little anxious because I’m not sure whether it shows that I am not good at something or I am ignorant or something. So when I saw all of these and, like half of them, I didn’t know. I may feel anxious and maybe, like, lose confidence sometimes.” (PhD student from China)

User experience might be clearer if definitions were signified by a different colour (not blue)

Because I wanted to know whether future users would discover this functionality by themselves, I asked participants what they expected the dotted lines to signify. Most participants guessed correctly.

“Interviewer: What do you expect those dotted lines to mean? Participant: Umm, I think dotted lines would mean that it will give me a definition.” (Researcher from Australia)

“It’s not exactly a hyperlink, but it might give you a definition or something like that” (Researcher from the UK)

However, a couple of participants thought the lines might signify hyperlinks, and another thought it was a misspelt word because it looked like Microsoft Word’s autocorrect feature.

“I’m honestly not sure whether I see just a highlight or it’s a link. Usually links are in blue, completely in blue.” (Medical student from Ghana)

“I think uh, there are some misspelled words” (ECR from India)

V.8 Cater to different kinds of user (readers vs dippers) by structuring guidance with headings, itemisation, hyperlinking to particular sections, and with optional content

Relevant website features: SRQR items are structured consistently, making information easier to find. Itemisation is used consistently, content is hyperlinked when useful.

Barriers addressed: Researchers may expect the costs to outweigh benefits

Overall, participants welcomed the structure and navigation features, including headings, consistent subheadings for each guideline item, navigation menus, and links to sections.

“I find the guideline more interactive than simply, you know, a PDF document. So I really appreciate that it was more structured and the guide was more easy to follow. And somehow having all these other, umm options, you know that the web offers, yeah [...] For me it was very useful for moving around more easily and more quickly” (ECR from Ecuador)

Participants liked having extra content as notes, but footnotes might not be optimal implementation

When an SRQR item contained extra context or explanation not relevant to all users, I moved this information to a footnote, but participants expected the footnote identifiers (superscript numbers) to be references, and did not like how reading the footnote took them to the bottom of the page, away from where they had been reading:

“I think it’s a citation. So uh linked to a reference. But I think it may also be foot note.” (Researcher from the UK)

“[Superscript numbers are] always a link to, like, a reference. If not a reference like some... [clicks it]... okay, so it’s telling me things. It’s probably not a reference like in Wikipedia. It shows you it’s a link. This one is telling me more things. So it’s a footnote.” (Research consultant from the Philippines)

“I worked hard to scroll down. I’m not gonna go back up. So, whatever, i’m not reading [it].” (Research consultant from the Philippines)

Menu navigation was useful but could be more prominent and go one level deeper

The redesigned SRQR guideline has a side navigation menu so users can easily navigate to the Introduction, Methods, Results, and Discussion guidance sections, or to the FAQs and citation information. A few participants did not notice the menu, or suggested it could be more prominent, and one suggested it would be useful if the menu listed all reporting items (i.e., not just *Methods*, but also the item titles within).

“when I scrolling I forgot to pay attention to the menu on the side. So [...] I suddenly realized it was a menu there [...] Maybe make it much more, like, the characters larger or, like, the colour, maybe change it, because now it is so light and I didn’t pay attention to that. But this menu is useful if I pay attention to this.” (PhD student from China)

Each reporting item has its own menu within the collapsible content, so users can jump to the item’s justification, examples, or related resources. The same participant described these item menus as a useful way to help “find the detailed part” they desired.

Section URLs could be more intuitive

When the mouse hovers above a reporting item heading, a button would appear to allow users to copy the URL to that section. One participant did not find this intuitive, and thought the button would take them to more information instead.

Participants may desire a summary of the guidance

In the think aloud, one participant said they would prefer to browse the checklist above the full guidance because “in my head, like the checklist probably has these things already” (Research consultant from the Philippines). They expected the checklist to act as a summary. This expectation may explain why another participant said “I just want to see what this thing is and at the moment I don’t know. Should I read the guidelines or checklist?”, perhaps because they desired an overview (Researcher from South Africa).

"So maybe not the full guideline, but I'm not sure how you can present it. Maybe one paragraph, or, uh, or you could present a mind map. Maybe you can present a mind map and you can put the keywords of this guideline. So if people want to see, for example, [...] introduction, literature review, methodology, blah blah. Something like this." (PhD student from China)

Participants found item titles hard to spot when scrolling

"The names of the sections are, like, black and a little plain. So I think [...] the name of each section should be somehow highlighted" (ECR from Ecuador)

"And they don't really stand out in my opinion, like, just about the layout of the page itself. You see, this [heading] doesn't stand out so much from the rest, just the font is a little bigger. For example, if I'm scrolling through it very fast, it may be I might miss it. So if those headings or subheadings could be made more visible..." (Medical student from Ghana)

Some participants did not expect / want all guidance on one page

"I didn't expect it to be part of the same page because, just looking at the side [menu] here, it makes it a very long page to scroll through. So I guess what I'm a bit more familiar with is, you know, you click on something, it takes you to the guidance and then each of [the items] might be on a separate page [...] So it's less of a scroll and more of a click." (Researcher from the UK)

V.9 Explain how the guidance was developed and why it can be trusted

Relevant website features: Brief description included on home page and at top of reporting guideline, links to full development information

Barriers addressed: Researchers may feel patronized

The relationship between the website and publication could be clearer

The top of the guideline page references the original SRQR publication. Some participants said this helped them trust the website, but others voiced confusion about the relationship between the publication and webpage, whether the guidance matched, and which they should cite.

“I know I have the information to look for the paper and I could find it. So yeah, it inspires trust and it’s amazing to have it here.” (ECR from Ecuador)

“Is this [content on the website] now this article? Because if it is the article then I’m not clear on that [...] [After reading through the item titles] Do they match?... Yes, they match.” (Researcher from South Africa)

The FAQs helped clear up this relationship, and some participants recommended they be signposted (or briefly explained) at the top of the guidance page.

“[The FAQ section about development] gave me more information about how [the guideline] came about. You know, I was asking [for] that when we first spoke, and here it was. It was nice to be able to see [that] somebody put this together. [...] I like that I could get more information about the background. So in a way it’s not really a frequently asked question. It is actually the background.” (Researcher from South Africa)

“I think it’s very fine to put it here [in the FAQs], but I’m just thinking about, like, could we put a link or a summary at the very beginning of the website? Because it may make me trust it much more quickly.” (PhD student from China)

V.10 Make guidance appear shorter by removing superfluous information, hiding optional content, splitting long guidelines, using concise language, and separating design advice

Relevant website features: SRQR has been edited. The only text presented immediately is instruction on what the author needs to describe. Additional information is hidden at first and can be expanded. Text is shortened through editing and by using active voice. In the case of SRQR, this reduced the text length by 60%.

Barriers addressed: Researchers may expect the costs to outweigh benefits

On first impression, the guidance may appear too long

Participants spoke positively of the various features to hide optional content (like collapsible sections and pop-ups). Nobody expressed wanting all information up front, and all preferred to have it only when needed.

“If you want to, you can extend and read through [the extra information]. I think that’s very useful.” (Medical student from Ghana)

“but if you don’t need [the extra information], uh, you can’t see it. But if you need it, you may click on it and just show it very directly.” (PhD student from China)

However, some participants still felt the guidance appeared very long, and this could be off-putting upon first impression.

“So I kind of know how long this page is based on seeing [the scroll bar] move. I don’t want to read the whole thing [...] I don’t want to spend 17 min on that” (Research consultant from the Philippines)

In the first iteration, the page was even longer because of content in the introduction before the guidance began (e.g., information about the guideline’s scope, development, or how to use it).

"There's so much going on [at the top of the page, before the guidance]. [...]"

Yo, this is getting annoying. All I want is the guideline or the document"

(Researcher from South Africa)

In the second iteration, I moved some of this introduction information into a collapsible box, but some participants still described the page as very long.

V.11 Create spaces for authors to discuss reporting guidelines with others

Relevant website features: Each reporting item has its own discussion board.

Barriers addressed: Researchers may not believe stated benefits; Researchers may misunderstand; Researchers may not know how to report an item in practice; Reporting guidelines can become outdated; Researchers may feel patronized

Each reporting item had a link to its own discussion page. Participants welcomed the opportunity to post comments, ask questions, and suggest edits on this page.

"So you do not want to just show people how to do it, but you want to provide a platform for the readers to discuss and to create their own ideas. I think that's great." (PhD student from China)

"Maybe what helps is if you can, say, share your thoughts here, [...] you feel a little bit more [that] *I've got a voice* instead of *I'm just being told that it's a great place to be and it's got all the answers.*" (Researcher from South Africa)

"Maybe if I identify something that I know is not useful for me. Yeah, in that case, maybe I would write something, you know, to other researchers to read it and maybe give feedback about it, but also for developers of the guideline."

(ECR from Ecuador)

The button could be clearer

The link to the discussion page was a button with an icon of a speech bubble, no text, and not all participants knew what this button would do or where it might take them.

One participant said this button would be easier to understand if it used text, and recommended it be placed at the end of the expandable content for each item.

“Maybe the blog? No, I don’t know.” (ECR from Ecuador)

“I thought it might be, uh, like a frequently asked questions or a or a further explanation” (Researcher from Australia)

“I think you can just put [a button] after all the [extra] content you have seen [in the expandable box], because if if people haven’t seen this content they will have nothing [to discuss].” (PhD student from China)

The discussion page could be easier to use

The discussion page requires users to sign in before they can post. This deterred some participants.

“Ohh I see I need to sign in. Ah, usually that is something that discourages me immediately.” (Medical student from Ghana)

“I would comment, if I didn’t have to log in” (Research consultant from the Philippines)

Additionally, one participant did not understand the discussion page’s purpose.

“Am I doing this to myself? Is it comments about my research that I’m completing the guidance for or is it for the website for other people to see these comments that I’ve made perhaps?” (Researcher from the UK)

V.12 Address communications to authors

Relevant website features: All resources and website copy are directed predominantly at authors.

Barriers addressed: Researchers may feel that checking reporting is someone else’s job.

Participants expected the target audience was researchers, but the target could be more specific

Everyone recognised the website was for researchers, as opposed to editors. Most realised it was for *health* researchers.

“It can be useful for academics. It can be useful for people who do research on daily basis or as a regular part of their job” (ECR from India)

Some participants specified the target might be *early career* researchers or university students.

“those who are just getting started in the career journey in research....[pause]...definitely students?...umm.... people who don’t really have that much experience in writing.” (Pre-PhD student from Malawi)

“Maybe like master research or higher degrees and also for some junior scholars” (PhD student from China)

“But I think your your target audience is completely different than the very basic student... because it’s more for the expert researchers, right... for the expert researchers” (ECR from India)

Some participants hinted that seeing terms specific to their field signalled that the website was for them:

“If I know any of these terms, then clearly I’m gonna make use of it” (ECR from India)

Ohh then maybe if it sounded like “Want help writing up **Qualitative research** or something like that, so that I know that this is actually addressing me. It’s relevant to me. Yeah, but [the way it is written now] way is kind of open. So it’s not kind of, it’s not relating to me. (Pre-PhD student from Malawi)

V.13 Structure guideline items to make them quicker to digest

Relevant website features: Items have consistent structure and use bullet points consistently

Barriers addressed: Researchers have limited time

Participants welcomed items' structure and said they preferred it to unstructured text.

"It was more structured and the guide was more easy to follow." (ECR from Ecuador)

One participant questioned whether they might prefer full sentences over bullet points. They used SRQR's item 12 as an example, where the text says:

"If the actual sample differs from the target sample, describe:

- the difference,
- why these differences may have occurred,
- how this might affect the findings."

The participant suggested changing this to "Why the difference has occurred and how it might affect the findings. So one line and you're done with it", instead of taking up four lines of text." (ECR from India)

V.14 Describe the scope of a reporting guideline at the top of every resource

Relevant website features: The intended scope of a guideline is clearly & prominently described. This definition includes contexts in which the guidance should not be used.

Barriers addressed: Researchers may not know whether a reporting guideline applies to them

Not all participants found the scope clear

One participant was not clear whether SRQR applied to survey studies, and said they wanted a definition of what *qualitative research* meant. Another participant wondered why the guidance discussed patient outcomes when the scope had not specified health research.

V.15 Search function on website

Relevant website features: Search function is easier to find as a recognizable icon in the navigation bar of every page. The home page includes additional ways to access search functionality.

Barriers addressed: Guidance may be difficult to find

The website featured search buttons (one in the navigation menu, one at the top of the home page), but the search buttons did not work, so participants could not explore the search functionality. However, many participants commented on the search buttons and instinctively knew what they were.

V.16 Include testimonials from research users who benefit from complete reporting

Relevant website features: SRQR includes dummy testimonials and quotes from research users

Barriers addressed: Researchers may not care about the benefits of using a reporting guideline

Quotes helped participants believe items were important, but some participants questioned their credibility

The redesigned SRQR guidance featured quotes in the margin of some items from people that use research, like evidence synthesisers, editors, or other researchers. These quotes were made up, as I didn't have time to collect real ones just for testing. Some participants said they liked these quotes:

“I think it makes this....how can I say?...practical from a different point of view. Like why exactly you need this [reporting item]. So now this person [in the quote] is telling from her own perspective how useful it is that you have [the item] described clearly, so it makes it such that if I’m trying to describe [this item], I’ll try to keep that in mind that.” (Medical student from Ghana)

“I like it because each of them tells me the why... umm...you know, kind of gives a plain language reason for [...] why it’s a useful thing. [...] That’s gives it, you know, humanity.” (Researcher from Australia)

However, some participants questioned whether these quotes were from real people.

“Maybe they’re real... maybe it’s legit” (Research consultant from the Philippines)

Quotes may not be valuable enough to be prominent

Some participants didn’t find the quotes useful, or described them as distracting.

“I don’t care what people think.” (Researcher from South Africa)

“I didn’t really pay much attention to it. I saw a few quotes and read through a few quotes, but I was like, *OK, so what does that add to me?*” (Pre-PhD student from Malawi)

“sometimes when I have to read [something complicated], but I don’t want to, [distraction] may be a big problem for me because, like, these comments are what people say, so it’s, like, more easy to read than the [reporting item]. So it may attract a lot of attention from me” (PhD student from China)

V.17 Include testimonials from researchers who were nervous about being punished for reporting transparently

Relevant website features: Quotes included alongside guideline

Barriers addressed: Researchers may expect the costs to outweigh benefits

Some of the fictitious quotes contained reassurance from researchers who had felt nervous when using a reporting guidelines because they felt unsure about being so transparent about parts of their work they knew were not perfect. There were only a couple of quotes of this kind, and few authors noticed or commented on them.

“That makes it relatable to a user, particularly a new user, because we can see that all of these people are, you know, they were first time users once.”
(Researcher from Australia)

V.18 Instruct authors to cite reporting guidelines so readers may learn about them

Relevant website features: Consistent instruction to cite reporting guidelines

Barriers addressed: Researchers may not know what reporting guidelines exist

The title *How to cite* was misleading to one participant

Participants expected citation instructions and most were not surprised to see them in a section called *How to cite*. However, one participant questioned whether the section would tell them how to cite the guidance, or how to cite resources in general.

“I want to see, like, if it says right away, a certain citation style like, are you talking about Chicago?” (Research consultant from the Philippines)

Some participants did not know whether to cite the website or the publication

One participant said they had sought the original paper because they “wanted to add the citation” (ECR from Ecuador) to their article. Another asked “Do you reference the EQUATOR network or do you reference the original article?” (Researcher from South Africa).

Some participants recommended the citation instruction be more prominent

“I think you can put citation in a in a box in a [coloured] box to make people notice it.” (PhD student from China)

"It is helpful of course, umm, but I think it got lost in the page." (ECR from Ecuador)

V.19 Explain when reporting guidelines do not intended to prescribe structure

Relevant website features: Explained at top of guidance

Barriers addressed: Researchers may struggle to reconcile multiple sets of guidance

Some participants included sub headings for each reporting item in their writing sample

Many participants noticed and welcomed the clarification that SRQR does not prescribe structure. Nobody objected to it.

"I think this is very clear. If you say that did not prescribe order nor structure. [...] Yes, that's great. [...] I like it very much because I think it helps me understand that the guideline is not a standardized one, but it depends on the user. To use this kind guide in their own context. So I think this one this sentence is is very important." (PhD student from China)

In the writing task, two participants used subheadings for the items they wrote. When asked about these subheadings in the second interview, both said they intended to remove at least some of the subheadings. One participant replied "One of our reviewers wanted more structure, so of course we leave some [subheadings] in, but not all" (ECR from Ecuador). Another said "I'll probably not include them [in my final article], but the information, that's what I'm I'm going to maintain, yeah." (Pre-PhD student from Malawi)

V.20 Provide links to other resources that explain how an item can be done

Relevant website features: Links included when relevant.

Barriers addressed: Researchers may not know how to do an item

Participants mentioned needs that are not addressed by the current links

Most participants voiced support for links to helpful resources. None were against them.

“I wasn’t really expecting this here. But then it’s useful. These two [links] seem useful.” (Medical student from Ghana)

However, participants wanted links to other resources, including resources to help them with “flow charts” (ECR from Ecuador), “writ[ing] in a concise manner” (Pre-PhD student from Malawi), “sample size calculations” (Researcher from the UK), and more item-specific “training” (Medical student from Ghana), possibly including “videos” (Midwifery student from Uganda).

V.21 Links between related guidelines

Relevant website features: Guidelines prominently link to other relevant guidelines and explain when they should be used.

Barriers addressed: Researchers may not know what reporting guidelines exist

The top of the guideline page included information on SRQR’s scope and links to other, related guidelines. Participants said they liked these links. Most found the explanations clear, including the instructions of when *not* to use SRQR.

“This part, *when I should not use this guideline*, I think it is much useful for me to understand whether I use the right guideline or not. Yeah. [...] I need this judgment.” (PhD student from China)

“[Reading out loud] *Do not use it for writing a qualitative evidence synthesis. use this instead.* OK, that could be useful.” (Medical student from Ghana)

Not all participants understood the term ‘related guidelines’ or differences between similar reporting guidelines

“why would you put *related* if if you’re saying *use this one, but then you could also use this one...* I think that really starts to get confusing.” (Researcher from South Africa)

“Initially it seemed confusing which guidelines to use when doing mixed methods research” (Researcher from the UK)

V.22 Educate authors about writing as a process

Relevant website features: Some SRQR items now link to relevant EQUATOR materials and courses.

Barriers addressed: Researchers may not consider writing as reporting

The website did not sufficiently address participants' need for writing training

The top of the SRQR guideline page had short instructions about how to apply the reporting guideline whilst writing but for many participants, this was not enough. Many participants expressed desire for training on *how* to write an article (as opposed to *what* to write).

One participant wanted training on how to “write in a concise manner” (Pre-PhD student from Malawi)

Another talked about their struggle to apply training and guidance to their own writing process.

“When I try to look at your guidance on your website, I really want to use it in my own writing, but it is very strange because when I try to, uh, connect the information on the website with my own writing, I found there there might be a great gap because I think everything on your website is very clear (actually they are very specific, those suggestions), but when I try to connect those information with my own writing, I found it just a little bit difficult to generate some some specific ideas to start my writing.

So I'm thinking if that's because the problem of my writing is not the lack of specific guidance but some other thing like my motivation or, I don't know... it's just a little bit strange.

And I also talk about this with my friends because lots of my friends, they are also PhD students and they are struggling at writing too. So ask them if they have some guidance, uh, if they have looked at some guidance and if [they] have put those guidelines in [their] own writing and and their answers were, like, quite similar with me and and they all talk about that, '*yes, we we look at lots of guidance we try to look at lots of those writing books to teach you how to write, to teach you how to structure your writing. But it's still very hard*'.

When you really sit down and start writing, actually you couldn't, uh, call up [the information].” (PhD student from China)

V.23 For each item, explain why the *information* is important and to whom (not just what constitutes “good” design)

Relevant website features: Information added when necessary

Barriers addressed: Researchers may not know why items are important

The labels *Justification* and *Why readers need this information* were ambiguous to some participants

Each reporting item included a section within its expandable content called *Why readers need this information* alongside examples of writing and links to other resources. Participants found the information useful but were less enthusiastic than they were for the examples.

“So yeah, this this section, justification, examples and resources. I found them very useful as they give more detailed explanations on each specific section, and particularly the examples” (Medical student from Ghana)

However, a few readers interpreted the title of this subsection differently.

"I thought it might give you an example of how to justify [the choices you made in your study]" (Researcher from the UK)

"I was asking myself: *which readers?*.... the one like me who is using the platform or the readers of my paper?" (Midwifery student from Uganda)

Not all participants found the justification compelling

Whereas many participants spontaneously said how useful examples were, none said the same about the justification section.

"[It was] one of those things that you just read and go, like, *ohh OK*. But not, like, something which I requested or you sit down and think through." (Pre-PhD student from Malawi)

Interviewer: Maybe this section could do a better job of reminding authors who is going to be reading their research, and how different those people might be. Or their different perspectives or experiences. Do you think if it had done that you would have found it a bit more convincing and motivating?

Participant: Umm. Definitely, definitely. From your explanation, I think it kind of puts it in context, [as to] why this section is more important. So I didn't read it at first, but I think it did not come up very clearly as compared to how you've explained it. And so, after hearing that explanation, I'd say that this the way the readers need information." (Pre-PhD student from Malawi)

V.24 For each item, provide clear instruction of what needs to be described

Relevant website features: Writing instruction occurs first for each item.

Barriers addressed: Researchers may not know how to report an item in practice

One participant felt the instructions could be a little longer and less ambiguous

Each reporting item began with instructions of what to write. Participants welcomed these instructions, as articulated by one participant when describing how they used the guideline for the writing task, “I used it specially for understanding what kind of information, you know, apart from the obvious ones, every section wants me to write, so it it was really helpful in that respect” (ECR from Ecuador).

However, one participant suggested the information could be a little longer: “I guess just a few more words to kind of, articulate [the instruction] better or give it a bit more detail [because it may be ambiguous to] somebody less experienced” (Researcher from the UK).

V.25 Gather and communicate evidence for benefits

Relevant website features: Dummy quotes provides evidence for experienced benefits.

Barriers addressed: Researchers may not believe stated benefits

Participants questioned whether quotes were credible

The website included (fake) quotes from authors exalting the personal benefit reporting guidelines have brought to their job (e.g., easier writing, smoother publishing). Some participants found these compelling, but others worried these quotes were biased, fake, or vague.

“This is encouraging because it’s a feedback from a person working at the journal with responsibility for publishing. The editor.” (Midwifery student from Uganda)

“If you’re using quotes on a website, my first thought is you could be biasing, you know, how useful the website is, because you’ve picked out great quotes so, you know, until I’ve looked at this and used it myself, I might not agree” (Researcher from South Africa)

“Maybe you can put some concrete stories, because for me I like to read stories instead of just reading some uh, broad words like I like it. Yeah.” - (PhD student from China)

“Yes, I think it’s much more useful [to] have some timelines to show when people say [things], yeah, because if I could see the time, I may also check like how long the company or the website may [have] last[ed] for [...], how many people use that and how they are feeling like from one year ago or two years ago to now.” (PhD student from China)

V.26 Provide advice regarding how to respond if asked to remove reporting guideline content by a colleague, editor, or reviewer

Relevant website features: Advice given in FAQ

Barriers addressed: Researchers may be asked to remove reporting guideline content

Many participants did not notice the advice

The FAQ section included some advice on what to do if a colleague or editor asked you to remove content from your manuscript pertaining to one or more SRQR items. Only one participant noticed this advice. When talking about what they would do if asked to remove content by “a reviewer or, even a co-author, it was amazing for me that you give these tips, you know, to what to do in those cases and what you can do to, umm, make it or to highlight the importance of have each section within your paper.” (ECR from Ecuador)

Another participant, who had not noticed the advice, nevertheless said they would return to the website should they be asked to remove content:

“I think first I will ask his or her reason why he or she want me to delete that if I think, I don’t agree with that. I may try to make some formal explanation to explain that and also I may come back to this website to double check some

reliable things which can support my view to let them believe that I should also I I need this paragraph." (PhD student from China)

V.27 Use plain language

Relevant website features: SRQR is edited to use plainer language.

Barriers addressed: Researchers may misunderstand

Participants did not understand all words

I had tried to use plain language on the home page and introduction to SRQR, but participants still questioned the meaning of some words (e.g., “synthesis”), including words commonly used by EQUATOR staff (e.g., “transparency”). Even the term “reporting guidelines” confused some participants, especially when used at the start of a sentence, where some participants interpreted the word “reporting” to be a verb, instead of part of a compound noun.

V.28 Link all resources to each other

Relevant website features: Guidance links to all tools and development article

Barriers addressed: Researchers may not know what resources exist for a reporting guideline

Participants noticed links to resources, such as the buttons to download checklists and templates.

V.29 Reassure that all research has limitations to encourage explanation over perfect design

Relevant website features: This reassurance appears on the home page and all guidance pages

Barriers addressed: Researchers may expect the costs to outweigh benefits; Researchers may feel afraid to report transparently

Two participants spoke about content that reassured authors to be honest about limitations. After noticing a reassuring quote from an editor, one participant agreed, saying:

“sometimes when we don’t report limitations and the reviewer identifies the limitations [instead] then then we lose credibility. So it’s better we report **limitations** so that the the reviewers say *ohh this person has acknowledged this limitation, then then this is a good study.* So I just liked it.” (Midwifery student from Uganda)

Another voiced support for the SRQR reporting item about limitations, saying “no study is ever done perfectly or done in a way that the next person would do it. And so we’ve become really, really good, I think, in my team, at making quite long limitation sections to try and avoid peer reviewers from finding everything that’s wrong”. (Researcher from South Africa)

V.30 Tell authors when to use reporting guidelines, or that reporting guidelines are best used as early as possible

Relevant website features: Stated prominently

Barriers addressed: Researchers may not know when reporting guidelines should be used

Upon realising reporting guidelines are for writing articles, all participants naturally thought about using them in the drafting/writing process (as opposed to retrospectively check manuscripts that are already written). Sometimes this was immediate:

“From what I have seen [after the 5 second test], I think probably the website should be about, uh, helping you try to discover... identify the guidelines that you will use for writing your study quickly.” (Midwifery student from Uganda)

For some participants, using the guidance for the writing task reaffirmed their opinion that the guidelines would help them in early stages of writing.

"I saw it as very important for me at this stage that I'm at. I'm writing a manuscript. But I'd say that it can be usable at any level, only that I found it very important at my level, which is at my paper writing stage." (Pre-PhD student from Malawi)

No participants talked about consulting reporting guidelines when planning a study

Perhaps because participants did not intuitively understand how they could use a reporting guideline to plan research, none talked about using them in that way.

V.31 Use consistent terms

Relevant website features: The website uses terms consistently

Barriers addressed: Researchers may misunderstand

Participants were confused when different words referred to the same thing

Two participants questioned why the home page referred to "guidelines" *and* "reporting guidelines" and asked whether there was a difference. Another asked whether "guidance" and "guideline" were the same. Another participant noticed that an SRQR item 18 uses the word 'integration' in two different contexts: whereas the item asks authors to integrate their work with others', the example uses the word 'integrating' differently when discussing how their study combined modes of teaching.

V.32 Avoid patronizing language

Relevant website features: Continue to avoid using patronizing language

Barriers addressed: Researchers may feel patronized

No participants mentioned feeling patronized. When asked, one participant described the tone as “suitable” and said “simple comments” and “explaining like this, [is not patronizing] because I also need a lot to learn.” (PhD student from China)

V.33 Centralised hosting

Relevant website features: A core set of frequently accessed guidelines are now presented on a single website.

Barriers addressed: Guidance may be difficult to find

That the website contains many different reporting guidelines was rarely mentioned, but one participant said “to have a repository where all those things are, instead of having to go and search for them. That makes sense.” (Researcher from South Africa)

V.34 Provide translations

Relevant website features: Translations are prominently listed above the guidance

Barriers addressed: Researchers may not understand the language; Researchers may misunderstand

One participant noticed that their language was missing

“So what about, you know, all the researchers that speak Spanish?” (ECR from Ecuador)

V.35 Tell authors how long the guidance will take to read

Relevant website features: Estimated reading time given

Barriers addressed: Researchers have limited time

Not all participants liked being told how long the guidance would take to read

The guidance advises readers that it may take 16 minutes to read. One participant liked this: “Oh, I know how it’s going to take. It’s helpful when you are performing research you have like tight deadlines and not much time.” (ECR from Ecuador)

But another felt like 16 minutes was too long, especially if you only found the guideline as part of manuscript submission “Uh, it will be frustrating. Now again, you waste a lot of time.” (Midwifery student from Uganda).

And a third (not a native English speaker) worried they would feel bad if the guidance took them longer than the stated time:

“I think this is, like, useful, but I’m not sure whether some... when some people read it, they may feel stressful because, from my teaching experience before, some of my students may say to me if, like, there is guidance saying that you may need, like, 60 minute to read it, sometimes, if they take longer, they may feel confused, or lose a little bit [of] confidence [and worry that] they read it so slow or they are not, like, normal one. I think this [estimated time] *may* be a good [information] because, like, nowadays I found, like, many websites use it. But actually, for me sometimes this part is useful, but sometimes it is not.” (PhD student from China)

V.36 Encourage explanation even when choices are unusual or not optimal

Relevant website features: Added to items

Barriers addressed: Researchers may feel afraid to report transparently

Only one participant noticed this instruction

Some reporting items in SRQR ask authors to explain their reasoning behind design choices. Only one participant noticed one of these sentences. They reflected that researchers often “follow this path of only mentioning [what we did] but not explaining how we did it and why it was important to apply this strategy” (ECR from Ecuador).

V.37 Explicitly state when no better guidance exists for a use case

Relevant website features: Reporting guidelines warn authors when no better guidance exists for a use case, and how the current guidance can be adapted instead

Barriers addressed: Researchers may not know what reporting guideline is their best fit

Only one participant commented on this instruction

In the introduction to SRQR, where its scope is explained, the instruction mentions that there are no better reporting guidelines for writing *protocols* for qualitative research, and instead recommends authors use certain items from SRQR. One participant annotated this explanation and said they liked it (PhD student from China).

V.38 Provide instruction as to how and where information can be reported without breaching word count limits or making articles bloated.

Relevant website features: Added instruction at top of reporting guideline and in some items where most useful

Barriers addressed: Researchers may struggle to keep writing concise

Only one participant commented on instructions of where content can be reported

Very little of the guideline text deals with *where* content can be reported (e.g., in the article body, in an appendix, a table, a figure). Only one participant noticed it and said “I like the reminder” (Researcher from South Africa).

V.39 Reassure when guidelines are *just* guidelines

Relevant website features: Prominently displayed at top of reporting guideline

Barriers addressed: Researchers may feel restricted if reporting guidelines prescribe design

Only one participant drew the distinction that reporting guidelines are not rules

Throughout the homepage and SRQR page I had tried to convey that reporting guidelines are recommendations, and I took care not to use words like *rules* or *standards*. Only one participant commented explicitly about this, but no participants talked about the guideline as if it were a set of rules.

“So I think this explanation here is very clear that it helps researchers to know that definitely they can have their own ideas and this guideline is, it is kind of like a supporting one, but not a rule, not a standardized rule.” (PhD student from China)

V.40 Ensure all resources and tools (e.g., checklists and templates) are in ready-to-use formats

Relevant website features: No changes made

Barriers addressed: Researchers have limited time; Reporting guideline resources may not be in usable formats

Even though the links to the checklist and template were not live, participants expected the resources to be in ready-to-use formats.

V.41 Decrease fear of judgement by making reporting guidelines design agnostic

Relevant website features: SRQR explicitly states that it makes no assumptions about design. Inadvertent design assumptions were edited.

Barriers addressed: Researchers may expect the costs to outweigh benefits; Researchers may feel afraid to report transparently; Researchers may feel restricted if reporting guidelines prescribe design

No participants commented on this.

V.42 Explain importance of complete reporting to the scientific community

Relevant website features: Continue to do this

Barriers addressed: Researchers may not know why items are important

The website explains the societal and community importance of complete reporting in a few places: the justification subsections of each item, quotes in the margin, and briefly on the home page.

Although participants commented on the quotes from research consumers, and on the *Justification* sections within each reporting item, nobody talked about the negative impact of poor reporting on the scientific community at scale.

V.43 Provide clear instruction of what needs to be described when an item was not done, could not be done, or does not apply

Relevant website features: Instructed where relevant

Barriers addressed: Researchers may not know what to write when they cannot report an item

This component was not tested and no participants commented on this.

V.44 Remove branding and messaging that may invoke feelings of judgement, complexity, or administrative red-tape

Relevant website features: A clean, simple interface for the home page and guidance pages. Text makes less use of judgemental phrases and fewer references to the negative consequences of poor reporting.

Barriers addressed: Researchers may expect the costs to outweigh benefits

No participants described the design as unpleasant or judgemental.

V.45 Search Engine Optimization

Relevant website features: The site has additional meta-data. Each reporting guideline page has its own meta-data. The site is optimized for mobiles.

Barriers addressed: Guidance may be difficult to find; Researchers may not encounter reporting guidelines early enough to act on them

This component was not tested and no participants commented on this.

V.46 Use if-then rules to direct authors to more appropriate and up-to-date guidance when available

Relevant website features: Reporting guidelines clearly and consistently point authors to more appropriate guidance when appropriate, using if-then rules. These links can be updated any time.

Barriers addressed: Researchers may not know what reporting guideline is their best fit

Although participants commented on the links to related guidelines, they did not comment on the “if...then...” structure of these links.