

Quality in Qual:

A proposed framework to
commission, judge and generate
good qualitative evaluation in wellbeing impacts

August 2021



Authors:

Jonathan Breckon & Ruth Puttick

About the authors

Jonathan Breckon is an independent consultant working on research and evidence use. He was formerly Director of Nesta's Alliance for Useful Evidence for nine years, a network that champions and promotes the smarter use of evidence in social policy and practice. Previously Director of Policy and Public Affairs at the Arts and Humanities Research Council, his research interests cover politics and psychology, including the relationship between research and policymaking.

Jonathan can be contacted at jonathan@jbreckon.com

Ruth Puttick runs a research consultancy providing advice to a range of charities, foundations, companies and governments, including Nesta, the Cabinet Office's Open Innovation Team, and Bloomberg Philanthropies. She previously worked at Nesta, leading large studies into public and social innovation, established the Alliance for Useful Evidence and successfully advocated for the creation of the What Works Network of evidence centres. More recently she was at the global management consultancy, Tony Blair Associates, where she helped establish and manage the policy and research team in the government advisory practice.

Ruth can be contacted at ruth@ruthputtick.com

About the What Works Centre for Wellbeing

We are an independent collaborating centre that develops and shares robust and accessible wellbeing evidence to improve decision making that is used by governments, businesses and civil society.

This work was commissioned and published by the What Works Centre for Wellbeing. It was funded by The National Lottery Community Fund.

Table of contents

Introduction4

- What are standards and evidence, and how can these be useful in navigating different methods?..... 4
- A note on the methodology..... 5
- How to use this report..... 6
- How the report is structured..... 6

Chapter 1

Why this work matters 7

- What is qualitative evaluation and why is it useful? 7
- Key messages from interviews 8
- Example A - Being explicit about ethics: a realist evaluation in area of child sexual exploitation 9
- Example B - Bringing the findings back home: listening to services users 10
- Example C - The value of mixed-methods: an evaluation of early support for military families 11
- Example D - Contribution analysis as a way of evaluating impact 12
- Limitations and barriers when using qualitative evaluation methods and when interpreting it..... 13
- The role of standards of evidence frameworks..... 14

Chapter 2

Existing help and guidance on qualitative evaluation methods..... 15

- What is *not* included in our list..... 17

Chapter 3

Developing a new framework 24

- Developing a framework for qualitative evaluation 24
- The six quality in qual prompts: questions and considerations..... 25
- Example E - Being open about limitations is a strength..... 27
- A few caveats..... 28

Chapter 4

Discussion, next steps and recommendations..... 30

- Is there a need for a new framework for qualitative evaluation methods? ... 30
- Explore what scale any new framework or standards of evidence should be graded on to establish the appropriate evaluation method 31
- Improve the guidance on the value of qualitative evaluation, and how it can be used ... 31
- Recommendations and next steps..... 31
- Prompts for further discussion and debate 32

Annex 1: Glossary of key terms 33

Annex 2: Further reading and useful resources 35

Annex 3: List of interviewees and interview question guide 36

Annex 4: Common pitfalls in qualitative research 37

Introduction

This discussion paper sets out a framework to help organisations choose and judge the most appropriate qualitative evaluation methods, with a focus on understanding wellbeing impacts.

The paper outlines the need for this framework and discusses the problems which can be encountered when assessing, commissioning and generating good qualitative evaluation, as well as the contextual factors that should be taken into account when choosing the qualitative evaluation approach that is most appropriate.

It discusses the value and pitfalls of qualitative evaluation methods, provides an overview of existing structured guidance and concludes with a future agenda for further work and discussion to help improve the commissioning, generating and judging of qualitative wellbeing evaluation.

What are standards and evidence, and how can these be useful in navigating different methods?

Identifying what "good" evidence looks like can be complicated. And it is often hard to know where to start: which qualitative evaluation methods should you use, when do you use them, how do you use them and how do you use them well?

A quick Google search for "books on qualitative research methods" brings up 151 million results¹. On top of the books, there are almost 1.1 million academic journal articles on qualitative evaluation methods², plus scores of discussion papers, blogs and practice guides published by research institutes, think tanks, NGOs, research consultancies and others (we list some of the best ones in annex 2 of this report). In short: it is a noisy and complicated landscape.

In response, we have explored the possibility of creating an easy to grasp framework which can be used by both specialists and evaluation novices, by large and small organisations, and by both those generating qualitative evaluation and those funders, commissioners, users, and others, who wish to understand and use it. This report explores the feasibility of such a framework, a guide to help assess, commission and generate good qualitative evaluation and evidence. This report is not the finished product, but a starting point to help stimulate discussion and debate.

1 Google search conducted on 15 March 2021.

2 Web of Science search conducted on 15 March 2021. Search term used was "qualitative evaluation methods" and generated 1,067,555 results.

A note on the methodology

This discussion paper is based upon interviews held during March and April 2021. We conducted a rapid literature review of other existing guides, frameworks and resources, interviews with ten leading organisations and individuals who have in-depth evaluation and methodological expertise, and drew upon our own personal experience of creating evidence standards, such as for Nesta³ and Pearson Plc, and mapping the wider landscape of standards of evidence used in the UK⁴.

The desk research helped us identify what guidance on qualitative evaluation methods is available and what we can learn from these resources. The focus is on wellbeing evidence but we sought relevant examples from other areas of social policy and practice, in the UK and internationally.

We also crowdsourced ideas through a "[Call for Evidence](#)" hosted on the What Works Wellbeing website. Over the three weeks this was open, we received seven responses, providing frameworks and comments for our analysis and consideration.

We conducted ten semi-structured, qualitative interviews. Our list of interviewees was developed jointly with the What Works Centre for Wellbeing, and individuals were selected based on their evaluation and methodological expertise and experience. We spoke to those applying and advancing qualitative research evaluation methods, and experts who focus more on quantitative methods, to ensure we fully understood the issues, challenges, opportunities, and the work already underway (see annex 3 for the full list of interviewees). All the interviews were conducted online via Teams or Zoom, and we took handwritten notes. As well as providing valuable insights into qualitative evaluation methods and frameworks, these interviews also helped to identify relevant literature, including some 'grey literature', unpublished or internal guidance and reports. The question guide is included in annex 3.

As with many analytical projects, we had a limited amount of time and the study was a rapid exercise conducted over less than two months. This meant we could not speak to a wider sample of individuals about qualitative evaluation methods, nor did we have scope to engage with the end users of any future guidance on how to select and use qualitative evaluation methods. However, our group of interviewees are experts and provided a rich depth of knowledge about qualitative evaluation, and quantitative evaluation, and their expertise was invaluable in shaping this paper. Going forward, wider solicitation and consultation will be essential for the creation of successful guidance, and we would recommend this is factored into any future phases of the project.

3 Puttick, R. & Ludlow, J. (2013) Standards of Evidence: an approach that balances the need for evidence with innovation. London: Nesta. DOI: https://media.nesta.org.uk/documents/standards_of_evidence.pdf [accessed 16 March 2021]

4 Puttick, R. (2018) Mapping the Standards of Evidence Used in UK Social Policy. London: Alliance for Useful Evidence and Nesta. DOI: https://media.nesta.org.uk/documents/Mapping_Standards_of_Evidence_A4UE_final.pdf [accessed 16 March 2021]

How to use this report

The paper suggests a proposed framework for thinking about quality in qualitative evaluation. It is aimed to be used in the planning, commissioning and conduct of evaluations, and for the appraisal of published evaluation studies. It also collates insights and some useful resources on good practice in qualitative evaluation, and offers realistic and pragmatic advice about how both evaluation specialists and novices, working in any field, can select, generate and commission good quality qualitative evaluation.

As well as being clear on what this paper is and who it is for, we should also state what it is not. It is not advocating for or seeking to create a hierarchy or a ladder of evidence. It is not pitching qualitative methods against quantitative methods, nor does it make judgements about the inherent value of different research methods. Neither is it a 'how-to' methods guide that talks you through the wide array of qualitative methods as there are many resources already out there which do that. And finally, what we present should not be interpreted as the final and definitive version. Like much research, the past two months have generated more questions than answers. This paper is intended to prompt discussion, and we hope it will be the starting point for much wider feedback and consultation.

How the paper is structured

- Chapter 1 outlines the need for this research and discusses the problems which can be encountered when assessing, commissioning and generating good qualitative evaluation and evidence.
- Chapter 2 provides an overview of the existing structured guidance and frameworks, and describes their merit and relevance.
- Chapter 3 draws upon the available resources, and insights from our qualitative interviews, to present a new framework for qualitative evaluation and discusses what should happen next in order to take this work forward.
- Chapter 4 provides concluding comments and a future agenda for work and discussion.
- Annex 1 provides a glossary with definitions for our key terms.
- Annex 2 provides further reading and useful resources.
- Annex 3 lists the interviewees and outlines the interview question guide.
- Annex 4 highlights the common pitfalls in qualitative research.

Chapter 1

Why this work matters

This chapter discusses the benefits and value of qualitative evaluation methods, how qualitative impact evaluation methods are used, what affects the selection and choice of qualitative methods, and the challenges and barriers to using qualitative evaluation methods.

What is qualitative evaluation and why is it useful?

Qualitative research can be an essential component of evaluation design. Qualitative evaluation methods can help move beyond indicating whether something works to helping understand why it works, when it works and how it works.

Evaluation methods draw on the same pool of methods as qualitative research so it can appear to have many similarities. But there are no specific sets of methods that makes one project an evaluation and another a research project⁵. Indeed, some applied research, as opposed to basic research, overlaps with evaluation and it can be hard to tell the difference. But there a number of different concepts and pragmatic issues that help to differentiate them (see Table 1).

Table 1: Research vs evaluation

Basic research	Evaluation
Subject area of research of interest to the researcher	Aims to solve practical questions and social problems, and subject area of interest to commissioner.
Standards set by intellectual peers	Standards more pragmatic and 'good enough' to answer the question
Researchers usually in single academic discipline	Researchers move around different policy and programme areas
Narrow group of methods used	Diverse methods
Use assessed by research peers	Use assessed by commissioners and evaluation peers

(Adapted from Kelly, 2004; Rossi et al., 2003⁶)

⁵ Kelly, M. J. (2004). Qualitative Evaluation Research. In C. Seale, G. Gobo, J. Gubrium, & D. Silverman, *Qualitative Research Practice* (pp. 462–477). SAGE Publications Ltd. <https://doi.org/10.4135/9781848608191.d38>

⁶ Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A Systematic Approach* (Seventh edition). SAGE Publications, Inc.

Key messages from interviews

In our interviews, there was universal recognition of the value which qualitative methods bring to evaluation. This includes:

- You can understand the phenomenon you are exploring to “understand the meaning and the context, the real world, peoples’ views, their framing, their conceptualisation of it, and barriers and issues from their perspective”. This means it provides a personal perspective and helps to bring in a “human voice”.
- It enables outcomes to be “unpacked” to understand the meaning they hold for service users and participants.
- It can help explain the “why and how” of a story. Qualitative methods were described by one interviewee as “helping unearth the secret sauce of the intervention”.
- The often exploratory nature of qualitative research means it can help identify unforeseen or unexpected issues or ideas. “It is extremely powerful for understanding the complexities and the sequencing and the unexpected things that might be a factor, the ripple effects of impact, and the complications of impacts”.
- It emphasises the importance of context, ethics (see example A on page 9), and being open about the role of the researcher for potential bias when interacting with the participants (although this is also relevant to quantitative research).
- It is often particularly useful when interventions are at an early stage, there is a small sample size, there is pilot testing or when it is too early for a quantitative evaluation question to be answered.
- It is essential to implementation. As one interviewee said, “qualitative methods can help understand the implementation strategy, barriers, enablers, what implementation outcomes are being prioritised and what it would look like if they were achieved”.
- It can foreground the views of beneficiaries and users of services, through workshops, interviews, observation and other methods, and help charities and providers see things differently and “bring the findings home” (see example B on page 10).

Example A

Being explicit about ethics: a realist evaluation in area of child sexual exploitation

Evaluating programmes to address sexual exploitation needs sensitive handling and raises important ethical issues. A good example of managing this was a research team at Bedfordshire University who evaluated the Barnardo's Families and Communities Against Sexual Exploitation project (FCASE), a pilot programme on preventative measures for families which was funded by the Department for Education.

In their report, published in 2015, the team clearly laid out their approaches to ethics, including approval from their university ethics committee and how they had followed ethical process outlined by The Economic and Social Research Council (ESRC) and The British Sociological Association Ethical Guidelines, such as stressing to the participants that their involvement was voluntary and that there was informed consent.

The design has been informed by a 'realist approach' to evaluation⁷, and included both qualitative and quantitative methods. A realist approach aims to identify which programme interventions work in what ways and in what conditions to produce specific sets of intended and unintended outcomes. A realist approach focuses attention on three key features: the context (such as the geography and demography, or the national and local policy context of safeguarding and child sexual exploitation); underlying mechanism (such as the direct involvement of professional staff as experts and advocates); and patterns in outcomes (such as reduction in family conflict).

This study highlights how qualitative approaches, including a sound approach to ethics, can be used to provide robust evidence of a pilot programme and evidence-based recommendations that are feasible to deliver and can be incorporated into both service / programme design and delivery.

Charities may not have the resources to have full ethical approval for evaluations, as universities and hospitals do. But they should refer to some of the ethical guidance of professional bodies, such as the [Good Practice Guidance](#) of the UK Evaluation Society, or the [Ethical Guidance](#) of the Social Research Association.

More details are available [here](#).

Examples of ethical guidance from [the SRA](#) and [Evaluation.org.uk](#)

Qualitative research is also essential for evaluating wellbeing initiatives. Such methods can help understand and explore subjective psychological experience of wellbeing. For example, an evaluation by the Centre for Ageing Better and Calouste Gulbenkian

⁷ Pawson, R., & Tilley, N. (1997). Realistic evaluation. SAGE.

Foundation used telephone interviews to understand whether psychological and emotional support for those close to retirement could really improve their wellbeing, self-kindness, and their attitudes to ending their jobs and getting older⁸. They also used quantitative data, but the evaluation benefited by talking to people and understanding how they felt and responded to the interventions. Similarly, an evaluation of Wellbeing Workshops for single mothers in South Wales, run by the Mental Health Foundation and Single Parents Wellbeing, wanted to dig down into understanding motivations for those joining the course. The evaluators found that questions of 'why' were best asked through interviews⁹.

Example B

Bringing the findings back home: listening to services-users

There can be immense value to involve service-users and other beneficiaries in your evaluation, to see things from their perspective and provide powerful personal stories to influence decision-makers.

The National Children's Bureau (NCB), with funding from the Nuffield Foundation, evaluated Independent Reviewing Officers (IROs) who represent the interests of looked after children. Their research involved national surveys of IROs, their managers and Directors of Children Services; analysis of inquiries from IROs to independent advice services; an analysis of resources needed for the IRO service; and case studies of four local authorities, including an analysis of care plans, interviews and focus groups with IROs, social workers and other key professionals. In addition, they used participatory methods to listen to young people to discuss the findings emerging towards the end of the study. There were differences between what the children thought the conclusions should be and what had been written. The researchers wrote up the debate which helped policymakers to understand the trade-offs between different policy recommendations.

The children were also involved in the presentation of results at the House of Commons. They acted out real scenarios described in the report to an audience including influential stakeholders from the Department for Education and Nuffield Foundation. The evaluation was carried out in partnership with the Centre for Child and Family Research at Loughborough University. The final report was published in April 2014.

More details [here](#) and the NCB Evaluation report is available [here](#).

⁸ Navigating later life transitions: An evaluation of emotional and psychological interventions Report 2019 by Calouste Gulbenkian Foundation UK Branch and Centre for Ageing Better <https://www.ageing-better.org.uk/publications/later-life-transitions-evaluation-emotional-psychological-interventions>

⁹ Single parents wellbeing: an evaluation of five Wellbeing Workshops in South Wales created and facilitated by Single Parents Wellbeing for single parents Mental Health Foundation, 2019 <https://www.mentalhealth.org.uk/sites/default/files/single-parents-wellbeing-evaluation-web.pdf>

Recognising the value of quantitative methods, and particularly mixed methods

Alongside recognising the usefulness and value of qualitative methods, our interviewees also stressed the importance of quantitative methods and highlighted how well the two approaches can complement each other. We heard that qualitative methods can do many things, but so can quantitative methods.

Example C

The value of mixed-methods: an evaluation of early support for military families

Military life can create additional challenges for children and families. An evaluation of the NSPCC programme to help children and families in the armed forces, such as drop-in centres and other sources of support, used a mixed-method evaluation involving two phases. The quantitative data could track outcomes, but the qualitative created rich insights. There was also a good use of phasing so that one part of the evaluation could inform another: the first phase, which informed the design of the second phase, involved desk research of routinely collected data and a series of focus groups and interviews with children, parents, staff and external stakeholders. During the second phase, they used an online survey of parents and carers using the drop-in services.

Qualitative data was analysed using a Framework Method approach (for an open paper showing how to use it in health research, see Gale et al., 2013¹⁰). The Framework Method is a systematic and flexible approach to analysing qualitative data and can be a helpful tool as it is intuitive and can be used by novices who have no experience of conducting qualitative research. In this evaluation, the Framework Method helped analyse the insights into the experiences and needs of military-connected families, including whether and how early help services strengthen and support families, and the factors that influence the achievement of positive outcomes for children.

More details are available [here](#).

We asked our interviewees if qualitative evaluation methods can address the question of impact. And what we heard is that it is hard (but not impossible) for qualitative evaluation to answer this, particularly if attribution is to be understood. There are a range of qualitative impact measures, such as process tracing or contribution analysis, to assist with thinking about impact (see Example D on page 12).

10 Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, 13(1), 117. <https://doi.org/10.1186/1471-2288-13-117>

Example D

Contribution analysis as a way of evaluating impact

The international development NGO Tearfund have used 'The Qualitative Impact Assessment Protocol (QuIP)' of their Church and Community Mobilisation project in Uganda, Sierra Leone and Bolivia between 2016 and 2018. This is an approach to impact evaluation that draws on contribution analysis which aims to assess causal questions and infer causality in real-life program evaluations. It offers a step-by-step approach designed to help managers, researchers and policymakers arrive at conclusions about the contribution their program has made, or is currently making, to particular outcomes. QuIP studies serve to provide an independent reality check of a predetermined theory of change which helps stakeholders to assess, learn from and demonstrate the social impact of their work. QuIP's approach places project beneficiaries' voices at the centre of the evaluation, enabling them to share and feedback their experiences in an open, credible, and respectful way.

The use of qualitative methods was valuable as the project had no real baseline, beneficiaries kept changing and the evaluators believed that impacts were difficult to measure quantitatively. The QuIP offered an opportunity to understand more about how people's lives, beliefs and attitudes had changed, and what had influenced any changes.

More details are available on the [Better Evaluation website](#).

But a common theme was that neither can quantitative methods fully address the impact questions. To fully understand whether something is working, why, when, how and who for, there needs to be both qualitative and quantitative methods.

As one interviewee said:

"I don't think on its own qualitative evaluation methods can get to the issue of how much impact and attribution, but those are only some aspects of impact, aren't they? But then, neither can quant on its own.

Even with the supposed gold standard RCTs, there will be issues you can't unpack, such as the meaning for participants and why you are seeing the impacts for some participants rather than others. I'd never want to rely on qualitative methods. You get hunches and hypotheses from qualitative research, but I wouldn't want to rely on it solely. But then, I'd never want to rely on quantitative solely, either."

Although we, like our interviewees, recognise the divide behind qualitative and quantitative methods as a "false binary", we cannot ignore decades of debate surrounding the value of different methods, and it would be amiss to ignore the perennial "paradigm wars" between some who use quantitative methods and cannot see the value of qualitative methods or vice versa. The Sage Handbook of Social Research Method

describes this debate as “the ‘embedded methods’ argument which depicts research methods as associated with a set of epistemological assumptions. A research method is thus a cipher for underlying philosophical ideas”¹¹. This report is not intended to get into these philosophical discussions, but recognising the debates and arguments surrounding different methods, and of particular relevance here, the devaluing of qualitative methods, is essential if any new guidance is to take root, be used and be useful. This is discussed in greater depth in chapters 3 and 4.

Limitations and barriers when using qualitative evaluation methods and when interpreting it

With our interviewees, we explored the issues and barriers which people may encounter when using qualitative evaluation methods and when interpreting research undertaken by others. These issues can be summarised as:

- A lack of understanding of what qualitative evaluation is and a lack of distinction, at times, about when research ends and “anecdote begins”.
- Linked to this, interviewees said that more is needed to guide people on how to use quotes, avoid “cherry-picking” to make certain points and ensure there is clarity as to what constitutes a representative quote.
- Good qualitative research can take time. Illustrated by one interviewee saying “people don’t understand why it takes so long to do 12 interviews. It is incredibly time-consuming, but doesn’t look like much”.
- Qualitative research methods not being taken seriously which includes it being disregarded or added into an evaluation as an afterthought, and not considered a central part of the design.
- Difficulty identifying qualitative researchers who can do qualitative evaluations, and do them well.
- Identifying what a “good” qualitative evaluation looks like. This can be overcome by using the checklists and criteria discussed in the next chapter, but it would also be avoided by researchers being clearer and more transparent about their methodological approach when they present their findings.
- Linked to the discussion above about the comparative merits of qualitative and quantitative methods, an issue to using qualitative methods is if they do not provide the data which convinces certain audiences. As one interviewee said, “a lot of charities are looking for the benefit [of their work] in pound signs, they want cost benefit analysis to understand returns. And they want quantitative evaluation for the numbers. [...] There is a fashion among funders that they want quantitative analysis, and the sad truth is, if charities focus solely on qualitative research, they may be limiting their options amongst funders”.

¹¹ Alasuutari, P., Bickman, L. & Brannen, J. (2008) Chapter 2 | The End of the Paradigm Wars? In Bryman, A. (ed.) The SAGE Handbook of Social Research Methods. London: Sage.

The role of standards of evidence frameworks

One of our initial lines of enquiry for this research is to explore whether a standards of evidence framework could be developed for qualitative evaluation. We know that standards of evidence frameworks are useful in helping both users and researchers navigate different evaluation methods. Analysis by Nesta in 2018 identified 18 standards of evidence frameworks being used in the UK, and the figure now is certainly higher, with frameworks being used across a range of different policy areas. These frameworks help judge the evidence behind policy, programmes or practice, enabling both users and generators of evidence to navigate between different approaches. Although the Nesta research shows a range of approaches and aims, each framework typically has a one to five scale and usually focuses on evidence of impact. As an example, Nesta's standards of evidence looks at the strength of the design of single evaluations or individual interventions, and is used across its grant programmes and impact investment fund, and has been adopted by the Big Lottery Fund, and Pearson Plc, the global education company¹². The range of standards of evidence has provided a useful prompt for this report, but there is a noticeable gap with no standards of evidence framework focussing on qualitative evaluation methods.

Despite the plethora of standards of evidence and other frameworks to guide the selection and use of different research and evaluation methods, it would be amiss to ignore the criticisms levelled at such frameworks. As an example, Freeman et al (2007 p. 25) argue that standards of evidence "constrain the generation of evidence rather than improve it"¹³. Much like the debate surrounding how research is generated, and the supposed merits of either qualitative or quantitative methods, there can also be controversy and debate around how knowledge is organised and curated. Our interviewees could see the value in creating a framework, but stressed that commissioners and researchers should not be reliant on them. Deeper internalised understanding, based on training and experience, of how to conduct good quality qualitative evaluation is just as important as any external framework. As one interviewee put it, any framework should not be followed "like a Bible in regimented way. It needs to be married to other kinds of support such as training and not used in narrow and regimented. Nonetheless, people coming to an area may find value from such guides". This echoes the wider literature where other researchers and practitioners have warned against dangers of checklists and 'technical fixes' (see for example Barbour, 2001¹⁴).

The next chapter provides an overview of the help and guidance which already exists on qualitative research and evaluation methods. Chapter 4 then draws upon these resources, as well as the benefits and issues identified here, to discuss the development of a new framework for qualitative evaluation.

12 Puttick, R. (2018) Mapping the Standards of Evidence Used in UK Social Policy. London: Alliance for Useful Evidence and Nesta. DOI: https://media.nesta.org.uk/documents/Mapping_Standards_of_Evidence_A4UE_final.pdf

13 Freeman, M., deMarrais, K., Preissle, J., Roulston, K., & St. Pierre, E. A. (2007). Standards of Evidence in Qualitative Research: An Incitement to Discourse. *Educational Researcher*, 36(1), 25–32. <https://doi.org/10.3102/0013189X06298009>

14 Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *BMJ*, 322(7294), 1115–1117. <https://doi.org/10.1136/bmj.322.7294.1115>

Chapter 2

Existing help and guidance on qualitative evaluation methods

Drawing on our rapid literature reviews and interviews, we have collated existing standards and other frameworks which aim to provide guidance on best practice in qualitative research and use. Whilst this list is by no means exhaustive, we have identified 15 such frameworks, toolkits, checklists and structured guides and we discuss them here.

Qualitative evaluation is not a core feature of many standards of evidence frameworks. Nesta commissioned work to map the standards of evidence used in the UK, quantitative methods were a feature of all the evidence frameworks, and 12 out of 15 explicitly consider the effect size of the impact. The use of qualitative methods was less prevalent, with only a third explicitly mentioning that it uses or considers qualitative methods¹⁵.

In summary, we define frameworks inclusively to include a range of decision-support tools to judge the commissioning, conduct or appraisal of quality relating to qualitative research and evaluation, including:

- principles
- checklists
- guidelines
- reporting standards
- critical appraisal criteria

The existing qualitative frameworks include a plethora of different items to judge quality (see Table 2 on page 16). This can range from the very basic and factual (such as asking if the study includes a title and abstract) to the more nuanced and subjective (is the study ethical? Are the authors credible?). One mapping exercise of all guidelines on qualitative research found 22 different themes across 25 different frameworks¹⁶. This mapping is being used to develop CAMELOT, the Cochrane qualitative Methodological Limitations

¹⁵ Puttick, R. (2018) Mapping the Standards of Evidence Used in UK Social Policy. London: Alliance for Useful Evidence and Nesta. DOI: https://media.nesta.org.uk/documents/Mapping_Standards_of_Evidence_A4UE_final.pdf

¹⁶ Munthe-Kaas, H. M., Glenton, C., Booth, A., Noyes, J., & Lewin, S. (2019). Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: First stage in the development of the CAMELOT tool. *BMC Medical Research Methodology*, 19(1), 113. <https://doi.org/10.1186/s12874-019-0728-6>

Tool, currently being developed for health researchers across the globe. Other reviews of guidelines have found a similar weight of items, although the wording and emphasis can be different (for example, see O'Brien et al., 2014¹⁷).

As well as positive pointers for quality and excellence, there are also guides to avoiding common mistakes such as a list of pitfalls in qualitative research and mitigation tips provided by the 'Qualres' project of the US Robert Wood Foundation (see Annex 4 on page 37). Although Qualres is aimed at academic researchers it is also a useful guide for evaluators.

Table 2: Common themes from frameworks for qualitative evaluation and research

1	Is there a title and abstract?
2	Is this a qualitative study?
3	Was there a statement of the purpose, aims and problem formulation?
4	Did the authors include/discuss a theoretical perspective?
5	Did the authors conduct a review of the literature?
6	Is a qualitative method appropriate?
7	Was the research design appropriate to address the aims of the research?
8	Were end users involved in the development of the research study?
9	Was context taken into account?
10	Sampling strategy: who are the participants, how were they selected and were the methods for selection appropriate?
11	What unit of study?
12	Was the data collected in a way that addressed the research issue? What data collection techniques and technologies were used?
13	Did the researcher spend sufficient time in the research setting?
14	Has the research team considered their role in the research process and any influence it may have on the research process or findings?
15	Have ethical issues been taken into consideration?
16	Are there clear reporting criteria (including demographic features of the study)?

Continued on the next page.

¹⁷ O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine: Journal of the Association of American Medical Colleges*, 89(9), 1245–1251. <https://doi.org/10.1097/ACM.0000000000000388>

17	Was the data analysis sufficiently rigorous?
18	Is there a clear statement of findings?
19	How valuable is the research?
20	What techniques were used to enhance trustworthiness?
21	Are the findings linked to other empirical data and other previous work?
22	Have authors discussed/assessed the overall rigor of the research study including strengths and limitations of the research?
23	Is there an audit trail?

Adapted from Munthe-Kaas et al., 2019¹⁸, and O'Brien et al., 2014, and Cohen & Crabtree, 2008¹⁹

On the next page, Table 3 shows frameworks and standards of evidence which explicitly focus upon qualitative evaluation methods and discusses their relevance to this study. The focus of this table is on structured guidance, such as frameworks, checklists and standards, that are relevant to qualitative evaluation. These are freely available online and produced as guides, reports, toolkits, webpages and open-access journal articles. The table does not include textbooks or 'how-to' methods guides on qualitative research evaluation of which there are countless examples.

What is *not* included in our list

We realise that there are many valuable resources and research published in peer-reviewed books and journals, for example the Cochrane Qualitative and Implementation Methods Group guidance series²⁰. But we have only included below the ones that are free to access. The guides included are UK-based or international and multilateral (e.g. WHO), rather than originating in other single countries outside the UK.

The list below is not exhaustive or based on a systematic review, but is representative of some key guides and frameworks relevant to quality in qual evaluation. We exclude general guidance on evaluations – and concentrate on ones that are explicitly aimed at qualitative methods. We do not include all available guidance and published recommendations on quality in medical and health qualitative research. One 2014 review found at least 40 documents and webpages, some of which can be very specific, such

18 Munthe-Kaas, H. M., Glenton, C., Booth, A., Noyes, J., & Lewin, S. (2019). Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: First stage in the development of the CAMELOT tool. *BMC Medical Research Methodology*, 19(1), 113. <https://doi.org/10.1186/s12874-019-0728-6>

19 Cohen, D. J., & Crabtree, B. F. (2008). Evaluative Criteria for Qualitative Research in Health Care: Controversies and Recommendations. *Annals of Family Medicine*, 6(4), 331–339. <https://doi.org/10.1370/afm.818>

20 Noyes, J., Booth, A., Cargo, M., Flemming, K., Garside, R., Hannes, K., Harden, A., Harris, J., Lewin, S., Pantoja, T., & Thomas, J. (2018). Cochrane Qualitative and Implementation Methods Group guidance series—paper 1: Introduction. *Journal of Clinical Epidemiology*, 97, 35–38. <https://doi.org/10.1016/j.jclinepi.2017.09.025>

as on targeting indigenous communities²¹. We do, however, provide some summaries of those recommendations, and the most relevant ones to evaluations areas outside medicine. Future work may attempt to conduct a systematic review of all available guidance, both in and outside of health – including the grey literature of reports and tools published in English.

Table 3: Other relevant guides, checklists, and frameworks on qualitative evaluation and research

1	
Name	Critical Appraisal Skills Programme (<u>CASP</u>) for qualitative research 2018
Type	Appraisal tool
Description	Relevance
CASP has a checklist of 10 questions to help you make sense of qualitative research. The authors have created it to be an 'educational pedagogic tool' sitting alongside training.	Despite its origins in health research, it is relevant for many areas. It has been going for over 25 years and is part of a range of other guides including on RCTs and systematic reviews. It has a free template to use.
2	
Name	Quality in qualitative evaluation: a framework for assessing research evidence (<u>supplementary Magenta Book guidance</u>) HM Treasury 2012
Type	Guidance / framework
Description	Relevance
Provides a framework with 18 appraisal questions based on a review of literature and interviews There are 10 categories of the 18 questions (e.g. design, sampling) and ends with some general features of research conduct (reflexivity/neutrality, ethics and auditability).	Aimed at government and policymakers, not NGOs, but potentially relevant to other evaluators and social researchers. More background information provided by work by NatCen and Cabinet Office ²³ .

Continued on next page.

21 O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine: Journal of the Association of American Medical Colleges*, 89(9), 1245–1251. <https://doi.org/10.1097/ACM.0000000000000388>

22 Spencer L, Ritchie J, Lewis J et al. (2003) *Quality in qualitative evaluation: a framework for assessing research evidence*. London: Government Chief Social Researcher's Office

3	
Name	Listen and Learn: How charities can use qualitative research , NPC 2016
Type	Guidance / report
Description	Relevance
Practical guidance includes 5 principles for good qualitative research: validity; reliable; confirmable; reflexive; responsible – and a simple checklist and table with examples of 5 areas of quality, such as detailed methods, appropriate sampling, stories to reflect views, strong analysis, and 'reports designed to persuade people for the need for a service' (p.16)	Highly relevant to charities and with practical and realistic advice for organisations, both large and small, including how to conduct evaluations internally, and matching of quant and qual methods. Has a framework but more for WHEN to use qual. This guidance is adapted from NPC's Inspiring Impact's cycle of impact practice – 'plan, do, assess, review' to indicate how qualitative data can contribute in different ways at each stage.
4	
Name	MRS Guidelines for Qualitative Research , Market Research Society 2011
Type	Guidance
Description	Relevance
Best practice guidance that covers qual research, ethnography, and deliberative work. A strong focus on good conduct, ethics and professionalism.	Although aimed at members of the MRS and market researchers, this is relevant to anybody working in NGOs or elsewhere – including those interested in their own professionalism and techniques for more user-engagement and deliberation It is more about high-level standards and principles, rather than a 'how to' guide to vet quality.
5	
Name	Guidelines for Good Practice in Evaluation , UK Evaluation Society 2019
Type	Guidance
Description	Relevance
The guidelines provide 8 overarching principles such as trust, transparency and equity. With a strong focus on integrity and ethics, it targets messages for different audiences, including commissioners, evaluators, institutions and participants.	Relevant to any sector, these principles do not specifically cover qual evaluation – it covers all approaches –but is applicable across many areas. Includes useful references on guidelines on ethics from international bodies and learned societies.

Continued on next page.

6	
Name	<u>Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups</u> ²⁴
Type	Reporting checklist
Description	Relevance
Provides checklist of 32 items. Based on synthesis of 22 other guides, and a total list of 76 items. All items were grouped into three domains: research team and reflexivity; study design; data analysis and reporting.	Checklist aimed at two common methods in qualitative studies: interviews and focus groups. Published in health journal but relevant in other domains, the review included non-health guides, such as the Campbell Collaboration. Useful and free <u>checklist published online</u> .
7	
Name	SRQR: Standards for Reporting Qualitative Research; A Synthesis of Recommendations ²⁵
Type	Reporting standards
Description	Relevance
Synthesis of other recommendations on quality provides 21 items. Based on a synthesis of over 40 existing guidelines, studies, and other advice on qualitative research and evaluation.	Health focused, and includes all qual research, not just evaluations, and relevant only up to 2013. But useful for its systematic synthesis for existing advice, guidance and research on quality in qualitative research. Provides a useful and free table with 21 items for Standards for Reporting Qualitative Research (SRQR).
8	
Name	<u>Using Qualitative Methods in Healthcare Research; A Comprehensive Guide for Designing, Writing, Reviewing and Reporting Qualitative Research</u> , Robert Wood Johnson Foundation, 2008
Type	Guidance
Description	Relevance
A website with resources that will be useful for people developing, evaluating and engaging in qualitative research projects.	This is more focused on social science research and on healthcare settings. However, there are valuable, freely available resources and guidance - including tips on how to avoid common pitfalls in qual research such as unclear focus, insufficient sample, or being 'too jargon'.

Continued on next page.

²³ Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care*, 19(6), 349–357. <https://doi.org/10.1093/intqhc/mzm042>

²⁴ O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine: Journal of the Association of American Medical Colleges*, 89(9), 1245–1251. <https://doi.org/10.1097/ACM.0000000000000388>

9	
Name	Appendix H: Quality appraisal checklist – qualitative studies, NICE, 2012
Type	Checklist
Description	Relevance
A list of 14 areas to examine in qual research including more advanced technical items whether the data is 'rich', or the use of second researcher to code.	Aimed at public health research, some of the requirements are quite advanced. It is a clear checklist and useful beyond public health. Part of Methods for the development of NICE public health guidance (third edition).
10	
Name	Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 2: how to make an overall CERQual assessment of confidence and create a Summary of Qualitative Findings table (Lewin et al., 2018 ²⁶)
Type	Guidance
Description	Relevance
Freely available paper provides practical help and a table for judging the confidence in qual research, based on four areas of consideration: the coherence; relevance, methodological limitations, and the adequacy of data.	This is part of an influential guide for the synthesis of multiple qualitative research, and it is still a helpful resource for thinking about single studies. It builds on other guidance and work for GRADE-CERQual. Mostly used in health research but relevant in other areas.
11	
Name	Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: first stage in the development of the CAMELOT tool ²⁷
Type	Map of tools
Description	Relevance
A review of existing tools for the critical appraisal of methodological quality. This mapping will inform the forthcoming CAMELOT tool. They identified 102 critical appraisal tools: and developed a framework that included 22 themes.	Designed for health research and aimed at helping the use of CERQUAL-GRADE, the mapping has cast its net widely using social science databases. The forthcoming CAMELOT tool may become the go-to guide for looking at methodological limitations. The 22 themes in the review are a useful and up-to-date list of criteria in other methodological appraisal tools.

Continued on next page.

25 Lewin, S., Bohren, M., Rashidian, A., Munthe-Kaas, H., Glenton, C., Colvin, C. J., Garside, R., Noyes, J., Booth, A., Tunçalp, O., Wainwright, M., Flottorp, S., Tucker, J. D., & Carlsen, B. (2018). Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 2: How to make an overall CERQual assessment of confidence and create a Summary of Qualitative Findings table. *Implementation Science*, 13(1), 11–23. <https://doi.org/10.1186/s13012-017-0689-2>

26 Munthe-Kaas, H. M., Glenton, C., Booth, A., Noyes, J., & Lewin, S. (2019). Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: First stage in the development of the CAMELOT tool. *BMC Medical Research Methodology*, 19(1), 113. <https://doi.org/10.1186/s12874-019-0728-6>

12	
Name	<u>WEIRD (Ways of Evaluating Important and Relevant Data) tool</u> ; Questions to guide assessment / critical appraisal of programme descriptions, implementation descriptions and other mainly descriptive types of evidence - Cochrane Effective Practice and Organisation of Care
Type	Methods tool
Description	Relevance
WEIRD tool for assessing the methodological limitations of non-empirical evidence sources looks at the 'real world' information covering more descriptive views of programmes, system reforms, policies, and implementation.	Although this tool is a work in progress and has not yet been formally launched, it provides some helpful guidance on how to appraise the quality of diverse and sometimes quite informal sources of evidence – rather than formal impact evaluations – which can be relevant when formal evaluation or other research is not available. Relevant to health, welfare and other areas – including the contexts around specific areas.
13	
Name	<u>Using qualitative research to strengthen guideline development</u> , World Health Organisation, 2019
Type	Guidance
Description	Relevance
A mini-series of 3 guides to look at how to incorporate qualitative data into guideline development and implementation, specifically qualitative evidence syntheses (QES), some of which have been published in the Cochrane Library.	Focused on health guidance and reviews of qual evidence, not single studies, but an accessible guide for anybody making products, like professional guidance, decision-aid tools, or implementation of evidence-informed practice.
14	
Name	<u>Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: guidance for researchers</u> ²⁸
Type	Map of tools
Description	Relevance
Guidance for those conducting randomised controlled trials on how to include qualitative evidence for feasibility studies. Qualitative research can be used to examine and address key uncertainties prior to a full trial. The guidance consists of 16 items within five domains: research questions, data collection, analysis, teamwork and reporting.	Relevance for those planning, commissioning or thinking about the use of qual research to aid experimental evaluations. The guidance is also useful for those looking at feasibility and early-stage scoping of evaluations, including answering the question 'can the evaluation be done?'. Focus mostly on medicine and health but applicable to other areas.

Continued on next page.

²⁷ O’Cathain, A., Hoddinott, P., Lewin, S., Thomas, K. J., Young, B., Adamson, J., Jansen, Y. JFM., Mills, N., Moore, G., & Donovan, J. L. (2015). Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: Guidance for researchers. *Pilot and Feasibility Studies*, 1(1), 32. <https://doi.org/10.1186/s40814-015-0026-y>

15	
Name	Qualitative process evaluation from a complex systems perspective: A systematic review and framework for public health evaluators (McGill et al., 2020 ²⁹)
Type	Methods tool
Description	Relevance
A framework produced by researchers at the London School of Hygiene and Tropical Medicine. Based on a systematic review of 21 process evaluations that use qualitative methods in a complex systems perspective, they recommend a 2-phase framework: phase 1 involves producing a description of the system and identifying hypotheses; phase 2 involves following the 'pathway of emergent findings in an adaptive evaluation approach'.	Aimed at public health evaluation, but relevant to other areas interested in a complex systems. Shows how little consensus there is in qualitative process evaluations and complex systems.

28 McGill, E., Marks, D., Er, V., Penney, T., Petticrew, M., & Egan, M. (2020). Qualitative process evaluation from a complex systems perspective: A systematic review and framework for public health evaluators. *PLOS Medicine*, 17(11), e1003368. <https://doi.org/10.1371/journal.pmed.1003368>

Chapter 3

Developing a new framework

This chapter draws upon the issues and challenges encountered when navigating qualitative evaluation methods discussed in Chapter 2, and the existing frameworks identified in Chapter 3, to suggest a new framework to improve the quality of both the generation and use of qualitative evaluation.

Developing a framework for qualitative evaluation

From our research, there are three potential uses of qualitative research within the evaluation process. Firstly, qualitative methods can help determine the focus of the evaluation. Second, they can help understand the process and implementation surrounding a programme. Third, qualitative methods can also help determine improvements and alterations to a programme.

As a starter for discussion, we propose that there are six 'prompts' to qualitative evaluation. These six areas are intended to guide those commissioning, conducting or critically appraising qualitative evaluation. The six are based on our reading of the 15 other guides discussed in Chapter 2, and attempt to condense and consolidate common themes. However, they are not based on any detailed conceptual or empirical analysis of other guides; they are primarily based on our own judgement and background in creating other evidence frameworks, and from our interviewees who stressed the need for practical usage, rather than long complicated checklists. So, we have attempted to keep the framework simple to help evaluators. We welcome other expert views on their form and content so we can adapt and iterate them following feedback from others.

They are:



Table 4: The Six Quality on Qual Prompts: questions and considerations

Prompt	The expectation	Questions and considerations
1	<p>Aims and questions:</p> <p>Are there clearly identified aims and evaluation questions. Have they shown awareness of other relevant studies?</p>	<p>Inclusion of clear evaluation goals.</p> <p>Discussion of why study is thought to be important and relevant.</p> <p>Reference to its contribution to wider knowledge and research.</p>
2	<p>Methodology & design:</p> <p>Have you included a transparent approach to your methods, including justifying why a qualitative design is relevant to answer your questions?</p>	<p>Description of methods, methodology and design, and how they match the evaluation question and research strategy.</p> <p>Justification of the use of a qualitative approach.</p> <p>Inclusion of what form of data used (e.g. transcripts for interviews, focus groups).</p> <p>Discussion of limitations of research design.</p>
3	<p>Sample & recruitment:</p> <p>Have you provided a clear plan for how participants will be selected and recruited? Is there an explanation for why some people have been chosen?</p>	<p>Defence of the choice of sample design and target sources of evidence?</p> <p>Explanation of how the participants were or will be selected.</p> <p>Discussion around recruitment, including why some people chose not to take part.</p>
4	<p>Ethics & relationship with evaluator:</p> <p>Have you set out how you will manage confidentiality and informed consent? Has the evaluator considered their own relationship with the participants and personal biases?</p>	<p>Consideration given to ethical issues including sensitivity about contexts and potential harm to participants.</p> <p>Participants provide informed consent which is (or will be) documented?</p> <p>How to deal with participant distress, wellbeing, concerns regarding safeguarding etc.</p> <p>Discussion of anonymity, confidentiality of data and procedures for protecting it,</p> <p>Examine the evaluators' own role, potential bias and influence throughout the study.</p>

Continued on the next page.

Prompt	The expectation	Questions and considerations
5	<p>Analysis & reporting:</p> <p>Is there an in-depth and nuanced description of the analysis process?</p>	<p>Discussion of how/why particular significance is given to specific aspects of the evidence.</p> <p>Communication of the nuance, depth of the complexity and richness of the data?</p> <p>Analysis of different alternative, contradictory or diverse views.</p> <p>Analysis of the relationship and potential bias of researcher.</p> <p>Mention of any analytical concepts or links to other concepts and theories?</p>
6	<p>Findings & conclusions:</p> <p>Are the claims credible, not 'over-claiming' but offering plausible arguments that can be traced to the methods and evidence generated? The reader can see how the researcher arrived at his/her conclusions.</p>	<p>Presentation of findings clear and explicit, and related to the research aims and questions.</p> <p>Discussion of the evidence both for and against the evaluators' arguments and findings/conclusions.</p> <p>Plausible arguments that are not 'over-claiming' but show link between data, interpretation and conclusions – so the reader can see the route to any conclusions be seen.</p> <p>Findings linked with other knowledge and experience, from other data sources, wider research literature, or from peer review.</p> <p>Discussion of the credibility of findings, such as corroborated with other research, input from stakeholders, or use of more than one evaluator.</p>

This list could have been much longer. A longer list of other questions, sub-questions and themes might be included. Indeed, the HM Treasury framework has over 18 appraisal questions, CASP has 10 and the forthcoming CAMELOT tool is based on 22 themes (see Table 3 on pages 18-23 for an overview of other frameworks). Furthermore, HM Treasury guidance (2012) recommends being explicit about the evaluators' ideology and theoretical background. This is commendable and vital for much social science inquiry, but possibly unrealistic for our audiences. As a result, we have aimed to be more succinct and user-friendly. We wanted to keep our framework simple and user-friendly.

Before expanding the framework, we would like to test the framework with the stakeholders of What Works Centre for Wellbeing; explore whether these six prompts

to qualitative evaluation are enough and cover the right areas. The six prompts provide a minimum quality threshold on which to build on. Brevity, we believe, will benefit those coming to these prompts; some of our interviewees expressed doubts about the utility of longer lists. Shorter is more memorable and pragmatic. We would want to test and iterate this smaller list with the Centre's stakeholders. We can provide more detail for those that seek it, such as the guidance, checklists and frameworks set out in Table 3.

In addition to these prompts, it is instructive that other organisations and scholars have recommended some overarching principles or high-level aspirations, such as those recommended by the 2012 HM Treasury framework, including defensibility of the design, rigour, transparency, credibility and contribution to wider knowledge (2012, p11). For example, one abiding principle in most frameworks is the importance of being transparent and explicit about your approach, including any potential limitations of your work (see Example E below). However, these are somewhat abstract terms and we have not included principles in our framework here, but instead focused on practical prompts. We can test the absence of these principles with users during any future work.

Example E Being open about limitations is a strength

Green Gyms, promoting natural exercise for community health, was evaluated by New Economics Foundation (NEF) on its role in maintaining physical and mental health. The evaluators used a short questionnaire survey which was collected once a month for three months, between 2015 and 2016, and a series of qualitative interviews with both volunteers and referral partners. The final evaluation report is written in clear, jargon-free language and is explicit about its methodology. This includes valuable details for evaluators such as use of QDA Miner for analysis, the use of a thematic and framework approach to identify themes and patterns in the data, and the inclusion of the full questionnaires and interview topic guides, for example.

As well as clear evaluation questions, and a theory of change, it is very transparent about its shortcomings, like the drop out of participants. The evaluation found that Green Gym is effective in both improving and maintaining physical and mental health. Key mechanisms include increasing physical activity levels and reducing social isolation by creating opportunities for people to work together in worthwhile outdoor activities.

More information can be found [here](#).

Another point of explanation: we have tried to avoid technical jargon, such as not including 'reflexivity' in Prompt 6 as this concept, whilst important, is not widely known outside the social science community. We can also clearly see these six stages being made into a simple template to help nudge evaluators to consider these areas.

Finally, we were tempted to put these six into an acronym that might be mnemonic. But

there is already an alphabet soup of other frameworks on qual research, such as SRQR, COREQ, CERqual, CASP and, one of our favourite names, WEIRD (see Chapter 2 of this report). Adding another might just add to the confusion.

Table 4 provides more detail about each stage, and lists some potential evaluation methods which could be used. This list of methods is not exhaustive, but is intended to help prompt thinking and debate.

A few caveats

We hope this framework helps to prompt discussion about what qualitative evaluation standards of evidence could look like. There some health warnings and caveats we would like to emphasise:

- Although the amount of evidence and knowledge will grow, the process is not linear. New insights, questions and issues will arise which will require new research and exploration.
- There is no 'end point' to evaluation. Finding 'what works' is only useful if we ensure it is still working²⁹.
- No method, qualitative or quantitative, is inherently superior. Effective, high quality evaluation comes from using the best method to suit the question. Method selection is about appropriateness to get different insights and learnings.

There is a balance to be struck in presenting this guidance about being prescriptive and helpful, but also considering the huge array of evaluation questions to be answered, the size and scale of an organisation, their existing experience of evaluation and the resources they have available. For this reason, it's not appropriate to be overly directive, and, for example, dictate how many interviews they should do. Guidance can only get people and organisations so far and there is an overarching need for individuals to interpret and apply it for their own context and purpose.

Furthermore, there is no getting away from the fact that we are all convinced by different forms of evidence, at different times, communicated in certain ways. This means that any single standard of evidence may not be deemed the correct one by everyone, at all times. Linked to this, the very curation of evidence and methods, is deemed controversial by some, which may mean people criticise the very suggestion of a framework, regardless of what it contains.

Before consultation of this framework is discussed, we want to highlight an issue central to this research project which is about quality of qualitative evaluation. As we have said before, no method is inherently superior and any method, design or data collection can be done badly. An evidence framework can help both users and researchers to navigate and select the most appropriate methods, but this does not guarantee quality in their execution. The final framework will need to clearly link to quality guidance which may be

²⁹ Puttick, R. (2012) We don't need what works. We need to know what is working. London: Nesta. DOI: <https://www.alliance4usefulevidence.org/we-dont-need-what-works-we-need-to-know-what-is-working/>

A framework to commission, judge and generate qualitative evaluation in wellbeing impacts

in the form of the checklists identified in Chapter 2.

The next step is for these six prompts to be shared and discussed to understand their resonance, usefulness and usability. This is discussed in the next chapter.

Chapter 4

Discussion, next steps and recommendations

This final chapter reflects on the research to discuss how to move forward with this work and achieve the central aim of improving the quality, usability and usefulness of qualitative evaluation methods.

Is there a need for a new framework for qualitative evaluation methods?

At the outset of this project, we anticipated being able to develop a new set of standards of evidence for qualitative evaluation, potentially comparable to the Nesta Standards of Evidence framework³⁰. From our desk and primary research, we have solicited input into this gap and found that more work is needed in order to develop a comprehensive new framework or a single set standards of evidence for qualitative evaluation. What we have proposed in Chapter 3 is therefore a preliminary, untested framework, intended to act as a prompt for discussion about what a framework could and should do, and indeed, we would ask the question: is one going to be enough?

Linked to this question about the need to further explore the scope of a qualitative standards of evidence framework, across our interviewees there was a recognition of the value of mixed methods. Our analysis shows that people value qualitative evaluation methods and recognise the mutual complementarity when used with quantitative methods.

This raises the following questions:

- Should there be new standards of evidence for qualitative methods?
- Should adaptations be made to existing frameworks to give qualitative methods a more central role?
- And/or should there be a new Standards of Evidence for Mixed Methods, recognising the value of both qualitative and quantitative methods to answer different questions at different times in the innovation and evaluation process?

30 Puttick, R. & Ludlow, J. (2013) Standards of Evidence: an approach that balances the need for evidence with innovation. London: Nesta. DOI: https://media.nesta.org.uk/documents/standards_of_evidence.pdf [accessed 16 March 2021]

Explore what scale any new framework or standards of evidence should be graded on to help establish what the “appropriate” evaluation method should be

Both in the literature and in our interviews, we often heard the importance of ensuring that the evaluation method is “appropriate”. More is needed to understand what appropriate means for different projects and organisations within the wellbeing field.

As an example, the Nesta Standards of Evidence provides guidance on how to evaluate different interventions, products and services. In that framework, the appropriateness of the evaluation method determines the intervention’s stage in the innovation process.

Aligning methods to the innovation process may not be appropriate for the What Works Wellbeing Centre or for the organisations it works with. Instead, the appropriate method may be determined by the evaluation question, budget, scale and wider context. Greater clarity is needed to determine what these factors are and to help organisations select the best suited evaluation method.

Improve the guidance on the value of qualitative evaluation, and how it can be used

In addition to a framework which would help both users and producers navigate qualitative methods, there is also a clear need for better guidance on how to use and apply these methods. As chapter 3 demonstrates, there is no shortage of other frameworks and resources on methods, but the volume of advice can be overwhelming and much of it is based in health research.

Our interviews identified the need for a navigation tool, potentially akin to the Alliance for Useful Evidence’s Experimenters Inventory, which would provide more detail on qualitative evaluation methods, their pros of cons and when to use them.

Recommendations and next steps

We dare not conclude this study by saying that “more research is needed”. However, it certainly is if this thorny issue of improving the usefulness and value of qualitative evaluation methods is to be tackled.

But what we are proposing is not simply more research to generate more guidance, but developing this output or set of outputs in a completely different way.

Over the past two months we have synthesised the available resources and we have engaged with those with deep methodological expertise, to propose a framework of qualitative methods and to pose questions for discussion. What is needed now is to test our framework out with those we hope would find value in it. Does it give them what they want? Is it useful and usable?

Too often guidance is produced with the end user in mind, but not the end user involved, and that is a situation we want to avoid.

As interim next steps, we propose the following activities, all of which should be imbued with human-centred design and user-led methods:

- Thorough interrogation of the proposed framework, including workshops and focus groups to help challenge, expand and populate the different stages with appropriate methods and tools. Once we are happy with the proposed framework, publish it. Put it out into the public domain and let people critique it, challenge it and, better yet, use it.
- And to help facilitate our understanding of usefulness, we propose supporting a range of organisations, big and small – those experienced in research and those at an earlier stage – to make sense of it and apply it. An evaluation of this exercise would be essential to help understand the user experience.

As longer term next steps, alongside developing the proposed framework, we also recommend:

- Deeper understanding of what is meant by an **“appropriate evaluation method”**, which could be driven by the evaluation questions, context, scale, either of the origination and/or of its work, budget, timescale or other factors. Understanding what these factors are is crucial in developing guidance to help organisations navigate and select the correct evaluation method for their needs.
- Exploring the creation of a **mixed-methods standards of evidence**, as detailed above, **which could recognise the value of both qualitative and quantitative methods to answer different questions at different times in the evaluation process.**

Prompts for further discussion and debate

We hope that this discussion paper is just the starting point. We have been privileged to have already engaged with ten interviewees who have rich and diverse evaluation and methodological experience. We now want to engage more people. To help facilitate this, alongside this discussion paper, we will produce a blog post, a reference guide which draws upon the resources discussed here, and a short note which sets out the six-prompt framework. All of these outputs are intended to help stimulate discussion and debate.

Please share your thoughts and feedback by emailing info@whatworkswellbeing.org

We look forward to hearing from you.

Annex 1

Glossary of key terms

This section defines the key terms used in this report. Rather than generate new definitions, we have used definitions used by the What Works Centre for Wellbeing. To supplement the glossary, we have looked to the literature and used those published by Springer within their Glossary of Qualitative Research.

Case study

Case study is a research approach. It creates an in-depth description and analysis of a bounded system – one individual, institution, or educational context. By concentrating on a single (or few) case(s), this approach can describe a particular learning, teaching process or research setting in great detail. The focus of a case study can either be on the case itself or on an issue which the case(s) illustrates. Case study uses multiple sources of data and data collection methods, and it often combines both qualitative and quantitative research approaches³¹.

Data

Data is information created in or collected from a setting (from participants, observations, artifacts, and so on, about people, phenomena, ideas, events, or the context) in response to a research question³².

Data analysis

Analysis is a major aspect of the research process. There are many ways to do data analysis in qualitative research, but the goal is largely the same – to understand the data that has been collected, and organize it into groups or categories, to prepare for data interpretation³³.

Data collection

Data collection refers to the process of collecting information systematically through data collection methods³⁴.

Data collection methods

Data collection methods in qualitative research include observation, interviews, open-response questionnaire items, verbal reports, diaries, and so on³⁵.

31 As defined in the Glossary of Qualitative Research Terms. Springer Publishing. DOI: <https://link.springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf>

32 As defined in the Glossary of Qualitative Research Terms. Springer Publishing. DOI: <https://link.springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf>

33 As defined in the Glossary of Qualitative Research Terms. Springer Publishing. DOI: <https://link.springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf>

34 As defined in the Glossary of Qualitative Research Terms. Springer Publishing. DOI: <https://link.springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf>

35 As defined in the Glossary of Qualitative Research Terms. Springer Publishing. DOI: <https://link.springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf>

Evaluation

Evaluation is the process we use to find out how well an activity or intervention has worked. It's about gathering evidence to help us judge the value and success of an activity, so that we can learn and make improvements for the future³⁶.

Evaluation in Wellbeing

It's about looking at your activities through a wellbeing lens – and gathering evidence of the changes in wellbeing that may have occurred as a result of your work³⁷.

Mixed-methods

Research design in quantitative and qualitative methods are used in a single study to gather and analyse data in order to answer a research question(s)³⁸.

Qualitative analysis

Making analytic generalisations based observation, dialogue and evocation, rather than direct measurement³⁹.

Quantitative analysis

Representing data in terms of numerical values and involves direct measurement⁴⁰.

Qualitative evaluation

Qualitative data is information which comes from observation, description or other non-numerical sources. Words are the most common form of qualitative data, which, unlike quantitative data can't be counted. It comprises people's opinions, stories and views expressed in their own words. Qualitative data includes what people tell you in interviews or focus groups, diaries, videos or photographs, and through other means⁴¹.

[springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf](https://www.springer.com/content/pdf/bbm%3A978-0-230-23951-7%2F1.pdf)

36 As defined by What Works Wellbeing. Available online: <https://measure.whatworkswellbeing.org/intro-to-wellbeing-evaluation/>

37 As defined by What Works Wellbeing. Available online: <https://measure.whatworkswellbeing.org/intro-to-wellbeing-evaluation/>

38 As defined by What Works Wellbeing. Available online: <https://measure.whatworkswellbeing.org/glossary/>

39 As defined by What Works Wellbeing. Available online: <https://measure.whatworkswellbeing.org/glossary/>

40 As defined by What Works Wellbeing. Available online: <https://measure.whatworkswellbeing.org/glossary/>

41 As defined by What Works Wellbeing. Available online: <https://measure.whatworkswellbeing.org/intro-to-wellbeing-evaluation/>

Annex 2

Further reading and useful resources

Guidance	Organisation/author	Summary	Location
How to Collect Qualitative Data the Easy Way	NCVO	A four-step guide to collecting qualitative data, including further information about using Facebook for qualitative research, ethics in research, how to run focus groups, and how to make sense of qualitative data.	Link
Measure your wellbeing impact: A practical guide for charities and social enterprises	What Works Wellbeing	A seven-stage guide with details on what is wellbeing, planning your evaluation, conducting research, and analysing results. There is also a "Wellbeing Measures Bank".	Link
Happy City Policy & Measurement: Wellbeing Measurement A Guide to Qualitative Data Collection	Centre for Thriving Places	The guide details ways in which you can collect qualitative wellbeing data, understand how qualitative data can complement quantitative wellbeing data, and help improve how to gather and analyse good quality wellbeing data.	Link
A Guide to our Evidence Review Methods	What Works Wellbeing	A detailed guide about Evidence Reviews, with information about measuring wellbeing outcomes, and the system for ranking the strength and quality of the evidence of what works to improve wellbeing, as well as what doesn't work.	Link
Glossary of Qualitative Research Terms	Springer (pub.)	A comprehensive list of definitions for qualitative research terms.	Link
Qualitative Research Methods in Program Evaluation: Considerations for Federal Staff	U.S. Department of Health & Human Services	A detailed primer on qualitative evaluation methods.	Link
Understanding and measuring outcomes: the role of qualitative data	Emma Miller and Ellen Daly	A comprehensive guide on qualitative data collection.	Link

Annex 3

List of interviewees and interview question guide

List of interviewees

1. Chris Carmona, National Institute for Health and Care Excellence (NICE).
2. Emilie Smeaton, The National Lottery Community Fund.
3. Jane Lewis, Centre for Evidence and Implementation.
4. Jen Gold, Cabinet Office.
5. Jennifer Wallace, Carnegie UK Trust.
6. Jon Franklin, Pro Bono Economics.
7. Mary McKaskill, Centre for Youth Impact.
8. Rosie McLeod, New Philanthropy Capital.
9. Simon Lewin, Norwegian Institute of Public Health.
10. Siobhan Campbell, Government Social Research Service.

Interview questions

The interviews took the form of a 45 to 60-minute call via Zoom or Teams. The interviews were semi-structured and the questions below acted as an aide-memoire, therefore, were not read verbatim or in-order, and prompts and probes were used throughout.

1. How do you use qualitative evaluation methods in your work, what types of qualitative impact methods do you use, and what affects the selection and choice of qualitative methods?
2. What are the benefits and value of qualitative impact evaluation?
3. What are the challenges and barriers to using qualitative evaluation methods?
4. We hypothesise that it is difficult for researchers to navigate and select appropriate qualitative evaluation methods to test well-being interventions and programmes. Do you agree?
5. Do you know of any guidelines, frameworks, standards of evidence, or other resources, which can help individuals select appropriate qualitative evaluation methods?
6. Do you know of any guidelines, frameworks, standards of evidence, which help interpret and assess the quality of a qualitative evaluation?
7. Where are the gaps in provision?
8. Is there a need for additional frameworks and guidance for qualitative evaluations?

Annex 4

Common Pitfalls in Qualitative Research

The US Robert Wood Foundation has provided a helpful table of common mistakes for their [Qualitative Research Guidelines Project website](#). They looked at rejected manuscripts submitted to a health research journal and examined reviewers comments listed some of the errors made – and links to resources to help avoid those pitfalls. Some of the mistakes are more related to qualitative social science research (such as the mistake of being ‘too jargony’) – but is applicable to evaluators. The table is reproduced below. The review of literature and some analysis is produced in a subsequent academic journal by one of the lead researchers (Cohen & Crabtree, 2008).

Table 5: Common pitfalls in qual research

Common pitfalls	Examples from reviewers' comments	Link to topic on Robert Wood Foundation website 'qualres'
Lacks focus	Unclear focus	Guidelines for Conducting, Analyzing and Reporting Qualitative Data
Too jargony	Use of unfamiliar terminology without providing adequate definition	Guidelines for Conducting, Analyzing and Reporting Qualitative Data
Sample insufficient	Lack of iterative sampling process/saturation Inadequate description of the sample	Sampling Issues Iterative Sampling Common Research Methods
Analysis lacks depth	Superficial Analysis Data don't support the results Analytical Categories are unclear Coding decisions are not clear	Common Analytic Approaches See also Miles and Huberman Qualitative Data Analysis
Methods lack adequate description	Each step of the sampling, data collection, and analysis is not described in sufficient detail	Evaluative Criteria Audit Trail Malterud's article in The Lancet is very useful
Data Quality Concerns	Imposes categories or ways of seeing on participants that may be irrelevant for them Evidence that wrong method was used Evidence that method was implemented suboptimally	Common Methods Constant Comparative Method Coding Data

(Robert Wood Foundation, 2008)