Six principles to strengthen qualitative assessments in development interventions
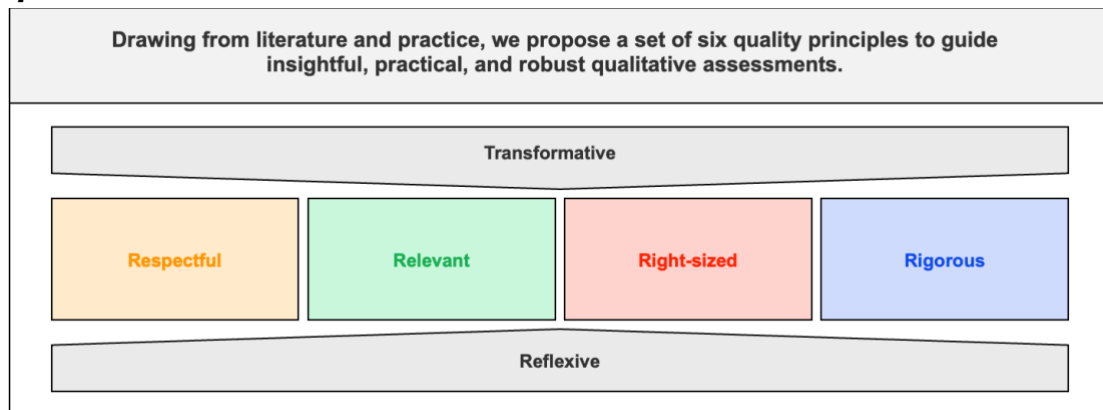
- MacArthur, Jess (corresponding author)
- Abdel Sattar, Rana
- Carrard, Naomi
- Kozole, Tyler
- Nicoletti, Chris
- Riley, Lauren
- Roglà, Jennifer
- Toeur, Veasna
- Willetts, Juliet

(Apart from the first author, other authors are listed in alphabetical order)

Abstract

Qualitative research and evaluation is often described as an art rather than a science. This intangible nature has left many programme teams feeling uncertain about how to justify and use qualitative forms of assessment in their monitoring, evaluation and learning practice. Existing guidance is theoretically based and often focused on applications in the global north. Building on an expressed need for guidance from a community of practice in Asia and the Pacific, this collaborative action research process aimed to create practical and tested guidance for program teams. The analysis was conducted as part of a Cambodia-based sanitation program assessment. Drawing from literature and the collaborative process, we propose a set of six principles to guide insightful, practical, and robust qualitative assessments. We provide examples regarding how the principles can be used to plan, conduct, and review qualitative assessments with a goal to strengthen future use of qualitative tools in programming.

Graphical Abstract

Drawing from literature and practice, we propose a set of six quality principles to guide insightful, practical, and robust qualitative assessments.

Transformative

| Respectful | Relevant | Right-sized | Rigorous |

Reflexive

Video Abstract

Behind the scenes: **Collaborative action research** | Developing principles to strengthen qualitative assessments

iDE ❖ UTS Institute for Sustainable Futures

Introduction

Within monitoring, evaluation, and learning (MEL), practitioners are often faced with difficulties in designing and implementing insightful, practical, and robust assessments, especially within the context of qualitative inquiries (Lincoln and Guba, 1985; Patton, 2002). Additionally, what counts as 'rigorous' is often difficult to contextualise in practice (Mayoux & Chambers, 2005). We use the word 'assessment' to reflect the breadth of research and analysis practices that review or evaluate program opportunities and outcomes.

Qualitative assessments are often criticised as not robust, and program teams report low confidence in collecting and analysing qualitative data (Patton, 2002). This insight comes not only from literature but also from the practice of the authors of this article. In many cases, evaluation teams adopt approaches that they are already familiar with instead of exploring new options and selecting the approach most appropriate for the research question. This can limit opportunities to hear from new voices, triangulate results, and gain a depth of insight. In other cases, teams face challenges in effectively adapting approaches to unique contexts and restrictions. These adaptations often feature trade-offs of rigour versus time, breadth versus depth, and utility versus independence. Lastly, few teams take time to conduct post-evaluation reviews of their selected approaches' efficacy and learn from their experiences (Mark, 2008).

There is a wide range of practical guidance available for evaluators broadly and in particular on qualitative methods. However, this advice is 1) primarily focused on northern applications, 2) often focused on evaluator competency, 3) designed for a particular type of evaluation, and 4) tailored for use in a specific sector such as education. This guidance is often theory-driven rather than grounded in empirical practice-based evidence. Little has been written about qualitative evaluation practice or the international development sector, where evaluations are often embedded into program implementation and conducted under resourcing constraints (staffing, time, and budgets) and differing conceptualisations of rigour (Bamberger, 2000). Building on these gaps and in response to an expressed need from a southern-focused community of practice on qualitative approaches to MEL, this collaborative action research has aimed to create a set of tested and practical principles to support qualitative assessment practice in Asia and the Pacific.

Led by this expressed need for practical and tested guidance, researchers in iDE Cambodia's SMSU3 program and the University of Technology Sydney's Institute for Sustainable Futures (UTS-ISF) conducted a Collaborative Action Research (CAR) study. The pilot assessment explored the impacts of a gender mainstreaming intervention for SMSU3 staff members, and a set of principles was formulated through a series of collaborative workshops.

In this short practice note, we explore existing guidance tools, introduce the six principles developed through the CAR process, and provide examples of how the principles can be used in practice.

Existing guidance for qualitative evaluators

The evaluation sector has a strong foundation of principles, standards, norms, and dimensions to guide evaluations. These guidance tools have been compiled by donors, associations or in connection with a particular methodological approach. In contrast to practice-driven approaches, these tools were derived from normative values as theory-driven approaches. Six of the most relevant guidance tools, five of which are from grey literature, are summarised in Table 1 in their most recent iterations. Two of the guidance documents are specifically focused on qualitative evaluations, while the other four are more broadly focused on assessments with applications to qualitative practice. Additionally, three of the tools are focused on evaluation in international contexts.

[Insert - **Table 1.** Existing guidance on qualitative evaluation]

These guidance tools contain between four and ten principles, which we have mapped onto the Lean Research Framework (Hoffecker et al. 2015). This framework held the most resonance with our CAR findings despite being the only reviewed guidance not explicitly focused on evaluation but research more broadly. As shown in Figure 1, *Rigorous* was the most identified principle within the six reviewed guidance tools, with *'credibility'* used in three of the six frameworks, followed by the principle of *Respectful,* which promotes a do-no-harm approach to assessments.

These existing tools offer a strong foundation of recommended practices in the evaluation sector. Nevertheless, there is an opportunity to balance theory and practice in the international development context and strengthen principles for practitioners conducting qualitative assessments. We determined that aspects of transformation and reflexivity were missing in all six frameworks, which aim to redress issues of power, tokenistic participation, and extraction inherent within many research designs.

[Insert - **Figure 1.** Circle packing diagram chart of the named principles in the six selected evaluation quality frameworks articulated through the Lean Research Framework.]

Collaborative Action Approach

The expressed need for this research arose during a series of webinars within a community of MEL practitioners based in Asia and the Pacific. During these webinars, community members shared their lack of confidence in conducting qualitative forms of assessment and their uncertainty about how to defend of justify the use of more creative forms of research and evaluation in their work, and ensure such work had rigour and quality. Such an expressed need is in the words of Paulo Freire a *'generative theme'* (Freire, 2000)*.*

Building on the six frameworks identified during our literature review, a collaborative team based in four countries conducted an interactive process to pilot and design a set of principles for use in qualitative assessments in the context of international development. While the research was facilitated by researchers in the north, the aim was to create practical tools with and for researchers in the south. The study was framed as Collaborative Action Research (CAR), an action-based approach that supports people with different "*responsibilities and roles to work together to achieve a shared common purpose"* (Coghlan and Brydon-Miller, 2014: 116). Integrating theory and action within a CAR model creates a unique opportunity to blend academic rigour and evaluation practice in supporting civil society organisations to expand notions of "acceptable" qualitative assessment techniques.

The study collected data at two levels: the meta-level and the evaluation-level, as visualised in Figure 2 The evaluation-level component based in Cambodia included reflection questions embedded into a qualitative assessment (n=176) and 16 reflection interviews with SMSU3 respondents. The meta-level included data from the research team through four workshop discussions and informal conversations between the lead researcher and research team members. The first workshop was conducted in Cambodia (March 2020) with a group of 15 Cambodian practitioners and the later workshops were conducted online with the diverse group of co-authors spread across four countries. Data from both levels was synthesised collaboratively to produce iterative versions of the framework while engaging with evaluation and design thinking best practices.

The core research team (named authors in this work) consisted of diverse group of development researchers and practitioners all with significant experience in the development sector and deep connections to practice across the global south. While the work was facilitated by a northern institute, the value of diverse voices was critical to the process. Reflection on the spectrum of participation in development activities (Cornwall, 2003; White, 1996), describes this research as a blend of transformative and self-mobilizing. In this framing, the diverse participants in the collaborative action approach have been owners and agents in the process which was facilitated by researchers based in a northern research institute.

*[Insert - **Figure 2.** Collaborative Action Research Process]*

Introducing the principles

The final six principles are designed to support practitioners undertaking qualitative assessments in the international development sector. The collaborative and field-tested process allowed for an iterative approach to designing the framework. Several notable iterations included aligning the principles with the Lean Research Framework and adding aspects of transformative and reflexive research to encourage a do-more-good approach through process and outcomes of assessment. A do-more-good approach aims to reframe research and evaluation ethics from focusing on harm, towards fostering approaches that can promote positive outcomes and processes of assessment.

The evolution of the principles was brought about through challenges in their application, for example, the difficulty of simplifying a breadth of guidance and different aspects of data collection, analysis, and utilisation into a concise framework that includes two framing principles (top and bottom) and four core principles. The core principles focus on more functional aspects, whereas the framing principles focus on researcher mindset. The final six principles are summarised in Table 2.

*[Insert - **Table 2.** Framework: Six principles of quality assessments and their definitions]*

This model includes considerations for transformative and reflexive research practices, both of which are missing from existing frameworks. Transformative research and evaluation approaches are becoming increasingly important in research that aims to foster social change both through the process and from the outcomes of assessments (Freire, 2000; Mertens, 2009). Additionally, reflexive assessment has emerged from action research practices which encourage self-reflection and critique to improve future practice (Fetterman, 2017; Freire, 2000). Both are underpinned by the foundational work of Paulo Freire whose participatory and reflexive approaches to social transformations highlight the value of personal reflection leading to action (Freire, 2000),

These two principles aim to redress critical challenges inherent in research processes. For example, transformative research aims to address the extractive nature of much research by focusing on creating benefit for the participants within the research process and outcomes; an approach championed by Robert Chambers in participatory rural appraisal (Chambers, 1994). The transformative principle can begin conversations on decolonizing measurement efforts but is not fully designed to address them. Similarly, reflexive research fosters reflection on the power dynamics inherent in research and the tendencies of research activities in development interventions to become tokenistic about participation (White, 1996). By depicting these two principles as frames (top and bottom of the framework), we suggest that these more radical principles should underpin and be reflected in the other four

principles. Qualitative assessments in the international development sector can only be strengthened by this type of purposeful focus and reflection.

Using the principles

The principles have been designed to be used by evaluation teams before, during and after assessments. The principles were built on the premise that tools, methods, and approaches can all be tailored to be good quality throughout data collection, analysis, and utilisation. The principles encourage teams to place less emphasis on 'what to do', and rather focus more on 'how to do'.

Before: Planning assessments

As a planning tool, the framework can be used as a conversation starter for teams in the planning process. Teams can use each of the 12 items as 'how might we' discussion questions to select and tailor appropriate tools, methods, and approaches as they plan assessments.

During: Conducting assessments

The framework can be used as debrief topics for research teams during an assessment. Daily debriefs are an important part of evaluation processes and ensure that expectations are being met. In this sense, debriefs should not just be undertaken during data collection but also during data analysis.

After: Evaluating assessments

The framework can be used to evaluate the efficacy of an assessment. One approach is to use the framework as a series of discussion questions to frame a reflection workshop. Similarly, additional questions can be embedded within the assessment to track how participants feel about the tools, methods, and content; both as observations from the research team and reflections directly from participants. For example, in our evaluation which used micro-stories, we included reflection questions for participants focused on the four central principles in the framework. Regarding the *Respectful* principle, we asked if the participants felt safe and comfortable during the audio survey. We also conducted quick reflection interviews with 16 participants who completed the audio surveys and inquired on perceptions of safety. An example of a rubric-style review of our micro-story collection is depicted in Table 3. This review was conducted in an online workshop format a year after the study was conducted. This ensured time to consider the impacts of the study findings on program implementation.

[Insert - **Table 3***. Example Assessment of SMSU3 midterm evaluation process**. Drawing from an online reflection workshop [October 2021], quotations are directly from the workshop or from reflection interviews with participants, percentages drawn from feedback from participants (n=176)]*

Limitations of the principles

While these principles have been developed in a practice-based dialogue between a multi-national collective of researchers and practitioners, they have only been tested in one context. Further research is required to explore their applicability in other geographies and with other forms of qualitative evaluation. As the principles were formed through the CAR process, further research may also evaluate their applicability for the entire design, implementation, and reflection process of an evaluation.

Conclusions

Building on established guidance for evaluators and drawing on a collaborative research process, the six field-tested principles introduced in this practice note have been designed to support practitioners in conducting quality assessments in the field of international development. They aim to reinvigorate a conversation in the sector on the opportunities to foster more robust assessments and to encourage teams to stretch beyond the familiar ground of focus groups and interviews while maintaining rigour. The addition of principles related to transformative and reflexive approaches also seek to support evaluators in 'doing-more-good' within the assessment process and to take advantage of the potential of an evaluative process to also be one that initiates beneficial change. Lastly, while the principles have been designed with qualitative assessments in mind, the process revealed that the principles can also be useful within the context of quantitative or mixed methods assessments.

Disclosure statement

The authors report there are no competing interests to declare.

Biographical Notes

Ms. Jess MacArthur's research explores the social dynamics of water and sanitation programming in South and Southeast Asia and focuses on gender research practices in the development sector.

Ms. Rana Abdel Sattar supports strategic decision-making for program operations, oversees the development of new WASH products and services in iDE Cambodia's Sanitation Marketing Scal-up (SMSU3) Program.

Ms. Naomi Carrard is a Research Director at the Institute for Sustainable Futures. Her applied research focuses on water, sanitation and hygiene (WASH) across the Asia-Pacific region

Mr. Tyler Kozole is the WASH Program Manager at iDE Cambodia. He leads iDE Cambodia's WASH and climate change programming and research initiatives, focusing on reaching households in climate-vulnerable areas.

Mr. Chris Nicoletti leads iDE's global measurement efforts, including rigorous impact evaluations, designing and implementing efficient management information systems, and effectively communicating data and results.

Dr. Lauren Riley is the gender equality and social inclusion manager at iDE focusing on market-based approaches to inclusion and equality challenges using both qualitative and quantitative methods.

Dr. Jennifer Roglà is iDE's Director of Research & Evidence focused on identifying questions the organisation is in a unique place to answer, garnering the resources and partners to answer them, and leading the charge in influencing the sector with the lessons that we are learning.

Mr. Veasna Toeur is a Monitoring and Evaluation Manager at iDE Cambodia. He leads the monitoring and evaluation team focused on improving WASH systems and gender equality in Cambodia.

Professor Juliet Willetts leads applied research to improve development policy and practice and to address social justice and sustainable development, including achievement of the global Sustainable Development Goals (SDGs).

References

AEA (2018) *Guiding Principles*. Available at: https://www.eval.org/p/cm/ld/fid=51

Bamberger M (2000) The Evaluation of International Development Programs: A View from the Front. *American Journal of Evaluation* 21(1): 95–102.

Chambers, R. (1994). The origins and practice of participatory rural appraisal. *World Development*, 22(7), 953–969. https://doi.org/10.1016/0305-750X(94)90141-4

Coghlan, D., & Brydon-Miller, M. (2014). Collaborative Action Research Network. *The SAGE Encyclopedia of Action Research*, 117–119. https://doi.org/10.4135/9781446294406.n58

Cornwall, A. (2003). Whose Voices? Whose Choices? Reflections on gender and participatory development. *World Development*, *31*(8), 1325–1342. https://doi.org/10.1016/S0305-750X(03)00086-X

Fetterman, D. (2017). Transformative Empowerment Evaluation and Freirean Pedagogy: Alignment With an Emancipatory Tradition. *New Directions for Evaluation*, *2017*(155), 111–126. https://doi.org/10.1002/ev.20257

Freire, P. (2000). *The Pedagogy of the Oppressed*. Penguin Books.

Hoffecker E, Leith K and Wilson K (2015) The Lean Research Framework. (August): 6. Available at: d-lab.mit.edu/lean-research.

Joint Committee on Standards for Educational Evaluation (2018) Checklist of the program evaluation standards statements.

Lincoln Y and Guba EG (1985) *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications.

Mark MM (2008) Building a better evidence base for evaluation theory. *Fundamental issues in evaluation* 1: 11–126.

Mayoux, L., & Chambers, R. (2005). Reversing the paradigm: Quantification, participatory methods and pro-poor impact assessment. *Journal of International Development*, *17*(2), 271–298. https://doi.org/10.1002/jid.1214

Mertens, D. (2009). *Transformative Research and Evaluation*. Gilford.

Patton M (2002) *Qualitative Research and Evaluation Methods*.

Phillips T and de Wet JP (2017) Towards rigorous practice: A framework for assessing naturalistic evaluations in the development sector. *Evaluation* 23(1): 102–120. DOI: 10.1177/1356389016682777.

Spencer L, Ritchie J, Lewis J, et al. (2003) *Quality in Qualitative Evaluation: A framework for assessing research evidence*.

UNEG (2016) *Norms and Standards for Evaluation*. Available at: http://www.unevaluation.org/2016-Norms-and-Standards

White, S. C. (1996). Depoliticising development: The uses and abuses of participation. *Development in Practice*, *6*(1), 6–15. https://doi.org/10.1080/0961452961000157564

Table 1. *Existing guidance on qualitative evaluation*

Title of Guidance	Source (Years Ratified)	Audience	Purpose	Parameters and Principles	Unique Attributes
Program Evaluation Standards	The Joint Committee on Standards for Educational Evaluation (1994, 2011)	Evaluators and evaluation users	To "guide evaluators and evaluation users in the pursuit of evaluation quality"	1. Utility 2. Feasibility 3. Propriety 4. Accuracy 5. Accountability	Focused on evaluation outcomes, primarily in the North American context.
Evaluators' Ethical Guiding Principles	American Evaluation Association (1994, 2018)	Evaluators	To "guide to the professional ethical conduct of evaluators."	1. Systematic Inquiry 2. Competence 3. Integrity 4. Respect for People 5. Common Good and Equity	Focused on evaluator conduct, primarily in the North American context.
Quality in Qualitative Evaluation: Guiding Principles	UK Government Chief Social Researcher's Office (2003)	Government appraisers of evaluations	To "critically assess" the quality of qualitative evaluations	1. Contributory 2. Defensible 3. Rigorous 4. Credible	Focused on qualitative evaluations, primarily in the United Kingdom.
Criteria of Trustworthy Naturalistic Evaluations	Framework for rigorous practice (Phillips & de Wet, 2017; Lincoln & Guba 1985)	Naturalistic (qualitative) evaluators	A "framework for assessing the rigour of naturalistic evaluations in the development sector"	1. Credibility 2. Transferability 3. Dependability/ Auditability 4. Confirmability	Focused on qualitative evaluations in the South Africa context applied as a meta-assessment framework.
Norms of Evaluation	United Nations Evaluation Group (2005, 2016)	UN Institutions	To "ensure that United Nations evaluation functions provide credible and useful evidence to inform and strengthen the work of the United Nations system in pursuit of its goals."	1. Internationally agreed principles, goals and targets 2. Utility 3. Credibility 4. Independence 5. Impartiality 6. Ethics 7. Transparency 8. Human rights and gender equality 9. National evaluation capacities 10. Professionalism	Focused on evaluations within the international development sector.
Principles of Research	Lean Research Framework (Hoffecker et al. 2017)	Researchers and designers in international Development contexts	"...to guide and improve the practice of field research with people and communities in the contexts of international development and humanitarian work…[often where] the impact of research activity on the lives of research subjects, communities, and local partners is often ignored."	1. Rigorous 2. Respectful 3. Relevant 4. Right-sized	Focused on the international development context and on the impact of research on participants.

Table 2. Framework: Six principles of quality assessments and their definitions

Principle	Definition
Transformative	Inspires positive change for participants and researchers through the process of the research.
	Inspires positive change for programs and organisations from the outcomes of the research
Respectful	Protects human and legal rights and maintains the dignity of participants and stakeholders.
	Proactively involves a diversity of participants and prioritises their satisfaction with the research process.
Relevant	Generates and disseminates rich and useable insights.
	Suitable for the cultural, geographic, and situational context.
Right-sized	Adopts relevant, simple, and convenient tools and techniques.
	Effectively and efficiently leverages time, money, and skillsets.
Rigorous	Employs a systematic approach to sampling, collection, analysis, and interpretation.
	Ensures well-founded, plausible, and justified insights, supported, and refined by existing evidence.
Reflexive	Engages openly about assumptions and other complementary and conflicting perspectives.
	Remains aware and honest about dynamics of power between the participants, researchers, and the broader stakeholders.

Table 3. *Example Assessment of SMSU3 midterm evaluation process*. *Drawing from an online reflection workshop [October 2021], quotations are directly from the workshop or from reflection interviews with participants, percentages drawn from feedback from participants (n=176)*

Approach	**Significant Change micro-stories**, Elicitation of micro-stories through cartoon visual on personal smartphones for 176 program staff of SMSU3.		
Methods	**Data Collection:** Audio survey (audio or text responses) [October 2020]	**Analysis**: Personas developed using both manual and computer-driven techniques [November – February 2021]	**Utilisation:** Online sensemaking workshop with key program staff and leadership [August 2021]
Tools	**Data Collection:** Qualtrics, Phonic	**Analysis**: Airtable, Excel, R Studio, PowerPoint	**Utilisation**: Surprises, Questions & Recommendations Canvas, Zoom, Google Slides, Personas (presentation and printed set)

Principles	Definition (approach…)	Score	Comments and Justification
Transformative	Inspires positive change for participants and researchers through the process of the research.	3	- "Besides learning a new technology, the team members have a chance to recall the gender knowledge when answering the questions." - "Stories can show evidence of change which reflected by each participant" - "What I like about the survey is that maybe the questions can empower women staff members." - "Now that we've done it this first time, it will be easier for next time" - 78% strongly agree – this survey helped me think more deeply about gender equality
	Inspires positive change for programs and organisations from the outcomes of the research	4	- Able cluster stories and create profiles/personas. Personas used in the planning of future gender mainstreaming strategies in a workshop - "MHM and other insights to be incorporated into Gender Mainstreaming training" - "Helped validate importance and impact of Gender Mainstreaming. Easier to use these insights to encourage other iDE programs to adopt this training." - 79% strongly agree - my answers can make SMSU3 better for staff; 73% strongly agree - my answers can make SMSU3 better for clients
Respectful	Protects human and legal rights and maintains the dignity of participants and stakeholders.	4	- "Some participants were uncomfortable sharing over audio", giving options for both audio and text mitigates these challenges - 82% of men and 76% of women felt fully comfortable.
	Proactively involves a diversity of participants and prioritises their satisfaction with the research process.	4	- Participants determine the type of micro-story they wish to share based on the card they select - Card sorting and audio recording allows for more engagement, but no purposeful interactive aspects between participants. - "Lots of feedback that the survey was fun to do" 65% strongly agreed the survey was enjoyable, 71% strongly agreed it was interesting - "Census sampling, so almost everyone participated"
Relevant	Generates and disseminates rich and useable insights.	4	- Rich data and new insights because of the breadth - However not able to go deep in the surveys and to probe for more. - "The personas analysis approach presents the data in a more digestible way" - "Provided key insights on how effective our gender transformative approach."
	Suitable for the cultural, geographic, and situational context.	4	- 79% participants agreed that the survey was extremely or very effective - "Detailed, participatory workshops ensured that the images, content, questions, and process all fit well within our context. Several iterations led to a solid final product."
Right-sized	Adopts relevant, simple, and convenient tools and techniques.	4	- Able to leverage staff access to mobile phones. - "Adding the audio recording option took some time to get right, but it worked well"
	Effectively and efficiently leverages time, money, and skillsets.	3	- Analysis does take time, but it is faster than traditional interviewing. Able to collect 176 responses in 2 weeks - Requires software licenses - "Effective at getting all field staff to participate, but did take some time from everyone"
Rigorous	Employs a systematic approach to sampling, collection, analysis, and interpretation.	4	- Very systematic methods for analysis and semi-quantitative analysis - "The approach provides clear qualitative and quantitative evidence" - Selected analysis tools to track data analysis process transparently and safely.
	Ensures well-founded, plausible, and justified insights, supported, and refined by existing evidence.	4	- Not ethical to share the stories back to participants, but can share composite stories - No way to check the bias of respondents. "Difficult to check if the stories are accurate/honest" - 69% strongly believed that others answered truthfully - Personas were designed using both manual and computer-driven approaches
Reflexive	Engages openly about assumptions and other complementary and conflicting perspectives.	4	- "Interactive workshops and sessions helped ensure that assumptions were being checked" - "Trialling tool a few times before deployment was key"
	Remains aware and honest about dynamics of power between the participants, researchers, and the broader stakeholders.	3	- Debriefs with research assistants helped to explore these dynamics - "Feedback that respondents were concerned about 'getting the right' answers."