Refining the intervention: qualitative pilot with authors

Commit ID: b2b401e

2023-07-21

Having defined intervention components and built a prototype (chapter 10) I wanted to refine the website by getting feedback from authors. Although I had included evidence from authors in my earlier work, this evidence came from secondary sources many of which had limitations. Many studies were free-form text surveys which resulted in thin description, and samples generally lacked diversity; almost all participants came from western academic institutions. To refine my website, I wanted to capture rich descriptions of experiences from authors from around the world, with different levels of experience, and different places of work. This was especially important as I hadn’t involved authors earlier in the design or development process.

# Methods

My objective was to refine the website by identifying parts that may not be functioning as I had intended them to.

Choosing a guideline  
  
\* Editing a guideline was time consuming  
\* Not possible nor necessary to do all, only needed one  
\* SRQR was good choice because general = large sampling pool, used by academics and non-academics, I was familiar with it

## Sampling strategy

To be eligible, participants had to be actively engaged in qualitative research and be able to attend an online interview conducted in English. I used purposive sampling, seeking participants that were all doing qualitative research but varied in their experience (years doing research), workplace (academia, industry, clinician etc.), geographic location, and stage of writing (drafting vs editing). I chose these dimensions as my previous work identified that different groups of people encounter different barriers. I was prepared to do additional sampling should I discover other dimensions of importance during data collection. Participants were offered £50 as remuneration for their time as a bank transfer (or Amazon voucher for UK participants).

I used information power to guide my initial sample size (see chapter 9 for an introduction to information power). I considered my aim to be narrow and sample to be specific but with adequate variation. My intervention and analysis was based on a behaviour change framework, and I expected my methods to provide ample opportunities for deep discussion. For these reasons, I felt confident that 10 participants would provide adequate information power whilst also being manageable within the time limit of my DPhil.

I recruited participants by a) posting adverts on Twitter which were retweeted by the EQUATOR Network and AuthorAid [1] (a network for researchers from low and middle income countries); b) emails forwarded by AuthorAid, the African Research Integrity Network [2] and 101 Health Research [3] (a publication service provider based in the Philippines); and c) a notification to authors using Penelope.ai [4], a free manuscript checking tool used by medical journals.

## Data collection methods

Defining intervention components by their target barriers and intervention functions lends itself nicely to designing a pilot study to gather feedback. Because I knew what each component was supposed to be *doing* I could design an interview schedule with questions and tasks to specifically explore intervention functions.

My interview schedule had 4 parts:

1. 5 second test
2. Think Aloud + Interview
3. Writing task (at home)
4. Writing review + interview

## 5 Second Test

GOT TO HERE

~For example, the purpose of the top of the home page is to communicate what reporting guidelines are and how they benefit authors. Instead of asking general questions (like “what do you think of…?”), I decided to ask more specific questions (“What do you think the website is about? How do you think it might influence your work?”). Furthermore, because the purpose was to communicate these things *quickly*, I decided to only give pilot participants 5s before asking these questions.~ Similarly, defining components in this way will facilitate future quantitative work to assess efficacy.

### Demographics questionnaire

After giving informed consent, participants provided their experience, place of work, english proficiency, and country of residence using an online form.

### 5 second test

Participants were invited to an initial interview conducted over Zoom. The lead researcher shared their screen and displayed the website homepage. Without warning, the page was closed after 5 seconds after which we used semi structured interviews to assess participants immediate understanding of and feeling towards the home page. Questions included:

“What do you think the website is about?”

If the participant mentioned the term “reporting guidelines”, we asked “What do you think reporting guidelines are?” and “What tasks do you think you can use reporting guidelines for?”

“How would you describe the design of the site?”

“What would make you want to learn more about the website?”

“What impact do you expect the website to have on your job as a researcher?”

### User protocols / think aloud tasks

#### 1. Finding guidance and resources

We gave participants descriptions of research and asked them to identify the most relevant reporting guideline from the website. We did not tell them how to go about this - participants could use direct links on the homepage, the search bar, or could follow the “wizard” questionnaire. We did this task 3 times. The first time, participants were asked to find a guideline for animal research. This was the easiest guideline to find as the description was displayed prominently on the home page. Second, we asked participants to find guidance for reporting cohort studies. This was a little harder as it required participants to read and understand descriptions of study designs to distinguish between different, related, epidemiology guidelines. Finally, we asked participants to find what guidance they would to report (# DECIDE: what impossible task?). This was a difficult question as there is not perfect reporting guideline for this kind of article. Instead, participants had to (# FIXME: complete task).

We also presented tasks that involved finding tools. We asked participants to imagine they had been asked to submit a “completed checklist” by a journal, requiring them to find the guidance, then find and use the associated checklist. We asked participants what they expected “to do lists” and “templates” to be, and when they might be used.

#### 2. Finding information within a guideline

We then asked participants to find information within the #DECIDE: guideline. We asked participants how they would report #FIXME, why it is important to describe #FIXME, and what they should write if they hadn’t done #FIXME. These questions required participants to find items #FIXME respectively, and to locate content within collapsible boxes.

### Plus - minus test

At the end of the first interview we gave participants a task to complete in their own time over the coming weeks. We provided them with the methods items of the #FIXME guideline in a #DECIDE Word file. Participants were to read the methods items of the #FIXME guidelines in their own time and highlight sentences that elicited positive or negative responses and to mark them with a “+” or “-” symbol. They could add notes if they wanted to.

### Writing evaluation (Performance test)

If participants were actively writing an article we asked them to use the guidelines to write their methods section. If they had something already written, we asked them to complete a reporting checklist. We asked them to do this in their own time, within two weeks. Once complete, participants sent their work to JH via email who then checked their reporting against the the #FIXME guidelines, noting which items had been reported fully and which hadn’t.

### Retrospective interview

Participants then attended a follow up interview two weeks later where JH asked open questions to explore the reasons behind participants + and - marks, and reasons for neglecting any items in their writing sample. The +/- test aimed to pick up non-specific responses, which would include parts of the text where the participant found difficult to understand. The writing check, however, would also reveal parts of the guidance that the participant had unknowingly \_mis\_understood.

Finally, we asked the participants again for their opinions on the website and guidance and how it could be improved.

## Data collection instruments and technologies

Interviews were conducted over Microsoft Teams, using its in-built video and audio recording. We created interview schedules (#REF) and piloted them amongst students in the department. The version of the website tested can be viewed at (#REF).

## Units of study

## Data processing

We used Zoom to automatically transcribe audio recordings, and then manually double checked the transcripts and added context from the videos, interview notes, +/- annotations and writing sample. We imported transcripts into NVivo (#REF), creating cases for participants and intervention components.

## Data analysis

We coded positive and negative experiences and grouped them by intervention component. We did this because we expected its output - negative and positive experiences grouped by intervention component - to be easier to act upon than if we grouped experiences by method.

## Techniques to enhance trustworthiness

* Double checking of coding
* Member checking
* Triangulation? Read the paper Charlotte recommended. Perhaps it is about mixed method studies? Perhaps that means I could count errors / number of people that did a task successfully? [https://www.bmj.com/bmj/section-pdf/186156?path=/bmj/341/7783/Research\_Methods\_Reporting.full.pdf]

## Ethical issues pertaining to human subjects

The study was approved by the Medical Sciences Interdivisional Research Ethics Committee at Oxford University. Participants gave informed consent via a web-form.

The context in which users encounter the website will effect how it is perceived.

Most visitors come from journal websites, having been instructed by an editor or JAI page to use a particular reporting guideline. Others may come to the site because a a guideline has been recommended by a friend. Very few arrive completely cold turkey.

In contrast to this study, where all participants were totally cold (unless they were already familiar with EQUATOR).

So when one participant suggested that “What are RGs” should come above the Frequently Accessed Guidelines (Or at least that the writing, checking, planning should) this may appear sensible on the surface. But given that most authors will come already with a guideline’s name in mind (even if they don’t really know what it is), this ight not be necessary.

Therefore, need a real world evaluation.

All could speak English = bias

1. AuthorAID - Home.

2. ARIN home page. ARIN

3. 101 Health Research & Statistics Quality. Efficiency. Ethics. Teamwork.

4. Penelope.ai. Penelope.ai