

Qualitative Methodology and Analysis Guidebook for Medical Educators

By Lindsay Demers, PhD
Director, Education Evaluation Core
Boston University School of Medicine

Caveat:

The overarching goal of this document is to provide clinician and scientist-educators with a methodological and analytic road map for medical education research. The level of detail included here is not sufficient to gain expertise in these areas; rather, think of this as a document you can use for guidance when designing your research and making analytic choices. If you find a topic that is of [Qualitative Coding Bootcamp](#) substantial relevance to your work, please email me for additional resources on the topic, as only the bare minimum of detail is included here.

Further, I have only included here what seems to be most commonly used within the department of medicine. It is likely you will encounter instances that are not covered in this guidebook. If/when this occurs email the Education Evaluation Core for a consultation (if you're within the DoM or part of CELP) or email ldemers@bu.edu for additional resources.

Myself and Dr. Ellen Childs (formerly of SPH, currently of Abt) published a Qualitative Coding Bootcamp on MedEdPORTAL. You can find that resource here:

https://www.mededportal.org/doi/10.15766/mep_2374-8265.10769

Commonly Used Qualitative Approaches in Medical Education

This guide provides a brief overview of commonly used qualitative approaches in medical education. Conducting qualitative research can be a very fun, intuitive and informative approach to answering your research questions. That said, it is important to be intentional and as rigorous as possible when selecting your approach. Below, I cover some of the basics of using qualitative approaches.

Sampling for Qualitative Methods

There are many different approaches when it comes to sampling in qualitative methods. The approach you choose will depend on the type of questions you are asking, the population you are studying, as well as other study-specific considerations. Here I will cover the three main sampling approaches we have used in the Education Evaluation Core.

Random Sampling

If you do not have specific criteria or characteristics that you want your interview subjects to have, you may simply draw a random sample from your population of interest. Unlike in quantitative approaches, using a random sample does not allow you to extrapolate the findings from your sample to your population. However, if you otherwise don't have specific ideas about who to talk to and why, a random sample can be useful to identify potential interview participants.

Purposeful Sampling

Although this sampling approach is quite broad in definition, generally speaking if you are purposefully selecting interview participants based on some criteria, this would be considered purposeful (as opposed to random) sampling. For example, say you did a survey of residents asking about their confidence in a given area. Maybe you only want to talk to people who said they had low confidence in this area to figure out why these people may be lacking confidence. You may also select people based on demographic characteristics – for example, if you are interested in the experiences of a group historically excluded from medicine, you might use that characteristic (be it gender, race, ethnicity, LGBTQIA+ status, etc.) as a way to determine your sample. Or, maybe you want to talk to Program Directors around the country to answer a research question about the structure of residency program.

Snowball Sampling

Just like the name implies, snowball sampling is an approach where the further along you go, the bigger your sample gets (imagine rolling a snowball and it getting bigger and bigger). In snowball sampling, you start by identifying a small group of potential participants, and then ask those participants to recommend other participants for you to interview. By having your interview participants recommend other potential interview participants, you may get ideas of who to talk to that you didn't previously have. This approach is also useful if you're unsure about who are the best people for you to talk to in order to accurately and thoroughly answer your research questions.

Once you've decided how to determine your sample, then you should decide which data collection approach will work best to answer your research questions. Typically, it is either interviews or focus groups.

Interviews

Interviews are a one-on-one dialogue between an interviewer and an interviewee. Interviews offer the research an opportunity to build rapport with their research participants, and can provide researchers with rich contextual information about the phenomenon being studied. There are three types of interviews:

- **Structured:** In this approach, the interview guide is pre-written and the interviewer does not deviate from the list of questions at all (it's basically a verbal survey)
 - Can be useful when you only have a specific list of limited questions you need answered
 - Requires less training on the part of the interviewer since there is no deviation from the interview guide
 - However, for obvious reasons, the data will be limited to only the pre-written questions and other possibly interesting phenomena that arise during the interview cannot be explored.
- **Semi-Structured:** This approach is the most commonly used in qualitative research. As the name implies, in a semi-structured interview the interview guide is pre-written, however if new and interesting lines of inquiry arise during the interview, the interviewer has the ability to go "off-script" to probe about these additional areas.
 - This approach usually requires a bit more training for the interviewer as well as a fairly good understanding of the research questions being asked so that they can appropriately ask additional questions
 - Assuming you are not psychic, you will not necessarily know what new topics might come up during the interview. However, you should do your best when writing the interview guide to anticipate where things might go so that you can write probing follow up questions. Anticipating as much as you can beforehand will help to ensure that the probing questions are not leading.
- **Unstructured:** Unstructured interviews are rarely used in research because the odds that your interviewee will organically bring up the topics you're interested in are not in your favor. This approach is more commonly used by clinicians with patients (e.g., tell me what brings you in today) and less so with research participants.

Because semi-structured interviews are the ones most commonly used in qualitative research, the remainder of this section is going to focus solely on this approach.

Designing your Interview Guide

Compared to designing a survey, designing an interview guide is relatively less stressful. When you think of a survey, if you have poorly written or unclear questions, once the survey has gone out there's really not much you can do to salvage the data. However, in an interview, the researcher is hearing the interviewees' responses in real time and can provide clarification, course-correction etc if the question is not being answered as intended. That said, it still is important to think carefully about what questions you want to ask in your interview guide and how to ask them.

One of the biggest pitfalls I have seen in interview guides is a tendency to ask leading questions – particularly in evaluation research. For example, asking "What did you like about X" versus asking "Tell me how you feel about X." In the first question the interviewee is being primed to provide a positive

answer (“what they like”) versus the second question which does not presuppose any particular thoughts or feelings.

Another question I get asked frequently is how long your interview guide should be. The response to this question is not straightforward. You may have a specific time frame within which you can conduct the interviews (e.g., 30 minutes) which can limit the number of questions you can ask. However, your interviewees are going to vary wildly in how long their answers to your questions are. For this reason, there is usually not a one-size-fits-all response to interview guide length. Generally speaking, if you’re getting into double-digits (10 questions or more) you’re likely not going to have time to get to all your questions unless your interview participants have time for a longer interview.

If you are worried about time running out before you have asked all your questions, be sure to order your interview questions based on priority so that the most crucial questions are asked first. Then if you have a talkative participant or run out of time for another reason, you’ll likely have captured the data that are most important to answering your research question. This advice assumes that your interview questions are not invasive or overly person – just questions about experience in an education intervention or program. If you are interviewing participants about sensitive topics you will want to be purposeful about the order in which questions are asked.

Key Considerations when Conducting Interviews

These are some best practices and things to keep in mind when you are conducting an interview study based on my experience.

- Assuming your interviewee consents, **AUDIO RECORD YOUR INTERVIEWS**. It’s best to use a small recorder (they are relatively inexpensive). You can theoretically use your cell phone’s recorder, however you would need to put your phone on airplane mode and then transfer the file to a secure computer and delete it before turning airplane mode off. By putting your phone into airplane mode you are turning off the wifi/5g and essentially turning it into a recorder.
 - If an interviewee does not consent to being audio recorded, it is best to have a third party present who can take notes while you conduct the interview. If that isn’t possible, take the best notes you can during the interview and immediately after the interview is complete, write down as much as you can possibly remember about the conversation.
- Be sure you have a plan for transcription. Transcribing interviews is tedious and time consuming. If you have any funding, I recommend you get the interview audio recordings professionally transcribed. As of June 2022, the rate I have for a professional transcriptionist is \$1.99 per recorded minute (so a 60 minute interview would be \$120 to transcribe). If you or someone from your team must do the transcription, I recommend using free software (e.g., Audacity) which will allow you to slow down the audio file and will make transcription easier.
 - If your participants reveal any identifying information during their interview or if they ask for something to be struck from the conversation, be sure whoever is doing the transcription is aware of this and redact anything of this nature before delivering the transcripts.
- If you are the person who developed and/or implemented the education intervention being studied, it is best for you not to conduct the interviews due to the potential for bias. If possible, you should identify an outside party who was not directly affiliated with the intervention to conduct the interviews. If you cannot find someone to fill this role, it is imperative that you have

a colleague review your interview guide to ensure the questions are not biased/leading in any way.

Possible Limitations of Using Interviews

The biggest drawback to conducting interviews is that they are very time-consuming for the researcher. The scheduling of the interviews, conducting the interviews, transcribing the interviews and analyzing the interview data can be very time-intensive. For this reason, I encourage people wanting to do qualitative work to find a colleague, trainee, administrator etc. who is interested in collaborating to help share the labor.

Focus Groups

After interviews, the most common qualitative method I have seen used in medical education is focus groups.

Designing your Focus Group Guide

My advice for designing a focus group guide does not differ much from my advice for the interview guide – you still want your questions to be clear and not leading. However, unlike in an interview, the goal is to get participants talking to each other, and you will want to keep that in mind as you write your guide. You also may want to include probes you can use if you want to elicit responses from focus group members who may not be as forthcoming and/or talkative as others.

Key Considerations when Conducting Focus Groups

- One of the trickiest parts of facilitating a focus group is managing group dynamics. As in all scenarios, you will have people who are very talkative, possibly bordering on dominant, and others who are less forthcoming. One of the major responsibilities of the focus group facilitator is to do their best to ensure that all voices are heard and to create space for those voices to come through. I do not have any one size fits all tips in this regard, other than to think in advance about what you might say to encourage participation among those less talkative – here are some generic examples:
 - “We haven’t heard from you on this topic, do you have thoughts?”
 - “I saw you smirk/nod your head/whatever, do you agree with his/her/their thoughts?”
 - “Was your experience in-line with what we’ve been hearing?”
- Keep in mind that the goal of the focus group is to encourage participants to talk to each other and have a conversation, versus a sequential one-on-one conversation between the researcher and each individual focus group member. If you find yourself falling into this pattern as you facilitate the focus group, you may need to leave more space for awkward silence which will give people the opportunity to speak. If as soon as a participant is done talking you start talking, you will be setting a tone of a one-on-one conversation between you and the participant which is not what you are striving for with a focus group.
- As with interviews, if you are the person who developed and/or implemented the education intervention being evaluated, it is best that you do not facilitate the focus groups due to potential bias.

Possible Limitations of Focus Groups

- Although the extent of this limitation will depend on the sensitivity of the topic(s) being discussed, one potential issue with focus groups is that participants might be less forthcoming or

honest because of social desirability or not wanting to appear to be “the odd one out.” For this reason, if you’re asking questions that might be sensitive at all in nature, you might consider doing interviews instead. If a focus group is the only option, your best bet is to emphasize to participants that their data will be de-identified and transcribed to increase the chances that they will feel comfortable disclosing to the group.

- Unlike in interviews where transcription is simply tedious, focus group transcription is tedious and difficult due to the multiple voices and potential for people talking over one another. Ideally the facilitator will minimize the amount of competition for air time among participants, but these audio recordings will still be more challenging to transcribe, especially for a novice. For this reason I strongly encourage focus group recordings to be professionally transcribed. Depending on the complexity of the recording, the \$1.99 per recorded minute figure I cite above may be an underestimate.

Analyzing Qualitative Data

In my experience researchers tend to really enjoy the process of analyzing qualitative data. Not only do they learn a lot about the thing they’re evaluating, it also is an opportunity to consider important contextual factors around the phenomenon of interest and to deeply understand peoples’ experience with it. However, this process can also be fraught with uncertainty and insecurity if this is new territory. To provide a concrete demonstration of how qualitative analysis typically goes, I’m going to use an example.

Example

Two faculty members and two residents collaborated on an interview study evaluating an intervention they gave to 4th year medical students. The research team conducted twenty 30-minute interviews of medical students for this evaluation. Then, they had the audio recordings professionally transcribed. So, the research team has all these transcripts – what should their first step be? How do they know that the 20 interviews they’ve done are “enough”?

The first step in the coding process will be for the four researchers to review 2-3 transcripts (the same 2-3) on their own. During this initial read through, each researcher will highlight areas in the transcript that they see as relevant to their research question(s). Based on the patterns the individual researchers observe, each one will come up with a list of potential codes, which are meaningful segments of data that are repeated throughout the data source.

When you are doing this initial coding process, you may come across something in the transcript that is of interest, but doesn’t repeat in your data. It is totally fine to make note of that interesting piece of data, but what you’re really looking for are concepts and ideas that seem to recur in the 2-3 transcripts you read through.

After this process is complete, all four researchers should get together to review the results of their initial readthrough of the 2-3 interview transcripts. Of particular interest in this meeting is discussion of the areas that most/all of the researchers highlighted as recurring concepts or ideas in the data (the codes).

When you’re doing this, you’ll likely find that there is a fair amount of overlap in what you and your research team found, but that you may use different words to describe the patterns you’ve identified.

This initial meeting is where you can hammer out all those differences and come to consensus on what words/language to use to define the codes.

At the end of this meeting, our hypothetical research team will have beginning draft of what will be their final codebook.

A codebook is simply a list of the codes (the repeating ideas or concepts mentioned by interview participants) and how you define them.

The next step is for the team to try and independently apply the drafted codebook to 2-3 additional interview transcripts. After this process the research team will meet again to address the following questions:

- 1.) Did the codes we initially defined based on the first set of transcripts we reviewed seem to also recur in these new transcripts?
 - a. If yes/no – is there a need to broaden or specify the definition of the code?
- 2.) Are their new patterns of responding that weren't present in the few initial transcripts you reviewed?
 - a. If yes, you will want to add these new codes to your codebook once the research team has reached consensus on how to define them
 - b. If you do add to the codebook, you will want to revisit the first 2-3 transcripts you coded with these additional codes in mind

After this second meeting, the research team will have a semi-final version of their codebook. Next, the two residents in our example research team will code the remaining interviews using the most recent version of the codebook. They will check-in with each other regularly to make note of any new codes that have emerged in the data and/or to discuss if the existing code definitions in the codebook need to be revised.

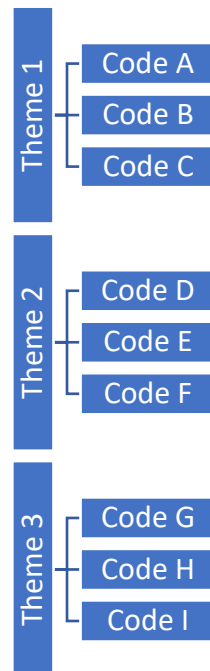
Once the coding is complete, the research team should meet again to discuss/finalize code definitions (if needed), and to assess if they have reached **saturation**.

Although there is no one size fits all rule about when saturation is reached, essentially it means that you have reached a point where you are not observing any new patterns of responding in your data.

If the two residents didn't have to make many (or any) additions to the codebook while coding the remainder of the interviews, there's a good chance the study has hit the point of saturation. However, if they found that they needed to add a lot to the codebook as they went along, that would be a sign to conduct some additional interviews. If it were me, I might start with another 3-5 interviews and then see what those data look like.

Once the entire research team feels good about their codebook and they are feeling confident that they have reached saturation, the next step is to organize the codes into themes.

The idea of codes versus themes can be confusing, try to think about them hierarchically with multiple codes nested within a theme.



Themes provide researchers with a way to talk about their findings more conceptually/at a higher level and provide a more manageable way for researchers to consider how interviewees' experience with the phenomenon of interest might be affected by various contextual factors and/or how these factors may be interacting/interrelated.

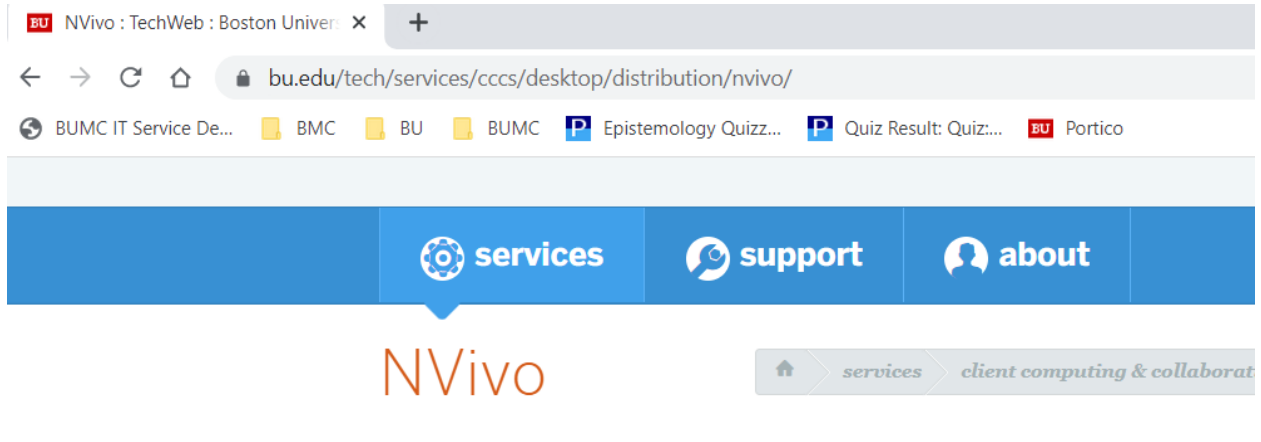
Interrater Reliability

There is not consensus on whether or not you need to calculate inter-rater reliability when you do qualitative coding. My personal opinion is that it should not be necessary – if you're interested in my TED talk on this topic please set up a meeting with me.

If you are in a position where you need to calculate inter-rater reliability for any reason, please just email me and I can talk you through the process.

Software for Coding

A commonly used program for qualitative coding is NVIVO. NVIVO is offered for free through BU. To install it onto your computer, follow along with these screenshots.



...scroll down on this page until you get to....

If you would rather not use the software for Mac OS X, you can use NVivo in a virtual installation of Windows using products such as [Bootcamp](#), [Parallels Desktop](#), or [VMware Fusion](#).

Download and Install

NVivo for Mac OS X and Windows can be downloaded and activated for free by active Boston University faculty, staff, and students. Please refer to these getting started guides on both [Windows](#) and [Mac](#) for additional help setting up NVivo.

[Download now](#) >

Renew/Replace a License

[Renew your license](#) >

The new license key is only available to those who have agreed to [the download terms](#).

Learn More

- > [Visit the NVivo Education Center](#)
- > [Contact the IT Help Center](#) if you have any questions or require assistance.
- > Sign up for a [training class](#)
- > View, print or download a PDF quick reference guide for [Windows](#) or [Mac](#)

After clicking “Download now” you will be prompted to enter your Kerberos/BU login and password. At that time, you can follow the instructions on the screen to download the program. Also note that IT offers trainings on NVIVO if you would like to learn that way. Youtube also has many good tutorials if that is more your style.

If you have a small data set, you may not need to use NVIVO. Although NVIVO does have a lot of neat functions that can help with data analysis, it is mainly a tool for data organization. If you can manage your data in Word or Excel, that is totally fine and equally legitimate as using NVIVO.