*African Journal of Inter/Multidisciplinary Studies (AJIMS)*
*Volume 7, Issue 2, pp. 1-12*
*e-ISSN: 2663-4589*

# Thin Data From Qualitative Interviews: On Probing for Quality and Relevant Knowledge

**Nomanesi Madikizela-Madiya and Zamakhosi Thina Shembe**

[1]**Department of Educational Foundations, University of South Africa, Pretoria, South Africa**

**Abstract:** There is an abundance of literature on the characteristics of deep and thick qualitative data, as well as the ways in which probing helps to elicit such data. However, there is limited engagement on thin data and its implications for quality in knowledge generation. In this article, we draw on our experience of a qualitative research project conducted in rural schools in South Africa, where interviews were conducted to explore how learners could be supported to perform better. We do not report on the findings from that research, but we highlight the pitfalls that rendered some of the data unusable due to its lack of depth. We provide examples of "thin data" and discuss its implications for the quality, credibility, and value of the research project. The paper argues that without sufficient probing or contextual detail, analysis risks becoming superficial, being unable to yield an interpretative and analytically generative account of the problem. The article contributes knowledge that will be useful to postgraduate students and inexperienced qualitative researchers. In addition to discussing the implications of thin data for the project, we provide possible ways in which research teams can avoid producing similarly thin data.

**Keywords:** probing; qualitative data; qualitative researchers; usable data

**CORRESPONDENCE**
Email: madiyn@unisa.ac.za

**EDITORIAL DATES**
Received: 09 July 2025
Revised: 14 November 2025
Accepted: 18 November 2025
Published: 31 December 2025



**DOI: https://doi.org/10.51415/ajims.v7i2.3120**

## Introduction

Data quality can sometimes be overlooked in discussions about academic research. Yet the quality of data is key to the success of research projects. In this article, we contribute to the discourse on research quality by examining the role of interviews in the generation of qualitative data. Seidman (2006) notes that the purpose of interviews is not just to evaluate or obtain answers to a research problem, but to understand people's experiences of a phenomenon and the meanings they ascribe to those experiences. Toma (2000, p. 179) uses the concept of "good data", arguing that it is "an evaluation of the product of the interaction between the researcher, the phenomenon, and the people under consideration, and the data being gathered". In other words, data from interviews should indicate a high level of interaction between the interviewer and the interviewee. Ruslin et al. (2022, p. 22) refer to the "*inter-view*", which they define as "an interchange of views between two (or more) persons conversing about a theme or a topic of mutual interest". This can occur in

a conversational manner, with the interviewer guiding the conversation to gain an in-depth understanding of the interviewee's experiences, hopes, and feelings in relation to the research questions. Data generated from such interviews is qualitative, thick, and reflexive (Gibson & Brown, 2009). Stahl and King (2020, p. 26) describe such data as that in which "readers would be treated to texts so rich in details that the event or the object of description is palpable".

While the role of interviews in generating qualitative data is irrefutable, the literature offers little insight into what we refer to as "thin data": data that lacks the necessary depth to address the research problem. Thin data is the information that lacks sufficient detail, context, or analytical richness to yield valuable material for analysis. It may be too general or superficial, limiting the study's capacity to generate valid insights, support critical analysis, or contribute to the formulation of evidence-based conclusions. Among the few researchers who have employed this concept is Morse (2009), who views thin data and writing from it as "cherry-picking". Morse (2009) points out that one of the dangers of cherry-picking is that the researcher may be entirely incorrect, as additional data might have highlighted missing traits, discovered hidden complications, and even diverted the analysis. In contrast, thick data refers to rich, qualitative information that captures the background, meaning, and emotions behind human behaviour. Thick data is qualitative, context-rich, and brought to light through human-centred methods that uncover the underlying reasons people behave in certain ways, providing in-depth insights into social meaning and human behaviour (Ang, 2019; Bornakke & Due, 2018; Wang, 2016). The literature emphasises that the interviewer plays a decisive role in producing data that has sufficient depth, but the opposite phenomenon – where certain practices are almost guaranteed to produce thin data – is hardly discussed. Van den Berg (2008, p. 180) notes that handbooks on qualitative research provide a variety of normative guidelines on how researchers should behave, "but neither the functioning of these guidelines nor the actual behaviour of interviewers is systematically studied". To this, we add that even in situations where the behaviour of researchers has been discussed, such as the need for probing, there are limited practical examples of the kind of data that may result from not following such guidelines. As such, inexperienced researchers have scant guidance on how to avoid wasting time in developing thin or minimal data, which sometimes happens.

We share our contextual experiences of having worked on a research project in South African schools, where some of the generated data was identified as thin, being unable to help answer the research questions. We present practical examples of such data, trusting that the knowledge we share may assist developing researchers. The article is presented in six sections. This introduction is followed by a review of the literature on this topic. The third section provides the context of the study, indicating where and how the research on which this article is based was conducted, and discusses probing as an analytical framework for the article. In the fourth section, we outline the steps taken to generate data for the project on which we reflect and briefly describe our approach to reflexivity in developing this article. The fifth section presents examples of thin data identified in transcripts of the interviews conducted for the original research project. We conclude by discussing the implications of thin data for the project and providing possible ways for research teams to avoid producing similarly thin data.

## Literature on qualitative interviews and data

Their ability to generate detailed information and offer insights on participants' feelings and opinions has made qualitative interviews the preferred data collection method in qualitative research (DeMarrais, 2004; Elhami & Khoshnevisan, 2022; Mwita, 2022; Roberts 2020; Stahl & King, 2020). The interaction between the interviewer and the interviewee has been identified as critical in determining whether the data from such interviews is of high quality. Interviewing skills are crucial in determining the quality of such interactions and generating rich data. Edwards and Holland (2013) note that the interviewer should be able to listen attentively to the interviewee's responses to ensure that the data given is responsive to the research problem. Similarly, Lavee and Itzchakov (2023) assert that when evaluating the quality of qualitative research, good listening should be considered a crucial basis and a main research measure in efforts to establish trustworthiness, credibility, and rigour. Learning how to listen actively is considered more important than getting through a list of predetermined questions during the interview. Roberts (2020) notes that well-known researchers such as Rogers, Piaget, and Freud all relied on silence and active listening techniques in their quest for knowledge. The ability to listen enables the interviewer to identify gaps in the provided information and to probe for clarity or depth. DeMarrais (2004, p. 52) argues, "Using interview questions and follow-up questions, or probes, based on what the participant has already described, the goal is to construct as complete a picture as possible from the words and experiences of the participant." In fact, the interview guide prepared by the interviewer is not necessarily "a standard protocol for an interview" (DeMarrais, 2004), but just a guide that works best with the addition of probing questions to

elicit details. Robinson (2023) identifies probing as central to developing thick data from qualitative interviews. He contends that "[i]f probes are used systematically throughout a research interview, the conversation between interviewer and interviewee moves into sensitive and salient areas of discussion that it is unlikely to access otherwise" (Robinson, 2023, p. 382). Therefore, the skill of the interviewer is crucial in ensuring that this probing contributes value to the process.

Apart from delivering thick data to the researcher, in-depth interviews that provide high-quality data are of value to the interviewees. Edwards and Holland (2013, p. 3) note that qualitative interviews provide an opportunity for both the interviewer and the interviewee to learn from the "interactional exchange of dialogue" between them. Where there is no such exchange, neither party may learn much from the interview. Interviewees should be empowered by the process; it should allow them to reflect on the research topic, express their perspectives, and develop their own insight into the topic. Gibson and Brown (2009) identify reflexivity, thick description, and naturalism as the characteristic features of qualitative data. They argue that reflexivity is a key issue in social research. Reflexivity refers to the process of reflecting on the role of the researcher in the construction of meaning and, critically, of data. Thus, "both researcher and researched bring with them concepts, ideas, theories, values, experiences and multiple intersecting identities, all of which can play a part in research interaction in the qualitative interview" (Edwards & Holland, 2013, p. 3). The researchers will have identified the research topic and, therefore, approach the interview with specific issues to discuss; they should, therefore, be strategic in practising reflexivity. Roberts (2020) argues that asking questions that are incapable of providing any value in relation to the research questions could result in useless data, wasting the participants' time and energy. Similarly, Rubin and Rubin (2012) urge that researchers ensure "thoroughness" when conducting interviews.

## Probing as a frame of analysis

Qualitative researchers agree on the importance of probing in generating thick and reflexive data through interviews and provide frameworks on how probing should be done. Jacobsen et al. (2025, p. 3) define probing as "… the asking of strategic follow-up questions that reveal deeper insights into participants' experiences, progressively uncovering potential hidden layers of participants' experiences or insights". Various types of probing may be used, as discussed below.

*Recapitulation and silent probes:* Alirezaei and Latifnejad (2020) identify the recapitulation probe and the silent probe, among others. The recapitulation probe returns the participant to the story after a digression, allowing the narrative to proceed. It is essential to do this, as interviewees often have a wealth of information to share and may easily stray from the research topic. The silent probe involves listening to the participant to draw meaning from what is being explicitly stated, as well as what may be implied or not stated (Alirezaei & Latifnejad, 2020). A researcher who is unable to listen attentively may miss important and useful information that could be drawn from the interviewee. However, when using the silent probe, the researcher should take care not to intimidate the interviewee. Jacobsen et al. (2025) state that "silence can create discomfort or confusion, leading to less authentic participant responses". Therefore, this form of probing may be enhanced using other forms as well.

*Descriptive probes :* Robinson (2023) notes that people's narratives, as presented during interviews, are a selection of episodes from their lives, and that these may be drawn from external events and actions or from their inner landscape of consciousness. Based on this understanding, Robinson (2023, p. 390) notes that descriptive probing may elicit "an in-depth description and contextualisation of an episode across outer and inner landscapes". Jacobsen et al. (2025) note that descriptive probing originates from narrative theory, which posits that narratives should be collaboratively co-constructed by the interviewer and the interviewee. Robinson (2023, p. 385) gives examples of questions that may be asked as descriptive probes, such as "Who else was present?", "When did it occur?", "Where were you?", "What were you doing at the time?" and "Could you tell me more about …?". To these, one may add the "why" questions, in which the interviewer seeks deeper information about the narrated situation. Robinson (2023, p. 385) notes that probes about the internal landscape may contain sensitive and subjective information, "and will be less likely to be divulged spontaneously". Examples of these kinds of questions are, "How did you feel at the time?", "What do you recall thinking about at the time?", and "What do you think that [pronoun or name] was feeling at the time?" Thus, Robinson suggests that the descriptive probe may also be used to ask about the subjective experiences of other individuals who are not necessarily present at the interview. These probes encourage the interviewee to think more deeply and provide information that might not have emerged had the interviewer not probed. In this way, they generate a deep and thorough understanding.

*Idiographic probes :* Jacobsen et al. (2025) note that ideographic probing is derived from the theory of autobiographical memory. It is used to "elicit detailed and specific memories as opposed to generic and broad recollections" (Jacobsen et al., 2025, p. 3). To explain this, Robinson (2023) argues that some interviewee information derives from past experiences, stored either as generic memories or specific memories. Probing should help the interviewee recall such memories to create rich data. Generic memories may be brought to light in statements such as, "I used to do [this and that]." Such memories do not provide specific information. Robinson (2023, p. 386) notes that the interviewer should be able to do "idiographic probing" to get the interviewee to provide specific memories that "are stored in narrative-like structures". Therefore, they "allow for being recollected and told in storied form more naturally than generic memories". Examples of such idiographic probing include, "You mentioned [X], could you give me an example of a specific time when that happened?" (Robinson, 2023, p. 387). In this way, the interviewee is prompted to recall specific events that occurred at particular times, rather than recalling only general memories.

*Clarifying probes:* The third domain in Robinson's (2023) meta-structure of theories governing interviews is self-disclosure theory, which requires clarifying probes. Robinson (2023) notes that some information is so sensitive that interviewees may not spontaneously disclose it until they trust the interviewer and are assured of privacy, confidentiality, and non-disclosure of their identity. Jacobsen et al. (2025, p. 3) describe personal information shared by participants as arranged in layers, "contingent upon trust and intimacy levels with the listener". Additionally, such information may be disclosed if the questions suggest that the participant will experience emotional improvement as a result of disclosure. Clarifying probes follow up on what the interviewee has said and involve asking questions that seek to understand the meanings of words, statements, or phrases. Such probes eliminate possible misconceptions and provide in-depth information. Robinson (2023, p. 390) provides examples of questions such as, "What do you mean by X?". The interviewer can also repeat a word or phrase used by the interviewee with a rising inflexion in the voice, suggesting a question.

*Explanatory probes:* Some probes aim to identify the perceived causes of events or phenomena in the participant's life or in the lives of others. Robinson (2023, p. 388) notes that causes may be attributed to others rather than to the speaker, which requires explanatory probes. From the above, it may be deduced that researchers may use a variety of probes to obtain thick and descriptive data. Probing can be a problem for novice researchers who "can derail the interview process by asking lengthy, closed, vague, or leading questions" (DeMarrais, 2004). In other words, novice researchers may be tempted to ask questions that lead to what they want to hear from the participant, seeking to "[confirm] their personal suspicions, thereby guiding the process in a way that validates their personal expectations instead of capturing the research participants' perspective" (Roberts, 2020, p. 3186). This practice can negatively affect the trustworthiness of the findings. However, Roulston (2010) affirms that where an interview appears to have "failed", inexperienced researchers should not despair, as "failed" interviews create fertile ground for asking methodological questions, which are seldom originally considered. The researchers involved in the project reported on in this article concur with Roulston (2010). Our experience of eliciting data that was too thin to be used helped us understand the role of the interviewer and recognise how our methodology in selected interviews distorted the purpose of the research project. Once the researchers were able to reflect on why they had obtained thin data, the meaning of quality in qualitative research became clearer.

## Methodology and context

The research on which this article is based was conducted in two rural provinces in South Africa: the Northwest and Limpopo. South Africa has a high learner dropout rate, especially of black learners from rural and township schools, where resources are insufficient. Grossen et al. (2017, p. 1) note that only 60% of the learners who enter the schooling system are recorded as completing Grade 12, while 40% "drop out of the system after repeated failure". The Department of Statistics South Africa (Stats SA, 2022) reports a worrisome school dropout rate in 2021 for various ages: 3% for 15-year-olds, 9% for 17-year-olds, 29.3% for 18-year-olds, and 46.3% for 19-year-olds. To address this problem, the Department of Basic Education (DBE) introduced and legislated the progression system. Progression is "the advancement of a learner from one grade to the next in spite of the learner not having complied with all the promotion requirements" (DBE, 2017). According to the National Policy Pertaining to the Programme and Promotion Requirements of the National Curriculum Statement (NPPPR), progression applies to learners in public schools who would otherwise be retained in a phase (either the Intermediate Phase (Grades 4–6), the Senior Phase (Grades 7–9) or the Further Education and Training Phase (Grades 10–12)) for more than four years. Since the inception of the NPPPR in 2013, Grade 12 results in some

provinces have declined, with the progressed learners counted among the contributors to the decline (Mogale & Modipane, 2021). Thus, this research project was launched to understand how to support progressed learners and those at risk of not progressing. The research team comprised both experienced and novice researchers. In this article, we do not present the findings of the research project's questions but reflexively report on and discuss issues related to the quality of the data and the reasons for variation in quality.

The project followed a participatory action research design. Participatory action research is a cyclic or collaborative approach to research that involves researchers and participants working together to identify a problem, collect and analyse data, and take action to bring about change. This process occurs in stages, from the identification of the problem to the next cycle, until all parties have exhausted the possible actions. Baum et al. (2006) contend that at the heart of participatory action research is the collective, introspective inquiry that researchers and participants undertake to comprehend and improve upon the practices in which they participate and the situations in which they find themselves. Payne (2017) asserts that participatory action research offers a way to capture the phenomenology of the population of focus, an aspect of analysis that is commonly neglected by community interventionists. This paper reports on the thin data generated from the first phase of the participatory action research project. This phase consisted of four sub-questions, each focusing on the potential role of a specific category of stakeholders in addressing the problem of learner dropout, including principals, teachers, circuit managers, learners, and parents.

The main research question was "how can a research project contribute to supporting the progressed learners and the learners at risk in Bahananwa and Maamusa circuits?". Supporting questions include:

- How can the circuit manager facilitate the programme of supporting the identified learners?
- How can teachers in the circuit assist the progressed learners and learners at risk?
- How can the school management teams support the project?
- What role can the researchers in the project play in supporting the stakeholders and the learners?

From the nine provinces in South Africa, two were sampled (Limpopo and Northwest) due to convenience, as the lead researchers had prior experience working in these provinces. Five schools in Limpopo and two in the Northwest participated. More schools in Limpopo were underperforming than in the Northwest and therefore formed the larger part of the sample. A total of 99 participants were involved in the project. The sample in each province comprised two curriculum advisers, one circuit manager and six principals – four from Limpopo and two from Northwest. Eight focus groups of learners were interviewed. These learners were identified by their teachers as either progressed learners or learners at risk, based on their current performance. The teachers, principals, and circuit managers were interviewed individually through semi-structured interviews. The research team members divided themselves into teams of two per category of participant. Two focus groups of parents were interviewed. The interviews were recorded with the participants' consent, and all were assured of anonymity and confidentiality. The project, involving 12 researchers, was ethically cleared at the researchers' institution. Permission to conduct interviews was obtained from the Departments of Education of the two provinces. Different sets of semi-structured and unstructured interviews were designed for each participant category. The project team members arranged to conduct their research during school holidays to avoid interfering with the normal school programme. Interviews were held at identified centres where extra classes had been arranged for matriculants. When the interviews were completed, the collected data was transcribed verbatim and then translated into English.

After the data was transcribed, the researchers met to analyse it and identify emerging themes in relation to the research questions. When going through the data, an experienced researcher found that some researchers had not probed some participant responses. This lack of probing led to thin data, which could not be blamed on the participants but on the limited interview skills of the researchers. The transcripts with these limitations were studied separately to develop this article. Thus, this article is based on the practice of reflexivity (Day, 2012) in relation to the project's data generation. It provides examples of how the method of data generation compromised the quality of the data and hence the rigour of the research findings. Thin data was identified from the interviews with all the categories of participants. Such data resulted from the reviewers' failure to follow up on or probe the statements made by the interviewees, and due to asking questions that seemed to be based on assumptions. The presentation of findings points to each instance in the examples provided below where thin data is identified. Furthermore, a suggestion is made in each instance regarding how the probing should have been done. The first

category of participants that we interviewed was teachers. Our aim was to elicit their views on how best to support the progressed and at-risk learners.

## Findings

This section presents the findings on teachers' views on learners' support, interview with the curriculum coordinator, views of the learners and views of the circuit manager.

### *Teachers' views on learners' support*

In one interview, two researchers asked a teacher about the subjects that were most challenging for progressed and at-risk learners. This was a good question for establishing a sense of where support was most required. The teacher listed the subjects as Geography, Mathematics, Agriculture, and Life Sciences. To follow up on this response, the following exchange ensued:

> **Researcher 1:** In Geography, what sections are challenging and why?
>
> **Participant:** Geomorphology, map-work, and climate and weather.
>
> **Researcher 2:** In Mathematics, what are the challenging sections?
>
> **Participant:** Trigonometry.
>
> **Researcher 2:** Challenging sections in Agriculture?
>
> **Participant:** Genetics, animal reproduction, animal nutrition, photoflood.
>
> **Researcher 2:** Challenging sections in Life Sciences?
>
> **Participant:** Genetics.

Researcher 1 posed a question comprising two parts, with the second part having the potential to facilitate a descriptive probe (Robinson, 2023). However, the interviewee did not answer this part of the question, and the researcher did not follow up on it, thereby rendering the data thin or non-descriptive. Overall, from this conversation, no useful knowledge was developed on addressing the question of how the teachers could support progressed and at-risk learners. The conversation yielded thin data, a direct result of a lack of probing.

> In another interview, the following conversation took place:
>
> **Researcher:** Do you have progressed learners and learners at risk in this school?
>
> **Participant:** Yes, we do have them.
>
> **Researcher:** So, when you separate them from those who are doing well or relatively well, do they feel excluded?
>
> **Participant:** Sometimes they do feel like they are segregated.
>
> **Researcher:** What factors lead to poor performance of learners?

This researcher started well by asking whether there were progressed learners in the school, even though this question was redundant considering the purposive sampling of the school. However, two issues may be identified that limited the value of answers in this exchange. First, the question, "So, when you separate them from those who are doing well or relatively well, do they feel excluded?" is baseless because there is no indication that the teacher separates the progressed learners from those who are doing well. In this case, it can be said that the researcher worded the question based on an assumption. This confirms Roberts's (2020) assertion that researchers may direct the outcome of interviews to align with their personal expectations; it is an example of cherry-picking, as described by Morse (2009) in relation to thin data. Also, the teacher could not be asked about the feelings of the learners because feelings are subjective, and only the person who has them can explain them. A better way of wording the question might have been, "How do you think the learners feel when you separate them …?" Even this question, however, could only have been asked after the researcher had established that the

teacher, in fact, separates these learners. As it stands, this is another example of a question that yielded thin data because it deviates from the conventions of generating knowledge through effective questioning.

Another opportunity to elicit thick data was missed when the researcher simply accepted the answer of "Sometimes they do feel like they are segregated …" without asking a follow-up question. The interviewer could have probed the response to understand how the teacher knows that the learners feel this way. Instead, the interviewer moved swiftly on to the next question. A follow-up question could have led to further inquiries about how the teacher could best support learners who feel isolated. In this case, the data was thin, lacking in detail, and was compromised in its ability to support the development of strategies for the learners concerned.

## *Interview with the curriculum coordinator*

The role of curriculum coordinators is to supervise subject advisers and to ensure that teachers are supported in implementing the curriculum plan. One of the coordinators interviewed indicated that he also ensures that poor-performing schools and learners are supported. Once this had been established, the interview proceeded as follows:

> **Researcher:** How is the reception when you visit poor-performing schools?
>
> **Participant:** It all depends on the principal. If he or she is welcoming and supportive, then the teachers and learners are most likely to welcome and cooperate with us. There are principals who do not appreciate us when we come unannounced, but we have to do our jobs.
>
> **Researcher:** What was last year's passing percentage?
>
> **Participant:** 72%, but in some other districts, there are schools that negatively affect the final percentage, because such schools would be underperforming, while others perform well.
>
> **Researcher:** What do you think of the progression policy?

Again, the researcher did not probe to understand how non-appreciation on the part of teachers was demonstrated. This was an opportunity to conduct ideographic probing (Robinson, 2023), where the participant would be asked to provide examples of when this happened, describe their feelings, and explain how they dealt with the situation. This would have provided important information for this study because non-appreciation would almost certainly constrain the coordinator's ability to ensure the implementation of the curriculum plan and to assist these teachers in supporting the learners concerned. Therefore, the answers given constitute thin data. In addition, the statement of "… there are schools that negatively affect the final percentage …" remained unexplained, since the researcher moved on to the next question without following up with descriptive probing. The question evoked only generic memories, rather than specific ones, with no prompts provided for the provision of examples, as suggested in the probing taxonomy (Robinson, 2023, p. 387). Therefore, no information was given that could have yielded insight into the processes that schools and the district use to support struggling learners. Nevertheless, the research team had the opportunity to gain further insight by interviewing the learners and understanding how they perceived support.

## *Interviews with learners*

In one of the focus group interviews, a learner indicated that there was no Mathematics teacher at her school, and therefore she was taught this subject only on weekends when she went to a local support centre for learners from various schools.

> A researcher conducted the following interview with several learners in the group:
>
> **Researcher:** So, you are taught by a different teacher [at the centre]?
>
> **Participant 1**: Yes.
>
> **Researcher:** So, how is the teacher at the centre?
>
> **Participant 1:** It's better when I'm here, but somehow, we do not understand because the English is too high.

> **Researcher:** Anyone [else]? [Indicating that other learners should respond to the same question].
>
> **Participant 2:** I understand, but when test time comes, it becomes difficult.
>
> **Researcher:** Are the questions difficult?
>
> **Participant 2:** Yes.
>
> **Researcher:** So, what do you do about this challenge?
>
> **Participant 2:** I try to understand the teacher.
>
> **Researcher:** Another one? *[Indicating that other learners should respond to the same question]*.

There is nothing in this interview to suggest that the learners and the researcher learned anything of value from it. The interactional exchange, as suggested by Edwards and Holland (2013), was not effective enough to lead to learning. The researcher left the responses hanging, simply asking other learners to answer rather than probing any of the answers given. For example, the researcher did not ask about the meaning of "… the English is too high". Did this learner mean that the teacher used complicated English words when teaching, or did she mean that the mathematical concepts were hard to understand when delivered in English? It would have been important to know what was meant, so that the researcher could report back to the teacher participants, and they could address the issue. Additionally, when a learner said, "I try to understand the teacher," the researcher did not ask how the learner attempted to do so. With this omission, the project could not generate insights into how learners who try could either improve their methods of trying or how they might be supported to try differently.

> The same interview proceeded as follows:
>
> **Researcher:** Another learner with a challenge?
>
> **Participant 3:** My challenge is Geography. I hear the teacher, but the way the question papers are written confuses me. The English confuses me.
>
> *[The researcher signals for other learners to respond to the same question..]*
>
> **Participant 4:** Maths Literacy; my teacher doesn't teach well.
>
> **Researcher:** Does the teacher come to class?
>
> **Participant 4:** Yes.
>
> **Researcher:** So, as students, what are your weaknesses?

This is another example of an exchange yielding thin data, since the researcher missed the opportunity to gain important information about how these learners could be supported. First, if Learner 3 understood the teacher but found the English in the question papers confusing, did this mean that the English in the question papers is different from that of the teacher? How could this be rectified, so that the learner receives adequate support? This information was not generated. Second, the meaning of "… my teacher doesn't teach well" was not followed up on. Thus, even if the teacher concerned were aware of this response, she would not know how to support the learner. Yet this learner appeared to trust the researcher, since she was comfortable sharing her concerns about the teacher and might have been willing to explain further had she been asked. The researcher had the opportunity to give clarifying probes for the meaning of what the learner said (Robinson, 2023, p. 388). Instead, the lack of probing generated thin data.

In another group interview, the researcher asked, "You as a person, what is your weakness?" The participant responded, "I do not create time to write my work because I spend more time on my phone." With this admission, the researcher could have probed to seek clarity on what the learner did on the phone, or perhaps what the learner thought they should be doing differently with the phone for their own benefit.

Also, among other questions asked was, "How much time do you allocate yourself to study?" The participant answered, "Two hours." The researcher then moved on to a different question. This generated thin data, defeating the purpose of asking the question, as it was not followed up with a self-disclosure probe (Robinson, 2023). For example, how are the two hours utilised? Are these hours dedicated to one or more subjects? What are the activities done during each part of the two hours? These questions might have provided valuable information on how the learner could be supported.

### *Interviews with the circuit managers*

As the authorities responsible for ensuring that learners and teachers are supported, circuit managers were expected to provide valuable information for the project. However, shortcomings were identified in some of the interviews, as shown below.

> **Researcher:** What would you say are the differences in schools regarding the issue of progressed learners?
>
> **Participant:** There are schools with learners who are willing to learn, and yet others are the total opposite. So, that creates differences in terms of others being progressed due to their failure. The culture of learning is non-existent, and this may be caused by the community that does not show value for education.
>
> **Researcher:** What kind of support do you think these learners need?

The first part of this response needed an explanatory probe (Robinson, 2023). The participant could have been asked to elaborate on the meaning of "learners willing to learn" and "the opposite". The identification of the willingness to learn and its opposite was not established. It is possible that, had this information been given, the indications of willingness to learn could have been used to support those learners who displayed the opposite tendency. However, the inability to probe produced data too thin to be of any value in this way. Also, the participant mentioned that there was "no culture of learning". The meaning of this concept should have been sought through a follow-up question of what the manager thought could be done to foster a strong learning culture. The manager also referred to the community as not valuing education. Since the community is composed of parents, peers, and others, the researcher could have probed for clarity about who did not value education and how they demonstrated this. Such clarity might have led to strategies that could foster a culture of learning in the schools concerned.

The above examples illustrate poorly developed interviewing techniques that yielded thin data during the project. Below, we discuss these findings to draw conclusions.

## Discussion and conclusion

We began this article by discussing the meaning and significance of depth in qualitative data and the role of interviewers in eliciting such depth. We referred to Seidman (2006) and Borgman (2012), who argued that the purpose of qualitative interviews is to understand people's lived experiences and to uncover the meanings they ascribe to these experiences. Personal experiences may be understood if the interviewer–interviewee engagement is interactive. The interviewer should drive the process by asking probing questions and being conversational (Robinson, 2023; Ruslin et al., 2022; Toma, 2000). We have shared examples of interview exchanges that yielded thin data, since they did not align with the descriptions of qualitative interviews given above. In this research project, some data was of little value to the communities concerned because it lacked detail and depth on how progressed and at-risk learners could be supported. As such, the data was not particularly responsive to the research questions and could not easily be analysed to address them (Kandel et al., 2011). The data did not treat the readers to "texts so rich in details that the event or the object of description is palpable", as suggested by Stahl and King (2020, p. 26). In other words, the shared data suggested that no detailed or intense interaction occurred between the researchers and the interviewees, resulting in thin data that did not meet Toma's (2000) definition of good data.

This article confirms the need for researchers to possess good listening skills (Alirezaei & Latifnejad, 2020; Lavee & Itzchakov, 2023) and to use their listening intentionally, based on the purpose of their research project. The aim of this project was to develop knowledge on how learners could be supported to perform better. The interviewers should have listened with the intention of gathering information to achieve this aim.

The interviewing guidelines provided by Robinson (2023) and Jacobsen et al. (2025) should have guided these researchers, enabling them to draw deep-seated information from the interviewees rather than generic information, by using a variety of probes. From well-worded probes, the interviewers could also have enabled the interviewees to reflect and learn from the interviews. In this way, Robinson's concept of the co-construction of knowledge might have become a reality. It may be concluded that listening and probing skills work together in qualitative research. In this project, some interviews failed due to the interviewers' lack of skills and experience. However, new questions emerged, and lessons were learned (see Roulston, 2010). These questions pertained to the interviewing and probing processes, and the lessons learned informed the development of this article. We shared insights on thin data, which does not help address a research question. The examples provided demonstrate that, even when data is drawn directly from participants, it may still not constitute good data and the interview may still not meet the requirements of a good qualitative interview, as defined in the literature (DeMarrais, 2004; Elhami & Khoshnevisan, 2022; Stahl & King, 2020). However, the findings are valuable pointers for guiding inexperienced researchers. They support Robinson's assertion that being able to use probes effectively should be regarded as a skill and that qualitative researchers should strive to develop this skill.

Four conclusions may be made from this article. First, probing is central not only to developing thick data from qualitative interviews, but also for avoiding thin, unusable answers to the research questions. Without thorough probing, valuable information remains concealed, and the research project may struggle to achieve its objectives. Second, it is essential that research teams have diverse levels of experience to facilitate cross-learning and ensure the success of a research project. Had this been a team of novice researchers only, the misfortune of having to use the shallow and meaningless data would have resulted in a less valuable study, based mostly on assumptions. Third, active and intentional listening (Roberts, 2020) is indeed more important than getting through a list of predetermined questions during an interview. Without this ability, gaps in the provided information cannot be identified. Thus, from the new questions we began asking ourselves as we engaged with the data, we concluded by concurring with Gibson and Brown's (2009, p. 1) statement that "[t]he 'success' of a research project is very much contingent on […] working with data to achieve something interesting and perhaps even important in relation to the substantive focus of a research project" (Gibson & Brownn, 2009, p. 1). We reflected on the process of our research and realised that more rigorous modelling of interviewing techniques should have been done. For instance, we could have conducted a pilot study in which we guided inexperienced researchers on how to probe. In qualitative research, probes cannot be developed ahead of interviews because they arise from specific statements made by interviewees and should be generated naturally during the interviews or conversations. This article demonstrates how thin data can be avoided through probing, as well as the significance of modelling probing for novice researchers.

**Declarations**

**Interdisciplinary Scope:** This article presents practical examples of thin data from a qualitative research project that was conducted in rural schools in South Africa. The paper defines and identifies different types of thin data, including that which lacks the necessary depth or detail for analysis, the superficial one, and the general one, which is irrelevant to the problem at hand. The examples provided in the article are useful to students and early-career researchers in education, as well as across various disciplines.

**Author Contributions:** Conceptualisation (Madikizela-Madiya), literature review (Madikizela-Madiya and Shembe), methodology (Shembe), analysis (Madikizela-Madiya), review and editing (Madikizela-Madiya and Shembe). The final version of the article has been approved by all the authors.

**Conflict of Interest:** The authors have no conflict of interest to declare.

**Funding:** There was no funding for the publication of this article.

**Availability of Data:** The article includes all the relevant data. However, more information can be obtained from the corresponding author upon reasonable request.

## References

Alirezaei, S., & Latifnejad, R. (2020). Reflections upon the conduct of phenomenological interviews. *Journal of Midwifery and Reproductive Health*, *8*(1), 1986–1987.

Ang, Y. Y. (2019). *Integrating big data and thick data to transform public service delivery*. IBM Center for The Business of Government. https://www.businessofgovernment.org

Baum, F., MacDougall, C., & Smith, D. (2006). Participatory action research. *Journal of Epidemiology and Community Health*, *60*(10), 854–857. https://doi.org/10.1136/jech.2004.028662

Bornakke, T., & Due, B. L. (2018). Big–Thick Blending: A method for mixing analytical insights from big and thick data sources. *Big Data & Society*, *5*(1), 1–16. https://doi.org/10.1177/2053951718765026

Day, S. (2012). A reflexive lens: Exploring dilemmas of qualitative methodology through the concept of reflexivity. *Qualitative Sociology Review*, *8*(1), 60–85. https://doi.org/10.18778/1733-8077.8.1.04

DBE. (2017). *National policy pertaining to the programme and promotion requirements of the National Curriculum Statement: Grades R–12*. www.education.gov.za/Portals/0/Documents/Policies/NATIONAL%20POLICY%20NPPPR%20Dec%202017.pdf?ver=2018-05-10-110455-677

DeMarrais, K. (2004). Qualitative interview studies: Learning through experience. In K. DeMarrais & S. D. Lapan (Eds.), *Foundations for Research: Methods of Inquiry in Education and the Social Sciences* (pp. 51–68). Erlbaum.

Edwards, R., & Holland, J. (2013). *What is Qualitative Interviewing*. Bloomsbury Publishing. https://doi.org/10.5040/9781472545244

Elhami, A., & Khoshnevisan, B. (2022). Conducting an Interview in qualitative research: The modus operandi. *MEXTESOL Journal*, *46*(1), 1–7. https://doi.org/10.61871/mj.v46n1-3

Gibson, W. J., & Brown, A. (2009). *Working with Qualitative Data*. Sage. https://doi.org/10.4135/9780857029041

Grossen, S., Grobler, A. A., & Lacante, M. (2017). Repeated retention or dropout? Disputing Hobson's choice in South African township schools. *South African Journal of Education*, *37*(2), 1–11. https://doi.org/10.15700/saje.v37n2a1367

Jacobsen, R. M., Cox, S. R., Griggio, C. F., & Berkel, N. (2025). Chatbots for data collection in surveys: A comparison of four theory-based interview probes. *arXiv Preprint arXiv:2503.08582*. https://arxiv.org/abs/2503.08582

Kandel, S., Heer, J., Plaisant, C., Kennedy, J., van Ham, F., Riche, N. H., Weaver, C., Lee, B., Brodbeck, D., & Buono, P. (2011). Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization*, *10*(4), 271–288. https://doi.org/10.1177/1473871611415994

Lavee, E., & Itzchakov, G. (2023). Good listening: A key element in establishing quality in qualitative research. *Qualitative Research*, *23*(3), 614–631. https://doi.org/10.1177/14687941211039402

Mogale, M. L., & Modipane, M. C. (2021). The implementation of the progression policy in secondary schools of the Limpopo province in South Africa. *South African Journal of Education*, *41*(1), 1–10. https://doi.org/10.15700/saje.v41n1a1853

Morse, J. M. (2009). "Cherry picking": Writing from thin data. *SAGE*, *20*(1), 3–3. https://doi.org/10.1177/1049732309354285

Mwita, K. (2022). Strengths and weaknesses of qualitative research in social science studies. *International Journal of Research in Business and Social Science*, *11*(6), 618–625. https://doi.org/10.20525/ijrbs.v11i6.1920

Payne, Y. A. (2017). Participatory action research. *The Wiley-Blackwell Encyclopedia of Social Theory*, *20*, 1–5. https://doi.org/10.1002/9781118430873

Roberts, R. E. (2020). Qualitative interview questions: Guidance for novice researchers. *The Qualitative Report*, *25*(9), 3185–3203. https://doi.org/10.46743/2160-3715/2020.4640

Robinson, O. C. (2023). Probing in qualitative research interviews: Theory and practice. *Qualitative Research in Psychology*, *20*(3), 382–397. https://doi.org/10.1080/14780887.2023.2238625

Roulston, K. (2010). Considering quality in qualitative interviewing. *Qualitative Research*, *10*(2), 199–228. https://doi.org/10.1177/1468794109356739

Rubin, B. H. J., & Rubin, I. S. (2012). Continuing the design: Making the research credible. In B. H. J. Rubin & I. S. Rubin (Eds.), *Qualitative interviewing: The art of hearing data*. (2nd ed.). Sage Research Methods.

Ruslin, R., Mashuri, S., Rasak, M. S. A., Alhabsyi, F., & Syam, H. (2022). Semi-structured Interview: A methodological reflection on the development of a qualitative research instrument in educational studies. *IOSR Journal of Research & Method in Education*, *12*(1), 22–29.

Seidman, I. (2006). *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. Teacher College Press.

Stahl, N. A., & King, J. R. (2020). Expanding approaches for research: Understanding and using trustworthiness in qualitative research. *Journal of Developmental Education*, *44*(1), 26–29.

Stats SA. (2022). Increase in number of out-of-school children and youth in South Africa. https://www.statssa.gov.za/?p=15520

Toma, J. D. (2000). How getting close to your subjects makes qualitative data better. *Theory Into Practice*, *39*(3), 177–184. https://doi.org/10.1207/s15430421tip3903_9

Van den Berg, H. (2008). Reanalysing qualitative interviews from different angles: The risk of decontextualization and other problems of sharing qualitative data. *Historical Social Research*, *33*(3), 179–192.

Wang, T. (2016). Why big data needs thick data. *Ethnography Matters*. https://ethnographymatters.net/blog/2016/01/20/why-big-data-needs-thick-data/