Chapter 6

DATA CREATION AND DATA ANALYSIS

Chapter 6

DATA CREATION AND DATA ANALYSIS

> *"Research and policy, rather than privileging lived experience and people's agency, override them as being too subjective and irrelevant. When it is claimed that rural people are awaiting development to rescue them from the slumbers of pre-modernity, let us ask whose development we are talking about."*
>
> (Ngonini, 2004:15 in Zafar, 2004)

Introduction

In the data creation process I tried to be aware of the issues of foregrounding participants' experiences as mentioned by Ngonini, so that data collection became an 'organic' process. One consequence when faced with conflicting choices of recording data or engaging 'normally' with participants, the choice seldom came down on the side of the data or research. Note pads and tape-recorders gave way to memory and journal entries; systematic probing of ideas gave way to conversation and decision-making.

Following Reason's (1993) argument, which runs counter to the acquisition mode of data 'collection', I too prioritised the other purposes of research. Reason (1993) asks what the primary purpose of our time is and how our research is aligned with this. For me the answer is at least aligned to social justice and alleviating the concerns of participants: this cannot be achieved when their own immediate needs are conveniently repressed for the sake of a preconceived research agenda.

Data Creation

I present here a note showing an example of data collection in one day, in the one school at the start of the project:

> 6th September (See Data LS1, LS2 & LT1)
>
> Meeting with students – Lusiba (35 Grade 7 students -aged 13-17).
> With the help of the principal I presented the idea of the project: recapping also on the previous day's community meeting. I asked them to draw a "Picture of my life" and write down a short story of "Something someone taught you" and 3 questions they would like answers to.
> In the meantime teachers brainstormed writing down students' "Experiences and talents" "What would they like to learn?" There was a very lively discussion: 7 women teachers and 1 man.
> I then had an interview with the principal and Grade 7 teacher – before going on to the Sinevuso Secondary School. In the evening I chatted to the Woza Moya manager.

In this example I had a number of purposes in mind:

- Presenting information about the project
- Sharing responses from the previous meeting
- Modelling learner-centred classroom strategies
- Exploring students' contexts and concerns. This could be used for curriculum design
- Encouraging groupwork with teachers and building relationships
- Seeking guidance and input from the principal and grade 7 teacher
- Providing safe and fun spaces for all voices to be heard
- Expanding my understanding etc

I present some examples of how this snapshot of data was analysed in the next section 'Data Analysis'.

In presenting this array of often serendipitously collected data, academics expressed concerns about the research focus, framework, rigour, validity, 'science' content – to

mention a few. In answer to this I draw on Arundnati Roy who, in a talk in South Africa, cautioned that academics not become 'intellectual technicians':

> "*Academics have been frozen into their disciplines and literatures. They need to take their research into the streets, the fields, the dam walls…(and) Thinking carves up reality".* (Roy, 2003)

One way of managing conflicting agendas has been to be very aware of events as they unfold. So when I say the data collection was often serendipitous this does not mean that it was not planned. It does mean however that plans very often changed at short notice. So, for example, I would visit the school with interview questions planned for the teachers only to hear that they were at a workshop so I found myself with a class of Grade 7s for two hours. I nevertheless recorded the activities of the classroom visits which then became data. After that I might have lunch with an NGO worker and give the principal a lift into town where we would have time for casual conversation etc.

Making use of such a range of data enriches understanding and provides insights into the processes themselves more deeply than one research strategy might do. For example, I used questionnaires with students to elicit interests in science and this provided useful information, yet there were gaps and inconsistencies that required further discussion. Questionnaires given to primary school students as homework assignments were less successful. In this case it seemed that the teachers and principal were very keen that students gave 'the right answers', as well, the issue of language was a problem. Questionnaires for teachers provided more limited data than interviews or discussions. I arrived at a similar conclusion to Kann and Quarmby (1990:43) in Botswana:

> *"…the fact that structured questionnaire approach works in time-conscious societies does not mean that it works equally well, or is the most relevant approach, for African societies that value the quality, and sometimes the quantity, of interpersonal communication more than they value time."*

Having created a wide range of data, the categorising process allowed for me to identify gaps that needed further exploration. Also, as understandings and theories developed, I could test them through discussion.

Categorising Data

A list of over 100 sets of data is given in Appendix A.

These have been sorted, for convenience into data from:

1. Students – Sinevuso SS
2. Students – Lusiba SL
3. Teachers T
4. Principals P
5. Community C
6. Other O
7. Researchers R

The types of data include: questionnaires; interviews and focus group interviews; field notes; journal entries; reports from researchers, principals and NGOs; classroom observations; community visits; community workshops and meetings; visits to neighbouring projects; Mbumba committee meetings and minutes; classwork; homework assignments; photographs; student photo assignments; drawings; songs; stories; histories; posters; plays; models; informal conversations; science festival week and festival day; evaluations; NGO CD and Annual Report. Some of these were tape recorded or videotaped. Where respondents wished to communicate in Zulu, transcripts were translated.

Difficulties of data collection

As well as difficulties with planning and language, a number of other issues became evident. With regard to the planning, I feel disinclined to label the frequent changes to plans as a feature of an 'unstable' environment. 'Unstable' has connotations of being dysfunctional or crazy as in 'unstable person'. This is far from the case in this community. An example is a community meeting that may be arranged for 9am. By 9:30 am there is no-one there. Then the *Induna* simply (or impressively) stands in the sandy road while he chats to me. I ask him if anyone will come. He says: "They will come now." Within half an hour the word has gone around that the *Induna* is expecting people at the school. All this while he has not said a word to anyone except to me. This cannot be described as an 'unstable' community. There were eventually 40 people at that meeting: all enthusiastically contributing while the *Induna* sat quietly on the side.

Other difficulties included: shyness of respondents, my own not wishing to intrude, radically disparate opinions, respondents' wanting to give 'right answers', managing the process and competing agendas. Table 6.1 summarises some of the strategies we used to overcome these.

Table 6.1 Difficulties of data collection

Data collection difficulties	**Strategies to address these**
Other priorities, competing agendas	Field notes; journals; accept other outcomes
Reticence of respondents	Prolonged engagement; informal methods; non-verbal inputs, culturally resonant communications
Wanting to give the 'right answer'	Informal methods; awareness of this; assurances to teachers/students
Invested perspectives	Checking; awareness
Disparate opinions	Checking; holding ambiguity
Not wanting to intrude	Seeking advice and guidance from community researchers; asking permission

Examples of these issues include students being very shy to ask questions, especially in the beginning. Later in the science festival students were far more prepared to express themselves confidently. The following caution of Fleer and Robins applied to our interactions:

> *"When young children are unfamiliar with adults who are talking to them they are likely to be more reserved, and suspicious of the motives behind any questions: "Is there a 'right' answer that is expected of me?" (*Fleer and Robbins, 2003:417)

Even friends in Woza Moya became anxious when we took notes or used a tape recorder, on one occasion, prefacing a remark with: "I don't know if this is what you want?"

Principals sometimes told us things that could not be backed up by our observations. Some of our researchers became very set on 'producing a product', and accepted comments and statements at face value. Facts come in a variety of categories: information, misinformation and disinformation. Even while building relationships and

checking understandings it is well not to consider everything as simply ‘information’: it is, interpretation, perception, and context, all operating as pressures on the research. It is helpful to be aware of personal influences – noting that they themselves are ‘data’! It is also necessary to acknowledge personal moods, feelings and motivations, as well as to check understandings with other participants.

Sometimes supposedly ‘measurable facts’ were all but impossible to come by with consistency. For example, we struggled to get ‘exact’ student enrolment figures, teacher and student absentee rates, fees policies. The HSRC (2005) report also discusses discrepant data, for example school fees: in some instances parents say fees are between R50 and R100 while principals say they are less than R50. Other examples were information on AIDS infections, income levels and gender abuse. Gordon reports similar contradictions in her rural study which she describes as:

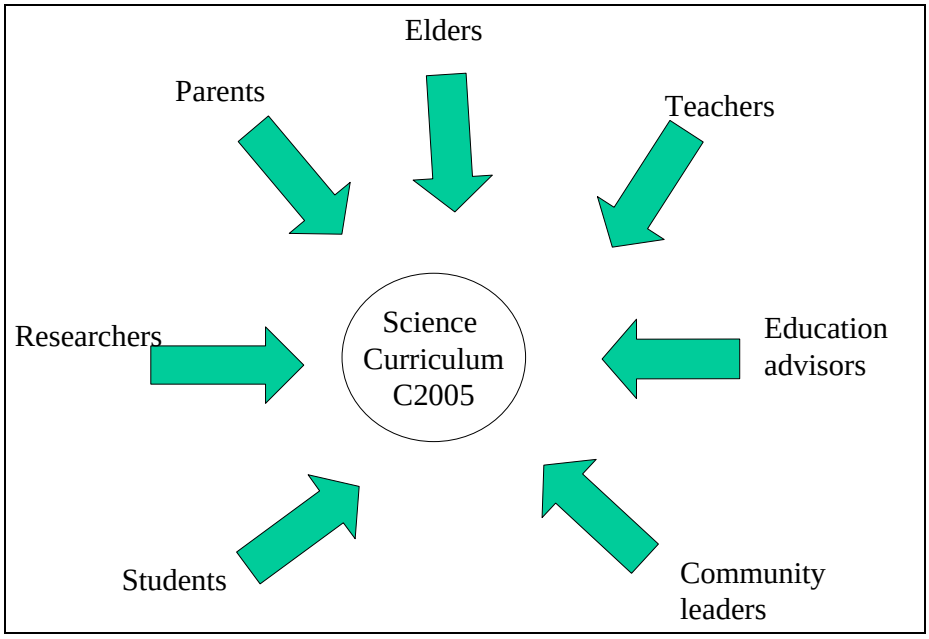
> *“…an indication of the difficulties facing urban researchers who attempt to interpret the underlying dynamics of rural lifestyles. It is possible that our understanding of rural visions is prosaic as our reflections are shaped by urban experiences, even though some members of the research teams were born in rural areas.”* (Gordon, 1997:9)

These types of discrepancy were also noted by Samuel (2002a:14). For example nearly 80% of the learners he interviewed said that their parents bought newspapers or magazines, while parents indicated the frequency was between 13% and 23%.

Data and project management

An important aspect of data creation was its integration with projects and curriculum design. I present an example of this process and how understanding evolved through iterative processes of discussions.

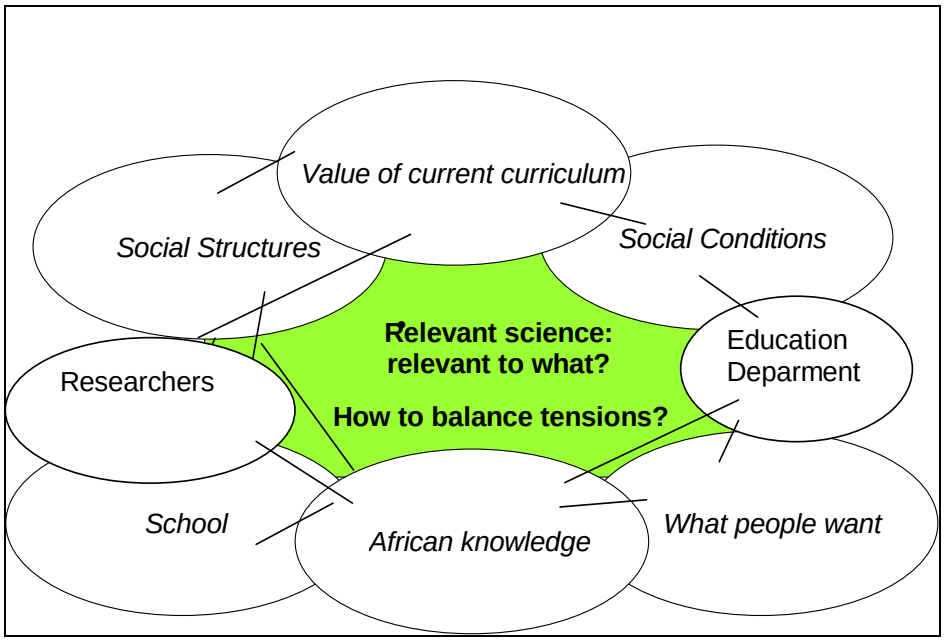
Figure 6.1 Curriculum Model 1



I used the diagram in Fig 6.1 at an early community meeting to explain participation in the envisioned curriculum design process. In some ways the categories held in the data analysis of inputs, but the actual process was far less neat, and so was the resulting 'curriculum'. Ironically it was the need for research data that pushed me out of a familiar curriculum-design mode into a more organic and chaotic process. I say ironical as at the time I thought the research processes would lead more or less directly (as Fig 6.1 might indicate) to a concrete curriculum package. I almost missed the importance of the *process* while thinking of *input* too readily as a step that precedes design! The diagram suggests a straightforward, unidirectional process, to a neat goal that falls centrally in the lap of C2005! The constituent contributors exist in equally balanced and distinct domains.

We moved into a 'visioning the future' exercise building on community knowledge from the past and linking this to current student learning. This is in line with the proposal of Shamos (1995) that science and maths literacy should be not so much about detailed content knowledge in science and maths as cultural enterprises (and ways of knowing and doing), deeply related to economic, personal and social life. The curriculum needed to be forward-looking to address new and future challenges and hopes. We were by now working with a different conception of curriculum, as shown in Fig 6.2.

Figure 6.2 Exploring curriculum Model 2

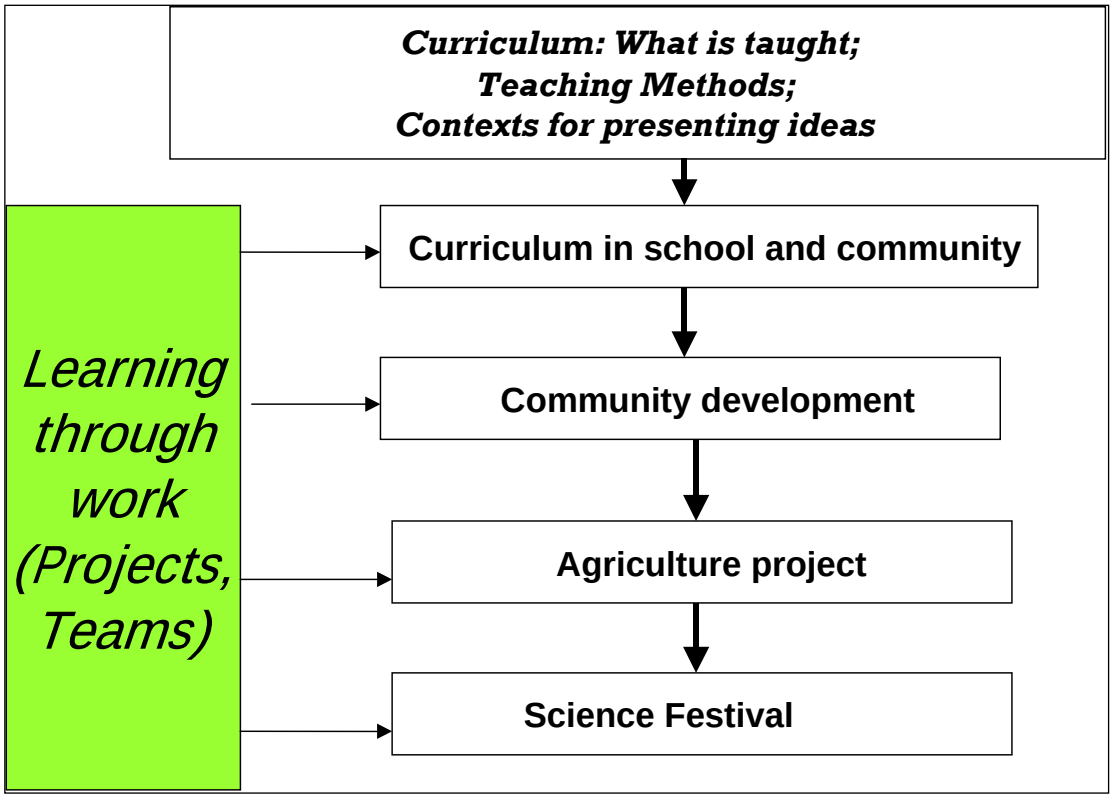


Complexity and interdependence became stronger features, compared to our 'stakeholder' categories. The changes charted in Figures 6.1 & 6.2 came about mostly through community meetings, sometimes directly from suggestions at the meetings, and often from an general uneasiness, only indirectly expressed, that the conceptualisation 'wasn't right' yet. This is how 'crystallisation' was so important in the research: crystallising weaknesses that were dimly seen and then the same with solutions. Looking back at the community meetings it is surprising how little was 'lost in translation' – literally between English and Zulu, and also between a researcher-educator perspective and a parents' perspective, even at relatively superficial levels (and I will talk about identity, community, insiders and outsiders later). The teachers, parents and elders graciously and enthusiastically engaged with vague and open-ended questions. And in retrospect, the basic ideas that shaped the project were there with individuals and groups in the initial meetings.

Through cross-referencing data and discussion, we synthesised the various inputs as we planned for the science festival. Our final understanding of a relevant science curriculum would be tried out in the festival week. "Learning through work" or "learning through projects" proved to be the key. In this, the initial ideas (of Fig 6.1-6.2) remained relevant: each category provided key information and in doing so was

strengthened as a feature in community life. They illustrate how planning, reporting, re-designing and theorising all contributed to data creation.

Figure 6.3 Learning through work is key



Our use of projects and 'learning through work' was a significant curriculum model. We used this for gathering baseline data in school-community assignments through to engaging with Heifer, Woza Moya and Food Gardens, through to the planning and celebrating of the community festival. This figure illustrates the input of project-learning into different facets of the research. There is also a sequential progression from school-based curriculum through community and farmers to a Science Festival.

Data Analysis

> *"It wasn't curiosity that killed the cat.*
> *It was trying to make sense of all the data curiosity generated."*
>
> (Patton, 2002:440)

The 'first pass' at interpreting the data depended heavily on discussions and journal reflections within the research team (including Sihle), personal interpretations of the recorded data and our experiences in the community. This is consistent with methods of grounded theory of constant comparative analysis.

The next strategy of analysis was organising data into categories to develop emerging themes - following the thematic analysis of Brand and Glasson (2004). “Finding patterns is one result of analysis. Finding vagaries, uncertainties, and ambiguities is another.” (Patton, 2002:437). Emerging themes were reflected back to the community in reports and drawings for discussion. Thirdly, themes were ‘jigsawed’ to make sense of the various emerging concepts and theories. In cases where there was convincing contradictory evidence: both or multiple facets of an issue have been discussed.

This type of analysis is dependent on the principles of mindful inquiry, carefulness of observation, not discounting the ‘irrelevant’, ongoing organising, checking understanding with everyone, investigating further and meticulous recording. Time commitments and language translation issues are both enriching and potentially problematic.

Analysis and language

Although I was concerned in my inability to speak Zulu that much meaning would be lost, there was, invariably, someone on hand who would serve as interpreter: teachers, the principal, other researchers or NGO facilitators. In the review of literature I noted many cautions about language in both research and science education. But in some aspects working in a foreign language can have some advantages. One is that people often took great care to explain things that they knew might be strange or not understood. The other concerns the original way an English Second Language speaker may use original – and very insightful expressions – that throw fresh light on a situation. In a new-language speaker it is as if the language is still fresh and there is less of the white noise of tired truisms and clichés. One example is the use of portmanteau words. Sihle in an interview report aptly and charmingly wrote: “I try to do better my job every time I am doing interviews. There are some problems of people being *obstracted* by the tape recorder. But we *ensured* them …(my emphasis).” Lewis Carol could not have described better someone’s combined distraction and obstruction caused by a tape recorder (Sihle’s Journal 7/10/2003. R45c).

Negative data / absent data

The concept of negative data contributes significantly in data analysis. This type of data only emerged in my awareness as time went on. Examples from students’

inputs are: no one asking for better teaching; no one asking for scholarships or help to get out of the community; no one seeking counselling. Photos of 'science in my life' omitted any themes of: energy sources, fire, traditional technology (containers, bricks, storage, building), and none of school. Teachers did not ask for help with lesson plans, they did not ask about results of student assignments. The community never asked for medicines or for (individual) personal help in any way. Similarly, there were *things that I did not see* in the community: for example no examples of art or craft. Patton, (2002:296) rates the ability to assess the "significance of non-occurrences" as particularly helpful in guiding further data collection as well as in providing insights into underlying causes.

Nature of data

Data itself is not only information. It is knowledge; expression of worldview/experience; community life; participation; action; transformation; belief; questions; and feelings.

I present here a brief example of data analysis. In this instance I identify and code key words which I draw out and code into themes.

Analysis example 1

6th September 2002, Lusiba: (See Data LS1, LS2 & LT1)

Grade 7 Drawings of my life: questions; what I've learned

Figure 6.4 Girl's drawing of 'My life'

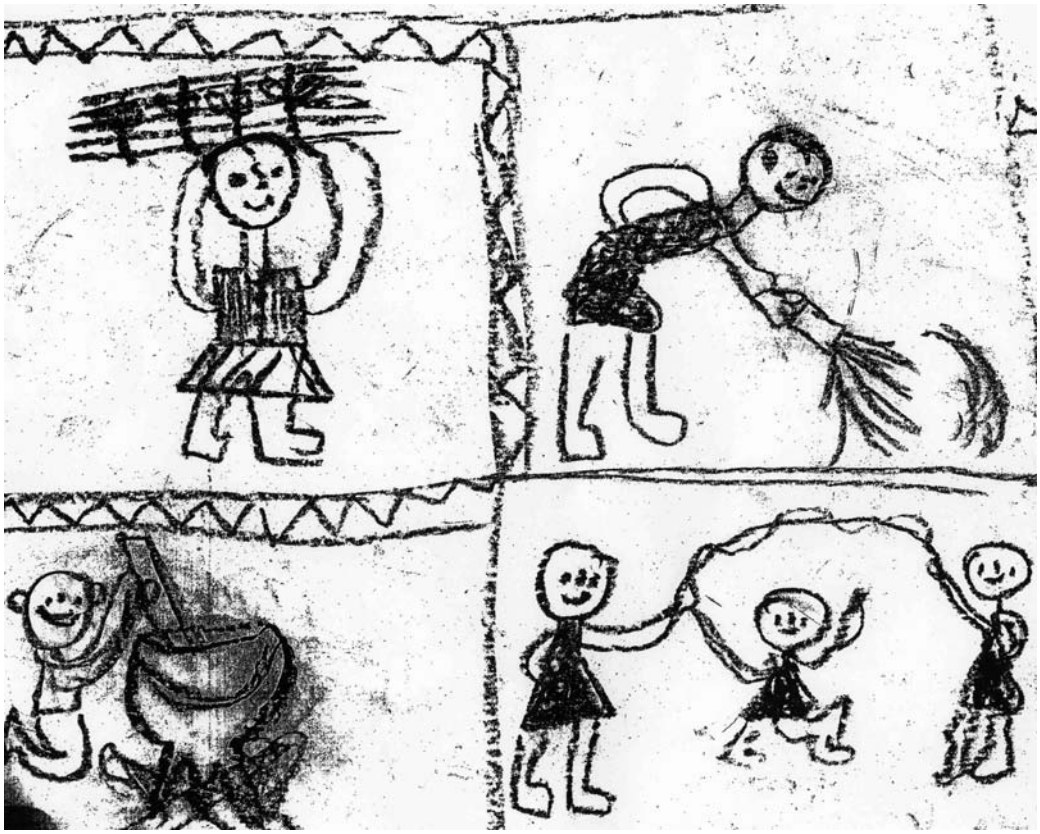
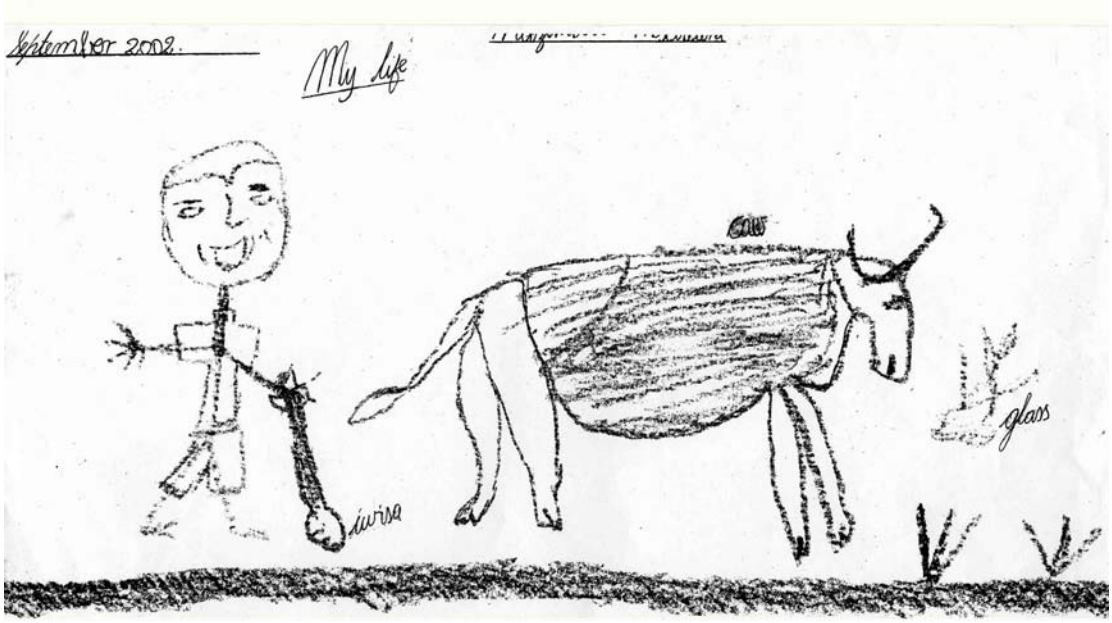


Figure 6.5 Boy's drawing of 'My life'



Student Questions

Is there a God?

Why God makes me sometimes unhappy?

Why are people talking about stripping bark because the **Nyanga** are helping and they are cheaper than doctor?

Why did the creator say people must **die**?

Why do people **kill themselves** badly?

God created heaven and earth but why do black people **die**?

My God why your world is like this?

Why the **teachers** don't work when they are late?

Why have only black people got **AIDS** Why I didn't hear that white people have got it?

(Many questions about **rain**: When I look up I don't see rain?)

Where did sand come from?

Where does **science** start?

Statements about learning

I try to **respect** other people (Many statements about respect)

I learn about **HIV**

My friend told me not to leave **school**

I like our **culture**

Science teaches us about beautiful things

Long time ago people were not happy because there were bad things happening: people were abused by Whites. Blacks didn't like Whites because of the hard life, and stay in fear because it might come to violence [as] in ancient time

Analysis

1. Initially there was translation and sharing of information from the students with the class teacher and other researchers.

2. Tentative categories: (Some of these emerge more clearly in other data sets and some categories overlap.)

Interest in:

Mystery (e.g. 'Where did sand…')

Religion ('Why God')

Natural phenomenon. (**rain**)

Concerns about:

Death (**die**)

Health (**die**)

AIDS (**HIV**)

Tension between

School not working and needing to go to school. (**school/ teachers**)

Political history and maybe the perceptions participants have of researchers (Whites. Blacks)

Traditional and conventional medicine or between school science and traditional practices (**Nyanga**)

Culture is important: including respect, ubuntu… (Drawings; **culture)**

Gender: (clear gender roles and many chores in children's drawings).

Aesthetics: (beautiful)

3. Identifying themes: It was only later after noticing science related so frequently to beauty (in the secondary students' photo assignments) that I noticed the aesthetic appreciation that the students have of science. Having identified the concern about AIDS, it became strange that the teacher group had not mentioned this at all.

4. In feeding back student concerns, teachers seemed more willing to acknowledge them (e.g. AIDS), and then mentioned them in subsequent data.

How does this relate to the research questions? Does it need to?

It was always clear from the start that 'Relevant Science' and 'Research Processes' would be the main foci of 'Findings'. The compelling weight of data relating to the other themes led me to include chapters as outlined below. Themes that finally emerged from the project, three years later, were already evident in this early data:

- Relevant science (with sub-theme on AIDS and Identity)
- Making the invisible visible (What became known)
- Indigenous knowledge (including worldview and religion)
- Community (Insiders and Outsiders)
- School
- Research process / Participation
- Ethics (see question about teachers here)

In weaving these ideas into the themes reported in the findings, I have quoted the data codes from the lists of data in Appendix B.

Conclusions

The data analysis has broadly resulted in the three forms described by Schlechty and Noblit (1982 in Patton, 2002:480): making the obvious obvious; making the obvious dubious; making the hidden obvious. I describe this in some detail in the Theme 'Making the Invisible Visible' in Part 2A.

In keeping with the spectrum of theoretical frameworks described in Chapter 2, I have drawn on a number of analytical structures. These include basic and thick descriptions, conceptual ordering, the emergence of themes, collaborative and participatory analysis - which together imply demystifying research, managing the process and reflexivity. Further, consensual validation was ongoing through collaborative writing, discussions, and a community report of knowledge claims and other outcomes.

Throughout the process, my concept of what constituted data evolved. I started with a focus on 'instruments', interview schedules, questionnaires and carefully designed activities, as well as journal entries. After a while I realised that *everything* is data - everything from the seemingly unrelated physical: Was it raining? Was it not raining? to the first draft posters that students discarded. This realisation broadened too in my understanding of interpretation. For example, when on the first day of the science week there were few students present, my immediate conclusion was: 'This intervention is not working'. Then, through conversations with community I came to realise there are many possibilities: some students had to seek work for the vacation, some were enlisted for chores at home, some travelled to family in Durban, some did not know about the festival… In an *ubuntu* framework, data creation is not simply response to set tasks, it is largely a story of relationship.

It is also now clearer that Model 2 (in Figure 6.2) illustrates the centrality of relationship. Separate entities are not simply abstract concepts or institutions that are linked: school, African knowledge, social structures etc. They exist through relationship. I am a 'person though other persons' applies to these entities as well.