Data Maps: Using Diagrams to Represent and Analyse Interview Data

Carmel M Diezmann
Faculty of Education
Queensland University of Technology
<c.diezmann@qut.edu.au>

Qualitative data collection and analysis have become increasingly important in educational research. This paper discusses the representation and analysis of qualitative data using data maps. The data were collected in interviews with children engaged in mathematical problem solving. Data maps are diagrams that provide wholistic displays of data and are produced using drawing software. There are three advantages in using data maps. First, data maps provide an overview of video, audio or text-based data and can facilitate the identification of critical points or events in an interview. This advantage also enables the selection of representative rather than \textit{ad hoc} excerpts of data in reporting. Second, the wholistic presentation of data maps supports visual reasoning, which is distinct from the sequential reasoning used with transcripts. Thus, a diagrammatic representation of the data provides an additional opportunity to identify patterns and themes in the data. Third, emerging questions can be explored by "cutting and pasting" existing data maps and creating new maps. These latter maps address the interactive assumption of qualitative analysis. This paper argues that a visual approach to data representation can enhance analysis and reporting. Through visual representation and reasoning, data maps provide insights that cannot readily be identified through non-visual analysis. Additionally, data maps provide the "global" context for the rich, thick descriptions that commonly appear in qualitative reports but that represent inherently "local" or subjective views of the data set.
In the past few decades, there has been a surge of qualitative studies that have provided insights into our understanding of teaching and learning (Webb & Glesne, 1992). The value of qualitative research is associated with their ability to present a wholistic but detailed understanding of the phenomena under investigation. Thus, qualitative research can be thought of metaphorically as "an intricate fabric composed of minute threads, many colors, different textures, and various blends of material" (Creswell, 1998, p. 13). Despite the potential of qualitative research, qualitative data analysis is a highly demanding process for the researcher. Creswell describes it as a "formidable task" and one that should be viewed with some trepidation (Creswell, 1998):

(The researcher will) engage in the complex, time-consuming process of data analysis ¾ the ambitious task of sorting though large amounts of data and reducing them to a few themes or categories ... it is a lonely, isolated time of struggling with data. The task is challenging, especially because the database consists of complex texts and images. (pp. 16-17)

Other researchers concur with Creswell's assertion of the difficulty of qualitative data analysis citing it as labour intensive (Bogdan & Biklen, 1982) and time-consuming (Pitman & Maxwell, 1992). The difficulties associated with qualitative data analysis extend beyond practical difficulties to conceptual difficulties. Whilst there are no shared conventions for qualitative data analysis, there is agreement that it should be an interactive and recursive process (Miles & Huberman, 1994) in which data are analyzed "using multiple levels of abstraction" (Creswell, 1998, p. 21).

The difficulties of qualitative data analysis are exacerbated by the limited tools and techniques available to support the analysis process (Miles & Huberman, 1994). However recent technological developments, such as NUD.IST software (1994), offer promise to researchers by facilitating the exploration of patterns and relationships in text. However, in contrast to the well-developed state of quantitative analysis, qualitative analysis is still in a "pioneering" state.

The majority of tools and techniques available to support qualitative analysis, predominantly involve the use of text-based representations of data and sequential reasoning (e.g., LeCompte, Millroy & Preissle, 1992; Miles & Huberman, 1994). Although there is no specific guidance on the visual analysis of qualitative data (e.g., LeCompte, Millroy & Preissle, 1992; Miles & Huberman, 1994), visual display techniques, such as graphs, tables and drawings, are commonly used in quantitative analysis (e.g., Tufte, 1983, 1990). The purpose of this paper is to outline a visual approach to data analysis that involves the use of diagrams that organise and display qualitative data, and show temporal relationships. These special purpose diagrams are henceforth referred to as "data maps" (see Figure 1). These data maps are produced using SmartDraw Pro (1995), an inexpensive and easily managed general-purpose drawing packages. An earlier conception of data maps is described elsewhere (Diezmann, 1996).

Data maps have three specific advantages. First, data maps can provide an overview of video, audio or text-based data and facilitate the identification of critical points or events in an interview. This advantage enables the selection of representative rather than ad hoc excerpts of data in reporting. Furthermore, it enhances the presentation of an audit trail,
and provides sufficient detail about the data and the analysis to enable the reader to evaluate the trustworthiness of the conclusions (Miles & Huberman, 1994).

Second, the wholistic presentation of data maps supports visual reasoning. Visual reasoning capitalises on the wholistic and interconnected representation of information and differs markedly from sequential reasoning (see Barwise & Etchemendy, 1991). Thus, the diagrammatic representation of information provides additional opportunities to identify patterns and themes in the data via visual reasoning.

Third, as the data maps are produced using software (i.e., SmartDraw Pro, 1995), emerging questions can be explored by "cutting and pasting" existing data maps and creating new maps. These latter maps address the interactive assumption of qualitative analysis by facilitating an exploration of "what if" scenarios within the data. This use of software extends thinking "beyond amplification" by engaging cognitive processes in qualitatively different ways (Pea, 1985).

Prior to illustrating how data maps can be used to represent and analyse data, the cognitive advantages of diagrams are discussed. This is followed by a description of data maps and an example of how data maps were used as a tool for analysis in a particular study.

**Cognitive Advantages of Diagrams**

Diagrams are abstract visual representations that exploit spatial layout in a meaningful way, enabling complex processes and structures to be represented wholistically (Winn, 1987). They have particular advantages related to working memory (van Essen & Hamaker, 1990), the conceptualisation of the problem (Yancey, 1981), and the informational content of the representation (Larkin & Simon, 1987). A further advantage of diagrams is that they provide a visual alternative to words (Mayer & Gallini, 1990). Additionally, changing the mode of a representation, for example from a linguistic to a visual form, can be a means of knowledge generation (Karmiloff-Smith, 1990). Due to these cognitive advantages, it is hardly surprising that diagrams are recognised as important sources of insight (e.g., Gleick, 1992), which augers well for their use in qualitative data analysis.

**A Description of Data Maps**

In essence, a data map presents a visual overview of an event in which the key behaviours of the participants are featured in chronological order. For example, a data map could be produced for an interview, a class lesson or a small group interaction. An example data map is shown in Figure 1. This data map depicts an interview with a student prior to an instructional program. In Phase A of the interview, the student worked on a mathematical problem-solving task independently, and in Phase B, there was interaction between the researcher and the student. The key features of the data map are identified to the right of the map. Each of these features is explained in the numbered description on Table 1.
<table>
<thead>
<tr>
<th>PRE-INSTRUCTION</th>
<th>1. Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase A</td>
<td>2A. Subtitle</td>
</tr>
<tr>
<td>Jon</td>
<td>3. Participant</td>
</tr>
<tr>
<td>1-read text (1-2)</td>
<td>4. Reading</td>
</tr>
<tr>
<td>2-non-digram strategy (4-6)</td>
<td>5. Shapes</td>
</tr>
<tr>
<td>3-answer X (7)</td>
<td>6A. Shading</td>
</tr>
<tr>
<td>Phase B</td>
<td>2B. Subtitle</td>
</tr>
<tr>
<td>Jon Researcher</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. An example data map.
Table 1

*Features and Explanations of the Data Maps.*

<table>
<thead>
<tr>
<th>Feature</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Title</td>
<td>This map represents a &quot;Pre-instruction&quot; interview.</td>
</tr>
<tr>
<td>· Sub-title</td>
<td>The subtitle indicates particular parts of the interview. In Phase A, the student worked independently (2A). In Phase B, there was interaction between the student and the researcher (2B).</td>
</tr>
<tr>
<td>· Participant</td>
<td>Jon was the student being interviewed.</td>
</tr>
<tr>
<td>· Reading</td>
<td>The maps are read vertically in the direction of the arrow.</td>
</tr>
<tr>
<td>· Shapes</td>
<td>Distinct shapes signify particular behaviours. In this map, the six-sided figure is used when the student appears to &quot;read text&quot;, whereas a rectangle with curved edges is used when the student presents an answer.</td>
</tr>
<tr>
<td>· Shading</td>
<td>A shaded shape, such as the third shape, indicates that some assessment has occurred. The outcome of the assessment is presented within the shape. The cross indicates an incorrect response to a problem, a tick would have indicated a correct response (6A). Later in this map, there are assessments of the quality of the student's performance on a particular task (6B). In this study, assessments are reported as Levels of performance, with Level 1 being the lowest level of achievement.</td>
</tr>
<tr>
<td>· Text</td>
<td>The text in the shapes represents a brief summary of a behaviour during the interview. For example, in this instance, the researcher asked: &quot;So with <em>The Koala</em> (a problem-solving task) what were you trying to find out?&quot; This query was shortened to &quot;explain goal&quot; in the shape to indicate the essence of the behaviour.</td>
</tr>
<tr>
<td>· Number</td>
<td>The number in the shape refers to the line number of the original transcript for easy reference.</td>
</tr>
<tr>
<td>· Block arrows</td>
<td>Interaction between the researcher and the student is indicated by block arrows. The arrows on the right indicate the researcher's behaviour and the arrows on the left indicate the student's behaviour. The highest positioned arrow is read first.</td>
</tr>
<tr>
<td>· Shapes connected with a cycle</td>
<td>Shapes that are horizontally aligned and connected with a cycle indicate an ongoing interaction between the researcher and the student.</td>
</tr>
</tbody>
</table>
**Data Maps and Qualitative Analysis**

Data maps are useful in qualitative analysis because they support the data analysis components of data reduction, data display, and conclusion drawing and verification (Sowden & Keeves, 1988). In Figure 1, the data were reduced by identifying the key behaviours of the student and researcher. Cross-referencing these behaviours to the text provides a means for the researcher to check to his or her interpretation of the events. If there is a change in the researcher's perception of the behaviour, the data display can be easily updated by modifying the text within the shapes, by the addition or deletion of shapes or by a change in shape. The visual data display supports the employment of a number of data analysis strategies that facilitate inference making, such as: counting, noting patterns and themes, imputing plausibility, noting relations, building a logical chain of evidence and constructing a causal chain (Miles & Huberman, 1994).

**Example Study**

The data maps were specifically developed to assist in the analysis of data in a case study that explored the effect of instruction on students' use of diagrams in novel problem solving (Diezmann, 1999). In that study, eight Year 5 participants completed five isomorphic pairs of novel problem solving tasks in pre- and post-instruction interviews, and, with their classmates experienced a 12 lesson program on diagram generation and diagram use in problem solving. Lessons and interviews were video-taped, and interviews were transcribed for analysis. The focus of the data analysis was on evaluating the effectiveness of the instructional program with reference to a number of criteria, including the generation and use of diagrams. Generally, there was evidence that the instructional program was effective. However the performances of two pairs of students were noteworthy and warranted further exploration.

These pairs of students, who are henceforth referred to as "like pairs", were assigned the same levels of performance on diagram generation and/or diagram use on a pre-instruction task but different levels of performance from each other on the isomorphic post-instruction task. Of course there are many plausible reasons for the differential outcomes of a like pair, such as motivation during instruction. However the assumption of similarity of performance according to an assigned performance level also warrants scrutiny. Thus, a question that emerged from the identification of like pairs was: How similar were the performances of the students within each like pair on the pre-instruction tasks? To address this question, it was necessary to compare the performance of each member of the like pair. Thus, a more fine grained analysis of the interview data was required to explore the circumstances of the assigned levels of performance. Data maps were initially developed for this purpose (see Diezmann, 1996, 1999). As the maps are computer-generated maps, they can easily be combined for comparative purposes or "cut and pasted" to explore particular questions that have arisen during the data analysis process. In the following example, the comparative and investigative uses of data maps are illustrated.

**Using Data Maps for Comparison**

Helen and Jon were one of the like pairs of students who were assigned identical pre-instruction performance scores for diagram generation and diagram use on a particular task but differed in their post-instruction scores on an isomorphic task. These students' performances are now described and data maps are used to compare their performances.

**Diagram Generation**
On the pre-instruction task (see Figure 2), Helen and Jon were each assigned Level 1 for diagram generation because they neither represented the metre marks nor the movement accurately (see Figure 3). There were three levels of expertise ranging from Level 1 to Level 3, which was the highest level. On the post-instruction task (see Figure 2), Helen was assigned Level 2 because she depicted the bricks accurately, although there was an error in the depiction of movement at the base of the well (see Figure 3). Jon, however, was again assigned Level 1 because he failed to show movement and incorrectly drew four bricks per row (see Figure 3).

<table>
<thead>
<tr>
<th>Pre-Instruction Task</th>
<th>Post-Instruction Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>A sleepy koala wants to climb to the top of a gum tree that is 10 metres high. Each day the koala climbs up 10 metres, but each night, while asleep slides back 4 metres. At this rate how many days will it take the koala to reach the top?</td>
<td>A frog was trying to jump out of a well. Each time the frog jumped it went up 4 rows of bricks, but because the bricks were slippery it slipped back one row. How many jumps will the frog need to make if the well is 12 rows high?</td>
</tr>
</tbody>
</table>

*Figure 2. The pre- and post-instruction tasks.*

*Jon: The Koala*  
[Image of koala diagram]

*Jon: The Frog*  
[Image of frog diagram]

*Helen: The Koala*  
[Image of koala diagram]

*Helen: The Frog*  
[Image of frog diagram]
The data map showed that there were three differences between Helen and Jon's performance that may impact on diagram generation (see Figure 4). First, the pre-instruction task differed in its cognitive difficulty for Helen and Jon. On the pre-instruction task, Helen spontaneously generated a diagram, whereas Jon generated a diagram after a specific prompt to *draw a diagram*. Even after the specific prompt, Jon was uncertain about whether or not a diagram could be generated and he had difficulty deciding how to represent the tree in *The Koala* task. His concern was that the height of the tree exceeded the height of his page. Jon's response to scaffolding suggests that this task was difficult for him.

Second, there was a difference between Jon and Helen's understanding of the problem structure. On the pre-instruction task, Jon experienced difficulty with the problem structure prior to his generation of the diagram when he resorted to using key words to determine the operation to use. There was no evidence that Helen had similar difficulty. Understanding the problem structure is fundamental to the generation of an appropriate diagram.

Third, there was a difference between Jon and Helen's understanding of representational systems. Jon's use of numbers with disregard for the problem information suggests that he lacks an understanding of the need for correspondence between the various representational systems. Again, there was no evidence that Helen had similar difficulty. Recognising the need for consistency is implicit in producing accurate translations between representational systems.

Thus, although Helen and Jon were each assigned Level 1 for diagram generation on the pre-instruction task, an analysis of the data maps suggests that prior to instruction Helen is more capable of generating diagrams in problem solving than Jon. Although there was no change in Jon's level of diagram generation on the post-instruction task, the data maps show that on the pre-instruction task his diagram was produced after scaffolding (see Figure 4), whereas on the post-instruction task (see Figure 5) he generated a diagram spontaneously. The change from scaffolded to spontaneous diagram generation indicates a positive change in understanding. In spontaneously generating a diagram, Jon is performing more like Helen in diagram generation, albeit not at the same level.

**Diagram Use**

There were four levels of diagram use ranging from Level 1 to Level 4, which was the highest level. Helen and Jon were each assigned Level 2 for diagram use on the pre-instruction task because there were errors in their use of the diagram. However, their errors were different. The data maps for this task show that Helen and Jon differed in two ways in their diagram use (see Figure 4). First, Jon required considerably more researcher support than Helen prior to using a diagram. Second, Jon was encouraged to quit by the researcher when it was apparent that his use of the diagram was not leading him towards the solution while Helen decided to quit independently of the researcher. Helen's action suggests that she has some metacognitive awareness. Knowing when to curtail solution strategies is an important metacognitive skill.

On the post-instruction task Helen's and Jon's scores for diagram use on the post-instruction task differed. Helen's score for diagram use improved from Level 2 to Level 3. Although Helen's method was appropriate, she made an error in the calculation of the solution. Jon was again only assigned Level 2 because there was an error in his use of the diagram. The data map for the post-instruction task (see Figure 5) showed that both Helen and Jon self-corrected during their use of the diagram. However Helen's self-correction resulted in a correct solution, whereas Jon's did not. As shown on the map, Jon also made other errors, which he did not correct, prior to, and following, the instance of self-correction. Thus, although Jon has some metacognitive awareness, it was not integral to his solution process.
PRE-INSTRUCTION

<table>
<thead>
<tr>
<th>Phase A</th>
<th>Phase A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jon</td>
<td>Helen</td>
</tr>
</tbody>
</table>

- **Jon**: read text (1-3)
  - non-diagram strategy (4-6)
  - 1st answer X (7)

- **Helen**: read text (1)
  - drew a diagram Level 1 (2-3)
  - 1st answer X (4)

<table>
<thead>
<tr>
<th>Phase B</th>
<th>Phase B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jon Researcher</td>
<td>Helen Researcher</td>
</tr>
</tbody>
</table>
Figure 4. Jon’s and Helen’s pre-instruction maps.
<table>
<thead>
<tr>
<th>Phase A</th>
<th>Phase A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jon</td>
<td>Helen</td>
</tr>
</tbody>
</table>

- Jon:
  - read text (167)
  - drew a diagram Level 1 (168-169)
  - 1st answer X (170)

- Helen:
  - read text (31)
  - drew a diagram Level 2 (32-34)
  - 1st answer X (35)

<table>
<thead>
<tr>
<th>Phase B</th>
<th>Phase B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jon Researcher</td>
<td>Helen Researcher</td>
</tr>
</tbody>
</table>
Similarities and Differences that Emerged from the Data Map Comparison

The purpose of using the data maps was to investigate the following question: How similar were the performances of the students within each like pair on the pre-instruction tasks? An analysis of the data maps suggest that there were similarities and differences between Helen and Jon on these particular tasks that were not apparent from their performance scores. On the pre-instruction tasks, when Helen and Jon received identical scores to each other for diagram generation and diagram use, there were differences in: the relative difficulty of the task for these students; their understanding of the problem structure; and their awareness of the need for consistency in translations between representational
systems. Additionally, there was a difference in their ability to quit the problem after incorrect answers were obtained. On the post-instruction tasks, when they were assigned different scores to each other for diagram generation and use some similarities were noted. Both of the students generated a diagram spontaneously and self-corrected at least once during the solution process. Thus, although performance scores may provide a useful indicator of a student's performance relative to a set of benchmarks, data maps provide an insight onto the students' performance relative to each other.

On both diagram generation and diagram use Helen's performance scores improved from the pre-instruction task to the post-instruction task which suggests that she benefited from instruction. However there was no similar improvement in Jon's performance score for either diagram generation or diagram use from the pre-instruction task to the post-instruction task. Which suggests the question: Did Jon benefit from instruction and if so, in what ways? This question can easily be explored by "cutting and pasting" the existing data maps to produce a new data map (see Figure 6) to compare Jon's pre- and post-instruction interview performance.

Using Data Maps for Investigation

**Diagram Generation**

The data maps clearly show substantive differences in Jon's diagram generation on the pre- and post-instruction tasks. On the pre-instruction task, Jon generated a diagram after trying other strategies and after specific scaffolding from the researcher. The map also shows that prior to generating a diagram Jon experienced uncertainty about how to represent a tree, which was an element in the problem. In contrast, on the post-instruction task Jon immediately drew a diagram and exhibited no hesitation about how to represent a wall, which was an element of the problem. Thus, despite Jon's assignment of the same level of performance on both tasks, the map shows that there has been considerable development in Jon's ability to generate a diagram to represent the problem structure.

**Diagram Use**

The map also shows that despite being assigned Level 2 on both tasks, there was some improvement in Jon's diagram use from the pre-instruction task to the post-instruction task. On the post-instruction task Jon engaged in self-correction which was not a feature of his diagram use on the pre-instruction task.

<table>
<thead>
<tr>
<th>PRE-INSTRUCTION</th>
<th>POST-INSTRUCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase A</strong></td>
<td><strong>Phase A</strong></td>
</tr>
<tr>
<td>Jon</td>
<td>Jon</td>
</tr>
</tbody>
</table>
Figure 6. A comparison of Jon's performance on the pre- and post-instruction tasks.
Conclusion

Qualitative studies are by their very nature diverse and accordingly require an analysis that suits the study. Data maps provide a visual approach to representing and analysing qualitative data that complements text-based approaches. The visual analysis of data is dependent on well organised and information rich data maps. The time invested in producing these maps however is time well spent. The maps add structure to volumes of data, facilitate the discernment of patterns and relationships, enhance rigour in analysis, fulfil the interactive assumption of qualitative analysis through "cutting and pasting", and add a layer of abstraction to the data analysis process. Though data maps are still being explored in other contexts, they have considerable potential to extend our thinking in ways that are unavailable when restricted to hours of video- and audio-tapes, volumes of transcripts and field notes or albums of photographs and work samples.
References


