This chapter outlines the design of the research undertaken: the population, the sample design, and the data collection for evaluation. It provides enough detail that any other researcher, in possession of the appropriate skills, would be able to follow the same procedures, and to produce corresponding results.

### 7.1 Selection of cases

Cases for a study can be selected using a variety of approaches, ranging between the mathematically rigorous sample design used in survey research and the single-case purposive selection often used in action research, when one case is selected without considering others. The language of sample design, though positivist in orientation, is useful even with qualitative work, because of its distinction between populations, sampling units, and samples. The use of these concepts hinges around whether any generalization is expected. In the usual type of action research, one social entity is being studied: it is both the sample and the population, and no attempt at generalizability (as opposed to generalization) is being made. If generalizability is to be aimed for (as in this case) the more that sampling units differ from one another, the more care must be taken with sampling.

A further reason for taking special care with the sample design in this case was that analytic induction would be used to form conclusions. The original form of analytic induction, as used by Znaniecki (1934/1968) was criticized by later writers, including Robinson (1951), Hammersley (1981), Lieberson (1991), Bogdan and Biklen (1992), and Katz (2001), as leading to finding conditions that are sufficient but not necessary. This weakness can be countered by comprehensively sampling the population for which generalizations are desired.

When designing a sample to answer a research question, several choices need to be made.

1. The choice of population
2. The choice of sampling unit
3. The choice of sample size
4. The choice of sampling method
5. Selection of cases.
Usually, the sampling unit and the population must be chosen first, and together, because the population is composed of sampling units – which in social inquiry are often individuals, but can also be households, businesses, or any other social or textual entity. When a study involves social groups, there can be several units of study (individuals and groups) but when a group is selected first, that is the sampling unit. The population is the entire collection of units from which the sample can be drawn. The selection is made using a sampling frame, which is either a list of every population member, or a method of choosing members.

Just as the sampling unit and population must be selected simultaneously because of their interactions, so must be the sample size and the sampling method. In a quantitative study, the necessary sample size can be calculated from the minimum tolerable sampling error, based on estimates of the distribution of responses. In such cases, and for many qualitative studies, a sample size can be safely chosen in advance. But when developing a theory or a method (as in the present project), not enough is known to be able to calculate sampling error, so there is no basis on which to select a sample size in advance – apart from economic considerations. An alternative approach is to continually assess the data, and decide on that basis when to stop collecting data. This can be done using a Bayesian approach, but only when fieldwork is very tightly coordinated (e.g. an online survey), or carried out by a single person.

The purpose of a sampling method is generally to ensure representativeness, and thus generalizability. This can be interpreted in two ways: either that all members of the population are represented, or that all relevant concepts derived from that population are represented. For the former, random sampling (or some similar method, such as quota sampling) is used. For the latter, a variant of purposive sampling is normally used. There are many such variants: Patton (2002, pp230-242) lists fifteen of them.

Finally, having chosen a unit, a population, a sample size, and a sampling method, the cases can be selected. The following sections outline the choices made for this study, in the above sequence.

7.1.1 Deciding the sampling unit

Because the scope of the Process is restricted to human futures, and because one of the original criteria for the Process was that it should be participative, this implied that the sampling unit would be a social entity: a group of people with something in common. In terms of the holonic principle (chapter 4, section 4.5) this amounts to a social holon.
7.1.2 Defining the population

Statisticians refer to an “infinite population” as being effectively equivalent to a large population (Kish, 1987), but the population used in this study was truly infinite. As proof: no matter how detailed the definition of a potential set of futures that can be studied, it will always be possible to add more potential futures – though some of these futures will be so similar that differences between them will be (by definition) infinitesimal.

This infinitude of possibilities has a practical implication: that it is not possible to draw a random sample, because there exists no sampling frame against which the representativeness of the sample can be assessed. (This is another reason why a positivist approach was not feasible for developing the Process.)

A tentative solution to the problem of infinite population is that, though the size of the population may be infinite, most of the variance in that population can be captured by using a relatively small number of classes of variables. (To use a statistical parallel, this is almost invariably the case in factor analysis, because of intercorrelations between variables.) Thus it was possible to stratify the sample in terms of variables expected to be relevant to the success of the Process. Given the decision to maximize the variety of the cases, the solution adopted was to develop a sampling frame that is a taxonomy of potential cases.

I propose that most of the variance in any area of human futures can be bounded by four classes of variable: (1) the concern (divided into activity and concept, only one of which normally applies in a study), (2) the social group involved, (3) its location, and (4) the time horizon. This corresponds to the taxonomy of prediction outlined in chapter 3.1 As it is not possible to prove this proposition, I have attempted to demonstrate it empirically, in section 7.1.3 below.

Each of these types of variable can be further subdivided, as shown in the following table, which is both a taxonomy and the basis of a sampling frame for this study. Each subdivision could be subdivided endlessly, but for this purpose there was no value in creating a taxonomy that was orders of magnitude larger than the envisaged sample of cases for this project.

---

1. This is the original sample design, as applied. Later it became evident that “situation” includes two separate components: place and context.
TABLE 7.1 A FACETED TAXONOMY OF HUMAN FUTURES

<table>
<thead>
<tr>
<th>1. A concern (activity or concept)</th>
<th>1.1 An action or habit (e.g. smoking)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.2 An occupation</td>
</tr>
<tr>
<td></td>
<td>1.3 An industry</td>
</tr>
<tr>
<td></td>
<td>1.4 The use of a product or artefact</td>
</tr>
<tr>
<td></td>
<td>1.5 The use of part of the natural world</td>
</tr>
<tr>
<td></td>
<td>1.6 An affect: emotion, value, or belief</td>
</tr>
<tr>
<td></td>
<td>1.7 The use of an abstract concept</td>
</tr>
<tr>
<td>2. A social entity</td>
<td>2.1 An individual</td>
</tr>
<tr>
<td></td>
<td>2.2 A family</td>
</tr>
<tr>
<td></td>
<td>2.3 An informal social group</td>
</tr>
<tr>
<td></td>
<td>2.4 A demographic category</td>
</tr>
<tr>
<td></td>
<td>2.5 A business</td>
</tr>
<tr>
<td></td>
<td>2.6 A non-profit organization</td>
</tr>
<tr>
<td></td>
<td>2.7 A government, or public sector agency</td>
</tr>
<tr>
<td>3. A location</td>
<td>3.1 A locality, suburb, or village</td>
</tr>
<tr>
<td></td>
<td>3.2 A town, city, or rural area</td>
</tr>
<tr>
<td></td>
<td>3.3 A state, province, county, etc.</td>
</tr>
<tr>
<td></td>
<td>3.4 A country</td>
</tr>
<tr>
<td></td>
<td>3.5 A continent or group of countries</td>
</tr>
<tr>
<td></td>
<td>3.6 The world</td>
</tr>
<tr>
<td></td>
<td>3.7 A non-contiguous set of areas with a common property (e.g. “Chinese cities”)</td>
</tr>
<tr>
<td>4. A time horizon</td>
<td>Range from [starting date] to [finishing date]</td>
</tr>
</tbody>
</table>

Using the above taxonomy, any futures project can be described as studying “the futures of [specified activity or concept] among [specified social group] in [specified location] during [specified time range].” Where one of the first three elements is unstated, the assumption is that there is no limit unless this is implicit in the other criteria. Hence the default concern is all activities or concepts, the default social group is the entire population, and the default situation is the whole world. The time horizon is different: its default assumption is not “the whole of recorded history,” but from the present to a date in the future that is indefinite, but probably no more than several decades.

For example, choosing a difficult and unlikely subject such as “undergrowth”, the focus of such a study might be bounded to “the future of the use of undergrowth [subject], on farms [activity = industry type], in England [geography = country], up to 2025 [chronology].” The social group was omitted from that statement, but it would obviously include people involved

---

2. The above taxonomy turned out to be very similar to the CIDOC-CRM ontology used in classifying cultural heritage for museums (Doerr, Hunter, and Lagoze, 2002).
with farming in England. (This discussion has assumed that the subject is already chosen. In practice, as explained in the next chapter, the precise choice of subject becomes part of the inquiry itself.) After completing the above taxonomy, I realized that it did not apply only to futures work. By varying the time horizon, it could also apply to any social research program (setting the time horizon to the present) and any historical research (setting the time horizon to a range of years in the past). Splitting the situation into the geographical area and the context of the inquiry would further sharpen it.

7.1.3 Empirical testing of the sampling frame

To refine the above sampling frame, I attempted to locate a logically comprehensive list of possible topics, using fault-tree analysis, but found that it was not possible to divide the entire conceptual world in this binary fashion. Had that been possible, it would probably have been done already, and used widely – but the existence of classificatory ambiguity in well-established classification systems such as the Dewey system and the Library of Congress system, and the gradual acceptance of faceted classification (Ranganathan, 1967) suggests that it is perhaps not possible. I therefore developed the above taxonomy using empirical data, from two distinct sources: a dictionary, and previous research.

The first source was a systematic sample of 100 headwords (on pages 9, 18, 27... 900) from the Oxford Illustrated Dictionary (Coulson et al, 1962). This particular dictionary was selected because it covers a wider range of word types than most dictionaries, including organizations, place names, and other concepts that normally appear only in encyclopedias. On each selected page I located the first headword (“X”) for which it would make sense to study the “future of X” (or of “the future of the human use of X”). For example, the first three entries were Adriatic Sea, alcohol, angelica; the last was undergrowth. In general, only nouns and verbs (as participles) could have their futures studied. For example, it makes no sense to study the future of ad interim, the first entry on page 9.

All 100 of the selected words fitted into at least one category in the above taxonomy. This provides some assurance that (practically, if not theoretically) the taxonomy is comprehensive. Two kinds of potential error exist in any classification: overlapping and omission – corresponding to the statistical Type I and Type II errors. As noted by Sneath and Sokal (1973), to minimize either type of error must fail to minimize the other. In this case, because it was more important to include false positives than to omit false negatives, the taxonomy is deliberately very broad.
The second source was 129 papers published between 1984 and 2004 in *Futures*, the longest-established journal of futures studies. These comprised all the papers whose abstracts included the phrase “future of X,” where X was anything specific. Because the purpose of this exercise was to unearth problems, no paper was excluded because it was “difficult” – unless its precise topic was unclear from the abstract. This proved more challenging than the dictionary-based classification, because of some complex multi-word concepts that did not fit the original classification of activity types. A particular problem was the proliferation of very abstract types of future, including the futures of transdisciplinarity, the self, political resistance, negotiating skills, global corporate responsibility, and Church-state relations. 53 of the 129 papers (41%) fell under this heading, and could not be accommodated by the dictionary-based taxonomy. The original facet of Activities was therefore replaced by the broader facet of Concerns, including both activities and concepts.

### 7.1.4 Deciding the sampling method

When the number of cases in a study is small (less than about 30, though there is no fixed number: it depends on the between-case variance) the Law of Large Numbers does not apply. One consequence is that random sampling cannot be relied on to produce a representative sample – even when the population is enumerable (Kish, 1987). Because of this inherent statistical problem, different methods of sampling have been developed for small-\(n\) case studies, generally involving purposive approaches to sampling. There are at least fifteen varieties of purposive sampling (Patton, 2002, pp230-242), of which the most relevant for this study were maximum variation sampling and theoretical sampling.

Theoretical sampling is used in grounded theory (Glaser and Strauss, 1967; Coyne, 1997), with each successive case chosen so as to try to disconfirm the findings from the previous case. In principle, this method would have been ideal for developing this Process. However, to use theoretical sampling assumes that enough is known about each member of the population that a likely disconfirming case can be identified in advance. While this may be true when the more public characteristics of a case are being studied (as is often the situation in grounded theory), such preliminary knowledge is not generally available when a method is being developed. The variables affecting successful use of the method with one case often can not be known in advance, and may be independent of externally identifiable characteristics of the case.

A further problem with theoretical sampling (in this instance) was the sheer amount of time it took to identify a suitable entity to study, then persuade it to co-operate, and finally arrange a time when its senior managers could meet together. The shortest time I achieved between
initial contact and beginning the fieldwork was four months. The longest was two years—and that did not include any time spent in identifying a suitable entity, before approaching it. The use of theoretical sampling would thus have extended the fieldwork inordinately.

Therefore, instead of theoretical sampling, I used a form of maximum variation sampling. The key difference between theoretical sampling and maximum-variation sampling is that theoretical sampling requires knowledge of the likely nature of each case before that case is selected: what characteristics that case may possess that are likely to disconfirm findings from a previous case. Maximum variation sampling is not sequential; it requires much less initial knowledge, and a form of stratification can be used to identify suitable cases that were initially unknown to the researchers.

A literature search using the Science Direct and Ebsco online databases revealed only 37 references to “maximum variation sampling” (including its less common synonyms, “maximum diversity” and “maximum heterogeneity” sampling), and most of those references simply reported using that method, without explaining it. The fullest methodological guide to this topic (apart from a brief paper of my own, currently under review) appears to be a single page in Michael Quinn Patton’s *Qualitative Research and Evaluation Methods* (Patton, 2002, p234).

The central principle of maximum variation sampling is to search for commonality among extremities. Cases are chosen, within a population that may not be enumerable, so that they are as different as possible from one another across any variables that are initially known. In the analysis, the researcher seeks principles that apply across all the cases; this sampling method is therefore well paired with analytic induction. The assumption is that, because these cases represent such a diverse group, any properties that they all have in common are also likely to be shared by other members of that population.

### 7.1.5 Deciding the sample size

Two factors impinge on the desired sample size: the information gained from each successive case, and the cost of adding another case. For the fullest development of the action research cycle, the number of iterations should be large. One would expect a law of diminishing returns to apply: that as each new case was added to the sample, fewer and fewer amendments would be needed for the process. A parallel was found in Landauer and Nielsen’s (1993) studies of software usability, as demonstrated by the formula

\[
P = N \left( 1 - \left(1 - L\right)^n \right)
\]

where
\[ P = \text{the number of problems found}, \]
\[ N = \text{number of known problems} \]
\[ L = \text{the probability of any given sample member finding any given problem} \]
\[ n = \text{sample size} \]

Graphing \( P/N \) against \( n \) produces a curve that tails off rapidly:

As the number of cases increases, the cumulative number of problems found grows more and more slowly. \( L \) of course is initially unknown; the above graph uses \( L=0.25 \), which seemed plausible both beforehand and in retrospect. (That is, the estimate was that each case study would uncover 25% of the problems – assuming of course that the number of problems would remain constant.) With \( L=0.25 \), six cases would uncover 82.2% of the problems.

Though Nielsen’s graph was designed for software usability studies, the same principle applies to the present situation: developing a method. At some point, the expected gain from adding a new case will be less than the effort of researching that case (which is linear), and the sampling should therefore stop at that point. This aligns with Lincoln and Guba’s recommendations on sample selection in naturalistic inquiry:

In purposeful sampling the size of the sample is determined by informational considerations. If the purpose is to maximize information, the sampling is terminated when no new information is
forthcoming from new sampled units; thus *redundancy* is the primary criterion. (Lincoln and Guba, 1985, p202)

By using sequential sampling (finishing one case before beginning the next) it is possible to detect the tailing-off of new information, and decide at which point to stop adding new cases.

Another factor impinging on sample size selection was the practical one of resources and time limits. Given the timetable laid down for a PhD thesis, and the amount of time involved in locating a suitable case, then working with the people, planning the workshops, developing the method, and reflection on the outcomes, it was clearly not going to be possible to work in detail on many case studies. (In fact, it turned out that my initial plan greatly underestimated the amount of time it would take to work with each case, mainly because of the many preparatory and follow-up meetings involved.)

After considering these factors, I decided to aim for a provisional sample of 6 cases. By choosing cases that were as widely varied as possible, I hoped to maximize the possibility of discovering problems with the nascent process. As Eisenhardt (1989) suggests, when developing a theory, the number of possible comparisons between sample members can be a more relevant measure than sample size as such – and the number of potential comparisons rises rapidly with sample size. With a sample of $n$, the number of possible paired comparisons is $n(n-1)/2$, which is 3 comparisons for a sample of 3 cases, 15 for a sample of 6, and 66 for a sample of 12.

### 7.1.6 Selection of the sample

The sampling taxonomy in Table 7.1 above is in terms of four facets: subject, social group, geographical area, and time horizon. In practice, any case chosen would include elements of all four: it would be a particular concern, for a particular social entity, in some situation, in some future period. However, the taxonomy was not detailed enough to enable selecting a sample that might uncover all suspected problems. Some variables related to the likely success of the Process were not specifically covered in the above taxonomy: particularly the size of the entity whose future was being studied. Thus the four facets of futures in the taxonomy were supplemented for the purposes of developing the method.

Some potential cases and classes of case were excluded, for various reasons:

1. Because of likely ethical problems, individuals and families were excluded. A further reason for excluding this category was that the workshop process being developed was not designed to apply to groups with very small memberships.
2. Because of linguistic problems, I chose only English-speaking cases (apart from the pilot case, which was part of a larger consulting project, using an Indonesian interpreter).

3. Owing to logistical and financial limitations, most of my cases were confined to South Australia, where I live.

4. Because this thesis was to focus on human futures, the physical future of geographic areas was excluded from the scope of study. For the same reason, I excluded the futures of animal species, natural habitats, geological futures, astronomical occurrences, and the like (though Soper, 1995, after Foucault, argues strongly that the natural environment can be treated as a human construct).

In terms of Table 7.1 (the taxonomy of possible futures) the goal was to include a range across each of the four types of variable: concerns, social entities, situations, and time ranges. Since many sampling variables can apply to a particular case, and there were more variables than cases, each case had to represent a number of variables. In addition to the above systematic grouping of social entities, some specific characteristics of social entities were particularly sought for inclusion in the sample, because they might reveal particular problems with the Process. Given a planned sample size of only 6, all of these would need to be additional characteristics of cases sampled to fulfil other criteria. I recognized from that outset that it would probably not be possible to include all of these criteria in the sample. The following table shows the number of sampled entities of each type; the desirable figure was at least 1 in each case.
A further consideration is that of Stake, who notes that, in case study research, “My choice would be to take that case from which we feel we can learn the most. That may mean taking the one that we can spend the most time with. Potential for learning is a different and sometimes superior criterion to representativeness” (Stake, 1995, p243). In the case of formative research (such as this) Stake’s criterion would appear to apply strongly. Though until near the end I had little choice of cases, I bore Stake’s consideration in mind when negotiating the terms and scope of each case study.

The following table lists the final sample selection case by case, showing which of the desired sampling characteristics were met by each case.
TABLE 7.3  CHARACTERISTICS OF SAMPLED CASES

<table>
<thead>
<tr>
<th>Characteristics of the entity</th>
<th>Case ...</th>
<th>RN</th>
<th>EM</th>
<th>Iraq</th>
<th>LS</th>
<th>CU</th>
<th>SC</th>
<th>Barossa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale of entity</td>
<td>medium</td>
<td>small</td>
<td>huge</td>
<td>small</td>
<td>medium</td>
<td>large</td>
<td>large</td>
<td></td>
</tr>
<tr>
<td>Facing major change?</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous futures work?</td>
<td>yes²</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strongly constrained by its enveloping entity?</td>
<td>yes</td>
<td>not applicable</td>
<td>not applicable</td>
<td>not applicable</td>
<td>to some extent</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Many factors unknown?</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High level of innovation?</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High level of conflict?</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content domain</td>
<td>Information</td>
<td>Technical</td>
<td>Governance</td>
<td>Governance</td>
<td>Financial</td>
<td>Social</td>
<td>Natural</td>
<td></td>
</tr>
<tr>
<td>- selected from FARTHINGS⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector of entity</td>
<td>public</td>
<td>private</td>
<td>..</td>
<td>NGO</td>
<td>private¹</td>
<td>NGO</td>
<td>..</td>
<td></td>
</tr>
<tr>
<td>Situation</td>
<td>..</td>
<td>..</td>
<td>country</td>
<td>..</td>
<td>..</td>
<td>state</td>
<td>local</td>
<td></td>
</tr>
</tbody>
</table>

Characteristics of the study

<table>
<thead>
<tr>
<th>Concern (activity or concept)</th>
<th>2003 war</th>
<th>governance</th>
<th>South Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>- if not the entity as a whole</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Riau</td>
<td>10 State</td>
<td></td>
</tr>
<tr>
<td>- if not entire area of entity</td>
<td>province</td>
<td>10 south</td>
<td></td>
</tr>
<tr>
<td>Time horizon (years)</td>
<td>~5</td>
<td>10</td>
<td>~5</td>
</tr>
<tr>
<td>Development method</td>
<td>workshop</td>
<td>interview</td>
<td>secondary shops</td>
</tr>
<tr>
<td>Language other than English?</td>
<td>yes</td>
<td>no</td>
<td>partly</td>
</tr>
<tr>
<td>Follow-up possible?</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

1. Officially not-for-profit, but behaves as if in the private sector (see Appendix 4).
2. Reported as existing (scenario planning by the US military) but unpublished.
3. No activity-based futures were used in this project.
4. See chapter 5, section 5.4.6
5. Followed up, but with a shorter than desirable delay after completion of the main study.

7.2  Fieldwork method

This section describes the fieldwork method used, covering issues of sequence of timing, data collection by cognitive mapping, the systematic use of iterative cycles, creating a suitable development environment, the fieldwork schedule within each case, the use of reflection in developing the Process, follow-up evaluation, and ethical issues in group work.

7.2.1  Deciding the sequence and timing of cases

In order to derive the maximum benefit from the action research cycle, it was desirable to finish the initial reflection on one case before beginning fieldwork with the next case. This was so that reflections arising from the previous case, and any deficiencies found with the Process,
could be addressed with the next case. As scenario networks are designed to be used over a period of at least several years, it would have been ideal to return to each case after that time, to review the outcome of the process with that case. However the exigencies of PhD completion meant that this cycle had to be accelerated for the later cases.

The original plan was to begin with cases that seemed relatively straightforward (e.g. private sector, fairly small, with few impinging stakeholders and little intrinsic conflict) and gradually work up to the other extreme. The most difficult expected was a large international voluntary organization, with many stakeholders and high conflict. Due to difficulties in finding social entities to co-operate in the process (as well as lack of initial knowledge about entities that agreed to co-operate) it was not possible to follow this sequence exactly, but deviations were not major. I had originally hoped to cover six cases in a year, but that was before I realized how much time would be spent on preliminary dealings with each entity. In the end, it took more than two years to cover the six cases (not including the pilot study).

7.2.2 Data collection by cognitive mapping

I used a method of data collection described by Laukkanen (1994). This is a method for producing a variant of cognitive maps, referred to by Laukkanen as “cause maps.” It is based on the “self-Q” technique of Bougon (1983) and elaborated by Bougon and others in Huff (1990). Eden and Ackermann (1998) describe a similar method. The latter method involves interviewing managers, often using two sessions. In the first interview the interviewer becomes familiar with the terminology used by the manager, while the second is devoted to creating a cognitive map of that manager’s understanding of the mechanisms of his or her work environment. Specifically, this involves a semi-structured interview, creating a set of statements of the type “A leads to B.” When all the As and Bs are connected, and the specific meaning of “leads to” (for that pair of concepts) is elicited, a cause map can be drawn.

Cognitive or causal mapping, as described by most writers on the topic, is an individual process, in which one researcher interviews one manager. Because this Process uses group workshops (to gain a wider range of perspectives) the method of elicitation had to be modified. Eden and Ackermann (1998) describe a method of strategy development, around which their proprietary software, Decision Explorer, is designed. This involves small-group workshops using walls to display data, in a similar way to this Process.

A practical difference in procedure from the Bougon/Laukkanen method in this project was that the participants wrote each concept on a small sticky note, then placed these notes on a
whiteboard or large sheet of paper, drawing arrows between the concepts, and moving the notes and arrows until those involved were satisfied with the set of relationships expressed. One advantage of the wall-based mapping process is that it gives participants a joint task to focus on. As noted by Doyle and Straus (1993), this tends to induce a spirit of co-operation among participants, temporarily reducing the conflict that often surfaces in a workshop situation.

### 7.2.3 The use of iterative cycles

In Lewin’s original (1946) paper on action research, the action research cycle involved three stages: planning, action, and reflection. Some recent writers distinguish four stages, with the addition of Observation between action and reflection (Carr and Kemmis, 1986), while Susman and Evered (1978) describe a five-stage cycle. Though I found no explicit argument in favour of a four-stage cycle over the three-stage version, it seems logical that observation should precede reflection – though the findings of Bem (1972 and 2002) and the review by Ajzen (2001) imply the opposite. If observation is not simultaneous with reflection, it is probably simultaneous with action, because it is the action that is being observed. From a constructionist viewpoint, it can be argued, paradoxically, that some reflection must precede observation: people must know what to expect before they can see anything (as empirically demonstrated by Piaget, 1971). In that case, some element of reflection could be present even as the first cycle is being planned, comparing expectations with a mental model derived from past experience. But regardless of the number of stages distinguished in the cycle, chronologically there are three: (1) planning, (2) acting and simultaneous observation, (3) formal reflection on completion of the action. It can be argued that all three stages involve reflection: on what might happen, on what is happening, and on what did happen.

The iterative cycle, or spiral (implying a third dimension, of improvement), is one of the essential characteristics of action research, but few writers have focused on the process of cycling, and the transfer of knowledge between cycles. For example, several of the most widely referenced books on action research (Reason and Bradbury, 2001a; Greenwood and Levin, 1998; Argyris, Putnam, and Smith, 1985; Hart and Bond, 1995) devote no more than a page to this aspect. It is addressed by Dick (2000) and by Carr and Kemmis (1986), but not in detail. Kock (1997, 2004) provides detailed treatments of the action research cycle, using Susman and Evered’s (1978) five-stage cycle of diagnosing, planning, action, evaluating, and
specifying learning. However, Kock’s positivist epistemology restricted the learning to a single-loop model, in Argyris’ terms.

7.2.3.1 Iterative cycles in fields other than action research

It is not only action research that uses the principle of improvement through iteration:

- Most educational practice also involves repetition: understanding a phenomenon by repeating it in different ways. This practice is well supported by theory. Kolb’s (1984) cycle of experiential learning, derived from the theories of Jung (1921/1971), has strong similarities to the action research cycle.

- Sancar and Onaran (2001) point out that the legal system, in its alternation between case judgements and legislation, uses a similar cycle to action research.

- Likewise, the quality management movement has introduced the concept of continuous improvement, particularly in manufacturing. Deming’s (1994) P-D-S-A cycle (plan-do-study-act) is strikingly similar to the action research cycle. The kaizen method first used in Japanese manufacturing is a less formal application of cycles (Mizuno, 1988). Hatten et al (2000) compare quality management with action research, finding many similarities.

- Charrette (a participatory urban planning method) uses a 24-hour design cycle, as noted by Lennertz (1999, p2):

  Four days is required to accommodate three feedback loops, scheduled at least a day apart. Three loops are the minimum required to facilitate a change in participants’ perceptions and positions. Only simple projects with little controversy should be attempted in four days. More complicated projects typically take seven days.

  This is a different type of charrette from that of Schuttler, as described by Glenn (1994).

- However, futures methods (apart from Delphi) do not make explicit use of cycles of learning.

From the extent of the above examples, it may be that most human learning occurs through such a cyclic process, as Kolb (1984) suggests. The potential contribution of action research is to make that process explicit, enabling a focus on improvement through systematic reflection (as applied by Langley et al, 1996).

7.2.3.2 Multiple loops

The theoretical weakness of the standard single-loop cycle (whether or not repeated) of action research, is, as Argyris and Schön (1978) point out, that it has no place for the original impetus of a project to be questioned – hence their concept of double-loop learning. An example of this is the development of the consensus group technique (as noted above) in which the feasibility of each successive approximation to a useful method was reviewed in the context of
unclear and changing demand from potential clients. In the language of causation, a double-loop approach is the equivalent of redefining causes as well as effects. In action research, this corresponds to a concept of cycles within cycles (chronologically), as noted by Dick (2000).

For the development of this Process, a double-loop method was used. The inner loop corresponds to the development of the scenario mapping methodology for a particular case. The participants in that case are involved only with this inner loop. In practical terms (as outlined below) the inner loop involves cycles within cycles: the outer cycle is the entire half-day workshop, while the inner cycle corresponds to each of the several tasks scheduled within that workshop.

The following diagram illustrates the linkage of the inner and outer cycles.

![The concept of cycles within cycles within cycles](image)

**TABLE 7.4  UNITS OF REPETITION WITHIN CYCLES**

<table>
<thead>
<tr>
<th>Cycle type</th>
<th>Unit of repetition</th>
<th>Observer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer cycle</td>
<td>Each case</td>
<td>Developer of the method</td>
</tr>
<tr>
<td>Inner cycle, Level 1</td>
<td>Each workshop in each case</td>
<td>Developer and participants</td>
</tr>
<tr>
<td>Inner cycle, Level 2</td>
<td>Each task in each workshop</td>
<td>Developer and participants</td>
</tr>
<tr>
<td>Inner cycle, Level 3</td>
<td>Each component of each task</td>
<td>Developer and participants</td>
</tr>
</tbody>
</table>

An alternative notation for the development of a Process is a triangular cycle, which involves passing information between three elements: the conceptual framework underlying the Process, the Process being developed, and participants’ reflections:

(i) Experience with developing the participative Process generates ideas that can improve that Process. These ideas can be used in a later workshop of the current case study, or the next case study.
(ii) The conceptual framework guides the development of the Process, which in turn refines the conceptual framework.

(iii) The conceptual framework generates reflections and ideas, which can be used to modify the framework.

This three-way process could be illustrated as follows. The figure can be read as six simple (subject-verb-object) sentences, with two beginning from each oval, one in each direction (with verbs as italicized above):

**Figure 7.3  THE TRIANGULAR CYCLE OF METHODOLOGICAL DEVELOPMENT**

Figure 7.3 is similar to Ulrich’s (2000, p252) “eternal triangle” of facts (analogous to the development process), values (ideas for the process), and system (conceptual framework). Though Ulrich’s description does not mention cycling between the three, his Critical System Heuristics (Ulrich, 1994 and 1996) is effectively a form of action research, and in a personal communication Ulrich confirmed that “cycling between the three corners of my ‘eternal triangle’ is an essential part of its application.”

Checkland’s FMA model (Checkland, 1999; Hindle et al, 1995) is also similar. In FMA, F is the conceptual framework used, M is the methodology, used (corresponding to the Process), and A refers to “area of application” – equivalent to the case in Figure 7.2. Reflection is not explicitly included in the FMA model, but M is permitted to change in response to F and A.

For the development of a method, there is one such triangle for each case: the triangle can be envisaged as surrounding the case, which would be labelled at the centre of the triangle, with all the information being passed around the case. The successive case-based triangles can be

---

envisaged as superimposed on one another, with gradually changing reflections, and gradual improvement of the process and framework (as noted following each case study in chapter 8). The following figure is thus a more complete representation of the development of the Process.

![Diagram of Development Through Repeated Triangular Cycles](image)

**FIGURE 7.4 DEVELOPMENT THROUGH REPEATED TRIANGULAR CYCLES**

Though the above triangular model was developed for this project, it is not specific to it, and could be applied to the development of any method of social inquiry.

### 7.2.3.3 Use of cycles in participative workshops

The Process took place through a series of workshops – generally, four half-day workshops for each entity. These share the concept of participation with Jungk’s “future workshops” (Jungk and Müllert, 1987), the “futures-oriented workshops” of Wood and Christakis (1984), and the large-group scenarios of Steil and Gibbons-Carr (2005) but differ in other ways.

The four-workshop format used the cyclical principle by having each workshop focus on a particular aspect of the entity that would determine its future, generally in this sequence:

- Workshop 1: What forces from the past will constrain your future?
- Workshop 2: What forces from your stakeholder groups will constrain your future?
- Workshop 3: What might you expect to encounter on the paths to your future?
- Workshop 4: How might all these forces affect one another?

Each half-day workshop consisted of around six short tasks (the Level 2 inner cycle listed in Table 7.4 above). These tasks varied for each workshop; a typical list of them is provided in Appendix 2.
The Level 3 inner cycle took place within each task, following this general pattern:

1. Brief plenary session: the facilitator explains guidelines for small groups in this cycle.
2. Small groups (3 or 4 participants) work on the task, often by writing concepts on small sticky notes and arranging them as an event tree.
3. After each small-group session, a spokesperson for each group reports back to the plenary session, and adds the group’s findings to the steadily growing diagram on the wall. Because each group has worked at the same task, many similarities are noted. Reasons for any differences are discussed, often resulting in changed perceptions.

Each sequence of the above three tasks can be seen as a small action research cycle in itself, corresponding to Lewin’s three-stage cycle: planning, action, and “fact-finding about the result of the action.” The initial plenary session involves planning, the small-group work is the action, and the plenary reporting session is the fact-finding or reflection. It might be argued that this is not a true action research cycle, because the day’s work sequence is pre-planned, and there is thus no chance for the concluding plenary to affect the next task. However, because I was keen to cultivate serendipitous improvements, and I was the facilitator (for most cases), I made an in-principle decision to facilitate quite loosely. This involved not insisting on keeping exactly to the timings and sequence laid down in each workshop’s agenda, and being prepared to improvise when necessary, using a repertoire of what Andersen, Richardson, and Vennix (1997) refer to as scripts. Thus in the circumstances, it was true action research, though not at a high level of participation, because the decision to vary the following task was not the participants’ but mine – though always in response to participants’ activities. Participants were helpful, but (with several exceptions) not well-enough informed about the previous development of the Process or about futures methodology to make decisions during the flow of the workshop.

7.2.4 Creating a suitable environment for development

The partial-prediction axiom in chapter 4 states that the medium-term future of a social entity (despite many participants’ preconceptions) is to some extent knowable, at least in broad terms. In graphical terms, the future is an area or envelope, rather than a point. The primary purpose of the Process is to help participants develop their capabilities for foresight, which can be envisaged as a weak force, easily overwhelmed by the pressures of daily life. It follows that the environment in which the foresight takes place is crucial: psychological “noise” must be minimized to allow the weak force to be perceptible. All aspects of the development environment should be conducive to foreseeing the envelope of futures that may occur.
The iteration involved in the inner loops is therefore an attempt to help participants view that envelope from different angles. This is similar to Morgan’s (1997) images of organizations, and to Linstone’s Multiple Perspectives approach. Linstone (1984) suggests applying three types of perspective that can be applied to any management problem: technical, organizational, and personal. Integrating data seen through each of the perspectives can reveal insight that is more than the sum of the parts, and help unrealized assumptions to surface. Likewise, the principle of this Process is that, since the future is difficult to perceive through the fog of time, the cyclical paradigm of action research can be used to enable multiple perspectives of the future.

7.2.5 Fieldwork schedule for each case

This section lists a generic schedule for the four main workshops in the Process (as described in principle in Chapter 5, with a detailed agenda in Appendix 2), preceded by some preliminary planning meetings, and followed (about a year later) by a follow-up meeting. Though no case followed this schedule exactly, departures from it were not significant.

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Time since previous</th>
<th>Purpose</th>
<th>Output</th>
<th>Evaluation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary meeting/s to plan workshops</td>
<td></td>
<td>Plan for workshops</td>
<td>Critical reflection, comparing cases</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2 weeks</td>
<td>Introduction; consider the presents</td>
<td>Map stakeholders and influences</td>
<td>Questionnaire and discussion at end</td>
</tr>
<tr>
<td>2</td>
<td>1 week</td>
<td>Consider the pasts</td>
<td>Left-hand side of scenario map</td>
<td>Discussion at end</td>
</tr>
<tr>
<td>3</td>
<td>1 week</td>
<td>Consider possible futures</td>
<td>Right-hand side of scenario map</td>
<td>Discussion at end</td>
</tr>
<tr>
<td>4</td>
<td>1 week</td>
<td>Combine previous weeks’ data, add attributional links</td>
<td>Annotated scenario map with layers</td>
<td>Questionnaire and discussion at end</td>
</tr>
<tr>
<td>5</td>
<td>1 year</td>
<td>Review of scenario map and flow-on effects</td>
<td>Revised scenario map</td>
<td>Entire meeting is evaluation</td>
</tr>
</tbody>
</table>

Preliminary meeting(s) for data collection and planning

This preliminary stage involved coming to grips with the nature of each new case. In practical terms, this was done by holding meetings with managers of that entity, as well as preliminary phone calls and email messages. I had not initially envisaged the volume of this initial contact, which for some cases occupied more time than the scenario workshops themselves. In terms of the action research cycle, this can be viewed as an anomaly: the cycle begins with planning,
but this stage preceded planning – it was more observation. This involved exploring the possible process with insiders, learning about the organization, and what outcomes might be expected: characterized by Wadsworth (2001) as “compass work.”

The information collected at this stage included:

1. Stakeholder groups – and motivations of each.
2. Impinging systems – and forces exerted by each.
3. Key actors (decision makers, and possible future ones) and their interests.
4. Broad social trends which may impact indirectly through many sources.
5. The goals, values, baggage, worldviews, and other factors that may influence the organization’s actions (first pass).

After this initial information had been collected, the planning could begin. This involved deciding on the number and nature of participants, the particular participants, the venue, the dates, and the content of the meetings – as tailored to that specific entity. Creating a suitable environment is particularly important, both in terms of factors relating to shape, size, and arrangement of the workspace (Doyle and Straus, 1993) and the effects of the presence of other people (Green and Hart, 1999).

7.2.6 The use of reflection in developing the Process

An essential component of action research is the use of reflection. As noted by Dick, without systematic reflection, an intervention cannot accurately be described as action research (Williams, 2004). In the present study, with its twin substantive and developmental components, two forms of reflection were used:

(a) Reflections on the futures of the social entity currently being studied.
(b) Reflections on the development of the Process.

The former type of reflection was done mainly by participants (stakeholders in the social entity) while the latter type was done mainly by the researcher, as well as the assistants used in some of the case studies. To ensure that reflection occurred and could be productive, the following steps were taken to encourage reflection by participants:

1. Fifteen minutes was allocated at the end of each workshop, specifically for evaluation. Both written and spoken data were collected.
A questionnaire (incorporating the “critical incidents” elements recommended by Brookfield, 1995) was distributed, asking several relevant open-ended questions. Participants were asked to anonymously complete these and hand them up.

After filling in their questionnaires, participants were encouraged to discuss their experiences, and these were recorded on audio tape.

Workshops for each case were mostly held one week apart, this being judged an optimum time for reflection: long enough to allow follow-up reflection, but short enough to clearly remember the workshop.

Participants were encouraged to email the facilitator or the organizer for that entity with comments about the previous workshop.

Each workshop (after the first) began with the facilitator summarizing the previous workshop, and requesting further observations on it.

In the follow-up sessions with each case (approximately one year later) participants had another opportunity to reflect on the process. This was done by my distributing a summary report before the revisit, and asking them to read this and think about it. They then had the opportunity to reflect further before the follow-up session began.

The fact that these opportunities for reflection were provided to participants did not mean that they were all taken up. As the primary interest among participants would always be with their own entity, rather than the Process, the above variety of opportunities for reflection was set up in the hope that some of these opportunities would be used.

As for reflection by myself, as Process developer and as facilitator: when facilitating, I kept a diary, filling it in during quiet moments in each workshop, reflecting on the efficacy of the Process itself. Because effective reflection generally involves comparison against a framework (Checkland, 2000), each workshop was compared (using the criteria listed at the end of chapter 3) with (a) the plan for that case, (b) any previous workshops for the current case, and (c) the corresponding workshops for previous cases.

For the Legal Services case, in which I was an observer, and employed separate facilitators, they and I discussed the procedure after each session. This was also possible in the Barossa case, for which I brought along a cartoonist/observer (another PhD candidate) and conferred with the convenor, who had a doctorate in environmental studies and a strong interest in futures work.
Before the sessions with each entity, I recorded (a) my expectations, and (b) the methodological decisions needed: what could be tested by varying procedures from the previous case. After completing the work with each case, I re-evaluated the recently produced scenario map and its annotations, asking whether it was as clear as it could have been, and whether there was a better way to present the information. Sometimes this was done in conjunction with key members of the entity – if they had time, and were still interested. I then considered whether some changes should be made to the method for the next case, and if so, what changes were appropriate. This parallels Dewey’s (1910/1991) reflection sequence: the expected, the observed, differences between the two, reasons for the differences, and resolution.

### 7.2.7 Follow-up evaluation

The plan called for the final session with each organization to be an evaluation and review of the process. This would be about a year after the main study: long enough to evaluate the usefulness of the process, but not so long that the study would have been forgotten. The follow-up was planned as a single meeting with the group of participants at each entity studied (or as many participants as remained available). Before the follow-up meeting, prior participants would be recontacted, and asked to list the main events that have affected the organization in the time since the scenario mapping exercise.

At the actual meeting, with a planned duration of several hours, the scenario map from a year earlier would be reviewed. Four elements would be under review, corresponding to the levels of the Kirkpatrick model for evaluating training (Kirkpatrick, 1994):

1. participants’ satisfaction with the project
2. current situation on the scenario map
3. changes in the behaviour of participants
4. change in the entity, brought about by the Process.

As noted in chapter 9, the Kirkpatrick model was later modified to suit the evaluation of the development of a social inquiry process. At the same time, the fulfilment of the execution criteria developed in chapter 3 would be assessed, by myself and participants working together, with any disagreement being recorded.
7.2.8 Ethical issues in group work

When humans are communicating, there is always a possibility that ethical difficulties can arise. As noted by Stufflebeam (1991, p258) on the development of evaluation standards, there are potential conflicts between ethical principles, so “evaluators need to recognize and deal as judiciously as they can with such conflicts.” The solution is not to refrain from all social inquiry, on the ground that harm may be possible, but to be aware of that possibility in the course of inquiry, and take all practicable steps to avoid harm to participants and their entities. In a workshop situation, harm is perhaps most likely when one participant has power over another (typically an employee/employer relationship), and the less powerful participant says something in the workshop that annoys the more powerful one. One suggested solution (by Murphy et al, 1998) is never to have employer and employee in the same room, but this could weaken the Process by narrowing the range of participant stakeholders. Also, it does not guard against a situation in which one of two peers present later becomes the other’s superior. As it is not possible to design a social process to avoid such eventualities, the best solution lies in the execution of the process, and in the facilitator’s sensitivity and situational awareness. As a member of the Australasian Evaluation Society, I follow that organization’s ethical principles (www.aes.asn.au). In consultancy work with groups, I also observe the code of ethics of the International Association of Facilitators (www.iaf-world.org) but could not fully do so in this study due to part of item 4: “We avoid using processes, methods, or tools with which we are insufficienly skilled.” However, the developmental nature of the Process was made very clear to all participants, both in writing and verbally. Also, because of my previous experience in facilitation, only some aspects of the Process were new; the group techniques were fairly standard.

7.3 Limitations and conclusion

Following the taxonomy of possible futures formed in chapter 3, this work is subject to limitations of place, language, and time. In terms of place, most of the fieldwork occurred in South Australia, and for that matter mostly in its capital city, Adelaide. However, a pilot study was done in Indonesia, and an international case was developed using secondary data. In terms of language, the research was done only in English – though the preliminary trial in Indonesia was bilingual, using both English and Indonesian. In terms of time, all the development occurred between 2001 and 2004, but mostly in 2003 and 2004.
7.3.1 Representativeness of the sample

There is also a possible limitation of representativeness. A potential source of difference between the sampled population and the final sample is the differential unwillingness of some types of group to participate in the study. As the development of the technique involved approximately one week’s time for those involved, groups without a strong interest in the future were unlikely to take part in the study. For businesses, this involved an investment of about one week’s salary for each staff member involved. For voluntary groups, it involved a decision by participants to give up time that they could have spent in other ways.

Within cases, this limitation was more obvious in some cases than in others, and is discussed under each case in the following chapter. A more serious problem is related to the representativeness of cases themselves. Social entities that were approached, but not (through their own choice – or delay) not included as a case, may have produced different results had they taken part in the study. Comparing the potential cases that resulted in case studies with those that did not, the clearest difference was that the former tended to be organizations facing imminent major change, while there was little evidence of this for the latter cases. Thus the tentative conclusion is that the sample used was biased toward organizations expecting major change in their circumstances. While this obviously made them more receptive to undertaking futures work, this would also apply to most other foresighting methods as well as SNM.

7.3.2 Researcher effects

In qualitative research, the “instrument” is the researcher; thus Guba and Lincoln (1992) refer to the “human instrument.” There is a strong argument that it was not valid for me (as a human instrument) to both develop and evaluate this Process. How can a reader of this thesis be sure that, even were I scrupulously impartial in administering the evaluation questionnaires, I did not subtly bias the participants’ reactions? (With a master’s degree in social psychology and 20 years’ experience in facilitation and moderation, this would not have been difficult to do deliberately – though that same background makes it clear that this could have changed responses to only a very limited extent.) Secondly, in the absence of experience of other methods of scenario development, might not a favourable evaluation by participants be a form of expectancy effect (Draper, 2004): the result of any futures studies exercise at all?

The safeguards against these effects were as follows.

1. The evaluation questionnaires completed by participants were anonymous.
2. Participants were asked to evaluate the method itself, rather than my facilitation of it, or their participation in it.
3. Evaluation was carried out at several stages during the workshops: immediately after each session, with a general review at the end of the final session main for each case, and the final evaluation approximately a year later.

4. Following the findings of Tannenbaum and Yukl (1992) on measuring the effectiveness of training, the focus was on the viability and potential usefulness of the Process, rather than on its being a pleasant experience for participants.

5. To reduce any chance of expectancy bias, during fieldwork I avoided forming detailed hypotheses (as distinct from reflections) about the Process. My working proposition was that it should be possible in some way to create a viable method within the four broad design parameters (participative, visually oriented, holonic, and scenario-based). Initially, even the issue of whether the approach should be quantitative or qualitative was left open.

Above all, because the research was formative rather than summative, my motivation was to improve the Process through successive iteration. The guiding objective was to improve the method to a point where it could be readily demonstrated to and used by others.

7.4 Review of this chapter

This chapter has covered the sample design and the fieldwork procedure used in the case studies. Though the sample selection method described in section 7.1 was faithfully followed, the “ideal” fieldwork method described in section 7.2 had to be varied somewhat to suit particular circumstances of each case. In summary, the sample selection involved:

- Creating a theoretical sampling frame using a faceted model.
- Maximum-variation sampling from an infinite population.
- The sampling unit was a social entity.
- A sample of six primary cases and one secondary case.

The fieldwork involved four half-day participative workshops, focusing in turn on

- (a) the entity’s past
- (b) its present situation
- (c) its future – from several sources
- (d) its future – integrating those sources.

Approximately a year later, each case was followed up to determine the extent to which its future had been anticipated, and any after-effects of the case study.