Reference papers **(Click on title to jump to paper or use bookmark)**:

- **Qualifying Quality - A Framework for Supporting Quality-Informed Decisions, Discussion Paper, Bill Allen, June 2002**
- **The Importance of a Quality Culture, Dennis Trewin**
- **Overview of ABS Quality Program, internal paper presented to the joint ABS/SNZ Methodology Senior Managers' Meeting, Narrisa Gilbert, November 2008**

Qualifying Quality - A Framework for Supporting Quality-Informed Decisions

Discussion Paper

Bill Allen

Australian Bureau of Statistics

June 2002

Table of Contents

"We assist and encourage informed decision-making, research and discussion within governments and the community, by providing a high quality, objective and responsive national statistical service."

Mission Statement for the Australian Bureau of Statistcs

1. Introduction

The role of the Australian Bureau of Statistics (ABS) is often interpreted in the context of disseminating important social and economic data, which then form the basis for informed decisions. These decisions can then be considered 'data-informed decisions'. While this is a worthy and admirable role for the ABS, this paper proposes an alternative paradigm.

This new paradigm parallels a change in the interpretation in the concept of quality towards a broader concept of quality, where quality is defined as 'fitness for purpose' and is judged by the user. In keeping with this change in focus, the role of the ABS needs to consider how to assist the user in both making this assessment and applying the results of that assessment appropriately. This is not to suggest that the ABS needs to sit hand-in-hand with the user, verifying the way the data are being used is appropriate, but rather considers how the ABS can assist the user through providing the user with the necessary information and education for them in how to use that information. In doing so, the user is prompted to consider not only the data values, but also the quality of the data. Thus, the decisions evolve from 'data-informed decisions' to 'quality-informed decisions'.

This paper describes a 'decision cycle' which is used to provide a framework for the process of making a quality-informed decision. Each stage of the decision cycle is overviewed, followed by a discussion of the corresponding role for the ABS. Two primary tools are also introduced: the Quality Declaration; and the Quality Assessment. The Quality Declaration is used to document the quality of a data source, while the Quality Assessment is use to document the fitness for purpose of a data source against a specific data need. Together, they provide the user with the basic information necessary to make a quality-informed decision through the application of appropriate risk management strategies.

2. Executive Summary

The proposed framework looks at the role of data and the quality of data from the perspective of the data user and their underlying decision-making processes. In doing so, it highlights the importance of properly defining data needs, making available descriptions of the quality of data through Quality Declarations, comparing the identified data need with the data source as part of a Quality Assessment and implementing appropriate risk management strategies into the decision-making process to take into account that the data need and data sources do not perfectly align.

The role of the ABS is discussed in the light of this framework, both in identifying existing activity and in considering opportunities to focus and build upon this activity to better facilitate quality-informed decisions in government and the community. This can be broadly summarised as follows:

- In terms of defining data needs, the ABS can assist through both general user education (e.g. seminars and training courses) and targeted direct ABS involvement through avenues such as ABS outposted officers, subject areas and user consultation.

- In describing existing data sources, the ABS has an important role to play both in completing and disseminating Quality Declarations for ABS data and through disseminating guidelines and templates so that other data providers are well placed to document their own data collections using Quality Declarations.

- Finally, the ABS also has a key role in the risk management process through the quality assurance of ABS collections and in assisting potential users of the data is the appropriate use of data in the decision-making process.

3. The Decision Cycle

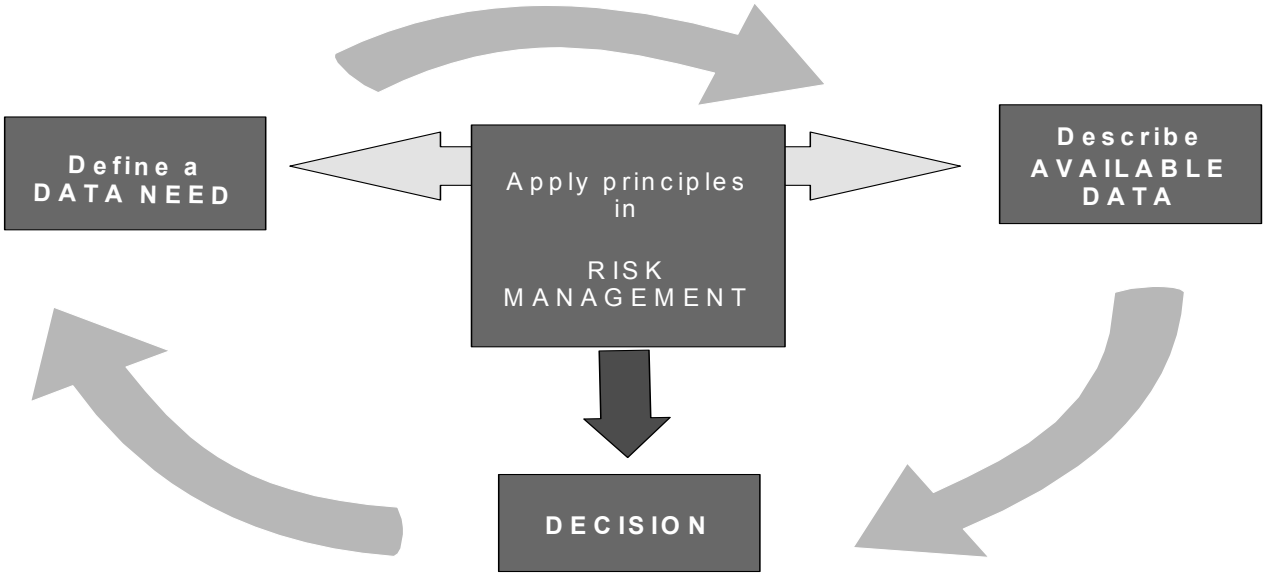
The decision cycle refers to the process which starts with a person wanting to make some underlying decision and finishes with the decision being made. A typical but non-preferred version of the decision cycle might go something along the lines of:

- Need to make a decision
- Want the decision to be based on data
- Look for data until something that looks close enough is found
- Use the data (as if it were perfect)

This decision cycle is focused on assisting in data-informed decision-making, but fails to take account of the quality of the data. Only a crude assessment is made of the fitness for purpose of the data and the degree to which the data are fit for purpose plays no role in the decision-making process.

Figure 1, below, describes a decision cycle proposed to facilitate quality-informed decision-making.

Figure 1 - The Decision Cycle



The decision cycle starts with the decision. The decision defines the desired outcome and the specific data need. This data need should be clearly defined in terms of the information required to facilitate the underlying decision. In essence, it describes the perfect data source for the decision required.

The next stage is to describe existing data sources. This description is called a Quality Declaration and forms the basis for the next stage as quality-informed decisions require an understanding of the quality (i.e. fitness for purpose) of the existing data sources. For the decision-maker, ideally, the range of possible existing data sources have already been appropriately documented and all they need do is reference the Quality Declarations for use in the next stage.

The final stage centres around the application of the data to the decision. The data need and data sources are compared against each other and mismatches are identified as the fitness of the data sources is assessed against the specific requirements of the underlying decision. This information is recorded in a Quality Assessment. Greater mismatches infer greater risks in using the data, so risk management principles are adopted which influence the final decision. Thus, just as the values of the data are used to make the final decision, the information about the quality of the data are also incorporated into the decision-making process to produce a quality-informed decision.

4. Defining Data Needs

The process of defining data needs draws its strength from two underlying models. The first model is the Inputs-Transformations-Outcomes (ITO) Model and is used to focus the data need on the underlying decision. The second model is a data quality framework and provides structure to the process of describing a data need.

The process of first defining the data need independent of existing data sources and

then comparing the need against existing data sources is essential to avoid the common mistake of defining data needs according to what data sources are available. Thus any shortcomings in the data are explicitly identified, considered and acted upon.

The process of defining a data need is detailed further under **4.3 Framework for defining data needs**.

4.1 The Inputs-Transformations-Outcomes (ITO) Model

The ITO model was developed by John Smyrk of the Australian National University and has been incorporated into ABS's project management framework. The model defines a project structure in terms of its objectives (outcomes), its deliverables (outputs) and how the project inputs are transformed, via outputs, into outcomes. Outcomes remain the ultimate objective, whereas the outputs are the physical deliverables that help achieve the outcomes.

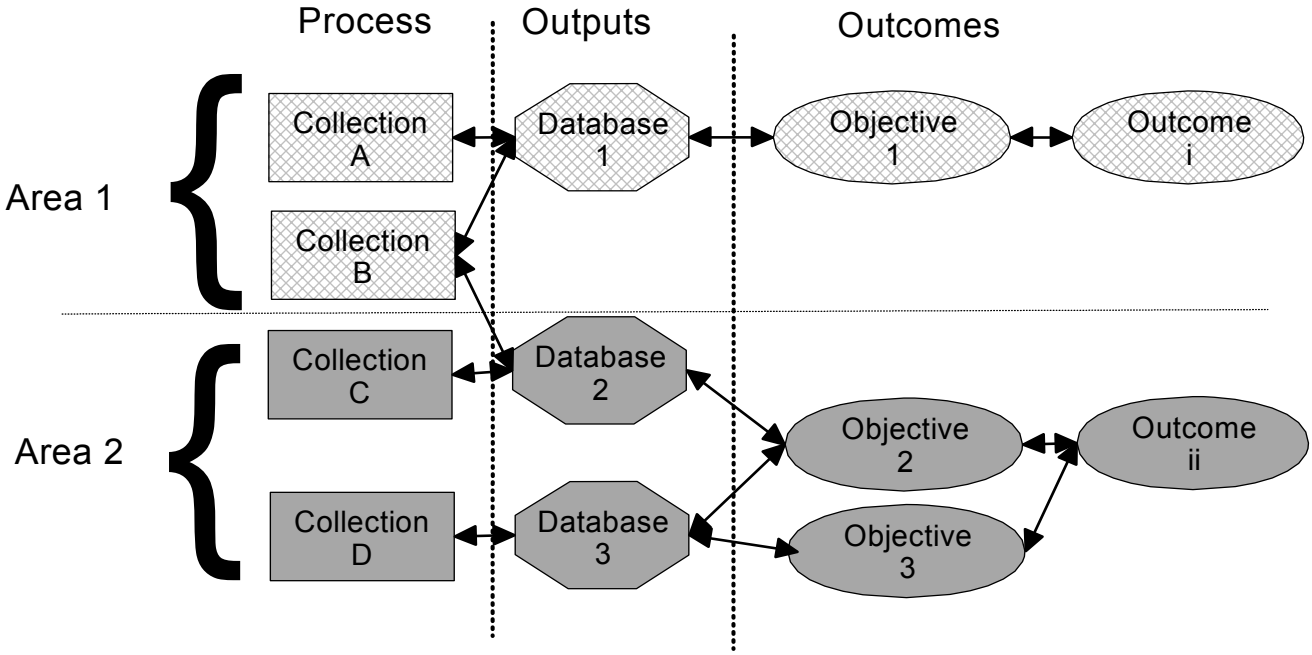
In the broad context of the ABS, the outcome might be considered to assist in informed decision-making, whereas the outputs might include deliverables such as statistics in publications and on the ABS Website, and consultancy devices. For policy makers, the outcomes would often correspond to 'real-world impacts', with the outputs corresponding to policy implementation and evaluation and the inputs corresponding to policy development.

In the context of quality-informed decisions, the underlying structure can be broadly interpreted as follows:

- the *inputs* correspond to the resources used in collecting data;
- the *process* corresponds to the data collection process;
- the *outputs* correspond to the data collected and the way the data are collated, stored and made available; and
- the *outcomes* correspond to the underlying reason the data are required.

This can be expanded further by recognising *objectives*, which relate specifically to the aims of a data collection. The relationships between processes, outputs and outcomes have been broadly summarised in *Figure 2* below:

Figure 2 - Processes, Outputs and Outcomes



In this example, there are two separate areas (Areas 1 and 2), responsible for specific objectives and outcomes. Area 1 manages Collection A and Collection B, while Area 2 manages Collection C and D, but also uses data from Collection B.

As shown in the above example, a database may access data from more than one data collection. Similarly, an objective may reference more than one database and outcomes may have multiple objectives. Also, some data collections will feed into more than one outcome. This reinforces the fact that data collections are not always custom-built for all needs and foreshadows the need for individual Quality Assessments.

The quality of the *processes* and *outputs* are measured using a data quality framework in the context of the *outcomes*.

4.2 The Data Quality Framework

The data quality framework proposed for incorporation into the decision cycle is based on a framework developed by Statistics Canada, which identifies six key dimensions of data quality:

- Relevance
- Accuracy
- Timeliness
- Accessibility
- Interpretability
- Coherence

This data quality framework has been published internationally (Brackstone G., *Managing Data Quality in a Statistical Agency*, (1999) Survey Methodology, Vol. 25, no. 2, Statistics Canada) and has been recommended by the ANAO as 'better practice' in specifying performance measures (*ATO Performance Reporting under the Outcomes and Outputs Framework, Australian Taxation Office*, Audit Report No.46 2000-01, pp63-64.) on advice from the ASB Statistical Consultancy Unit.

More specifically, the six dimensions of quality can be described as follows:

Relevance - The ***relevance*** of statistical information reflects the degree to which it meets the real needs of clients. It is concerned with whether the available information sheds light on the issues most important to users. Relevance is generally described in terms of key user needs, key concepts and classifications used and the scope of the collection (including the reference period). These components are then compared against specific user needs to assess relevance.

Accuracy - The ***accuracy*** of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. It is usually characterised in terms of error in statistical estimates and is traditionally decomposed into bias (systematic error) and variance (random error) components. It may also be described in terms of major sources of error that potentially cause inaccuracy (e.g. sampling, non-response).

Timeliness - The ***timeliness*** of statistical information refers to the delay between the reference point (or the end of the reference period) to which the information pertains, and the date on which the information becomes available.

Accessibility - The ***accessibility*** of statistical information refers to the ease with which it can be referenced by users. This includes the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of *accessibility* for some users.

Interpretability - The ***interpretability*** of statistical information reflects the availability of the supplementary information and metadata necessary to interpret and utilise it appropriately. This information normally covers the availability and clarity of metadata, including concepts, classifications and measures of accuracy. In addition, interpretability includes the appropriate presentation of data such that it aids in the correct interpretation of the data.

Coherence - The ***coherence*** of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time. Coherence encompasses the internal consistency of a collection as well as its comparability both over time and with other data sources. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys.

4.3 Framework for defining data needs

The process for defining a data need can be broadly summarised as describing the data need in terms of what an ideal data source might look like. The ITO model is used to help ensure that the data need remains focused on the underlying decision that needs to be made (i.e. the outcome), while the data quality framework is used to help clarify the data need by ensuring that the data need considers all the aspects of quality.

Table 1, below, provides a list of typical issues that need to be considered in defining a data need. This clarification of the data need indicates what is considered 'fit for purpose' and sets the standard against which data sources can be evaluated.

Table 1 - Typical issues to consider when framing a data need

Dimension	***Examples of Data Need Requirements***
Relevance	● How will the data be used in the decision-making process? ● What concepts do we need to measure? ● What population are we interested in? ● What classifications are we interested in?
Accuracy	● What are our sampling error requirements? ● What level of estimates are required?
Timeliness	● How soon do I need the data? ● How recent does the data need to be?
Accessibility	● How are the data made available? ● Will the cost of the data be prohibitive? ● In what forms are the data available (unit record file versus aggregates, electronic versus hardcopy)?
Interpretability	● Are the data sources documented sufficiently to known that the data source matches my needs?
Coherence	● What comparisons will be required over time? ● What comparisons will be required with other data sources? ● Is it important to match with certain standards?

4.4 Role of ABS

The ABS has a key role in the defining of data needs through its leadership of the National Statistical Service. This role can be broadly summarised as assisting users of data (i.e. government and the community) to define their data needs for:

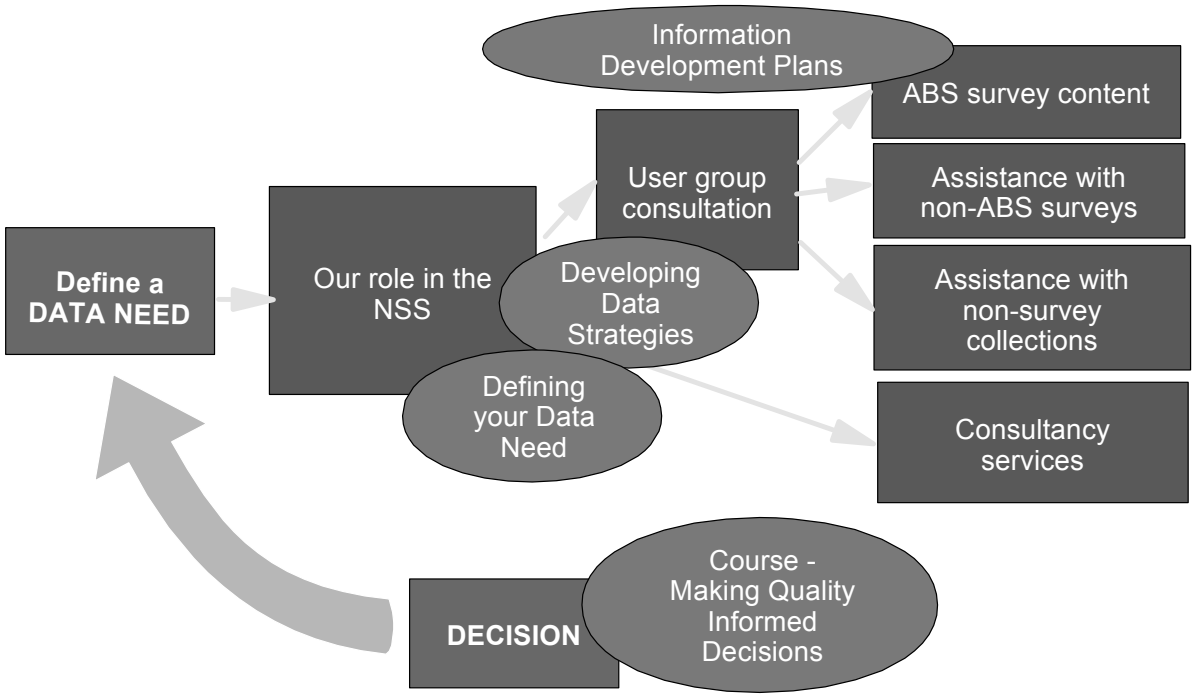
- the development of new data collections; and
- the assessment of existing data sources.

This assistance may be provided either through general user education, or specific involvement through mechanisms such as user groups, outposted officers or the provision of consultancies. Typical examples of specific involvement include:

- provision of internal and external seminars and training courses (e.g. the recently

piloted 'Making Quality-Informed Decisions' or courses on defining data needs or developing data strategies);
- user consultation on the content of ABS surveys;
- direct assistance to other government agencies with non-ABS surveys (e.g. methodological or subject matter assistance); and
- direct assistance to other government agencies with non-ABS administrative collections.

Figure 3 - The Role of the ABS in Defining Data Needs



5. Describing Data Sources

The purpose of describing the quality of data sources is to provide people with sufficient information to assess the 'fitness for purpose' of the data source against their specified data need. This description of the quality is called a Quality Declaration.

5.1 The Quality Declaration

The Quality Declaration focuses on the *process* of the data collection and the data *outputs* by providing a primarily descriptive overview (i.e. *qualitative measures*) of the data collection, supported by a small number of performance indicators (i.e. *quantitative measures*) for those characteristics where it is appropriate (e.g response rates).

For each data collection, the Quality Declaration provides information about the collection's methodology and processes. The Quality Declaration provides the necessary background for primary and secondary users (whose knowledge about

the data collection may be restricted to the information contained within the Quality Declaration) to complete an informed Quality Assessment (as described in **Risk Management**). The Quality Declaration will also be useful as a general reference document, for those needing to gain a broad understanding of the data collection.

It is the role of the Data Collection Managers to complete Quality Declarations for each of their collections. However, there may be data collections referenced within ABS for which there is no Data Collection Manager, primarily because the responsibility for managing the data collection exists fully outside ABS (e.g. managed by another government department). In this case, two main options exist:

- gain the cooperation of the area responsible to complete a Quality Declaration; or
- allocate the responsibility for completing the Quality Declaration to an area within ABS (such as the primary user of the data from the data collection). In this case, it is likely that the area may still need to seek advice from the area rsponsibile for managing the data collection.

5.1.1 Defining the content of the Quality Declaration

A draft template has been developed to assist areas to complete Quality Declarations and can be found in **APPENDIX 1 Template - Quality Declaration**. The Quality Declaration template draws upon the data quality framework described above in **4.2**.

The template identifies a range of characteristics for each of the six dimensions of data quality. These characteristics provide an overview of the associated quality-related issues. For example, the characteristics selected to represent 'accuracy' include the level of sampling error, the response rate, adjustments to data, levels of training and comparability in data values with related data sources. For each of these characteristics, examples of typical qualitative and/or quantitative measures have been provided.

The characteristics included in the template represent a first draft of what could be included in a Quality Declaration. However, the content is still very much open for debate. For example, it might be argued that some of the issues surrounding data security need not be there, while timeliness would benefit from including the timing of the first publication from the data collection and the release (where applicable) of a confidentialised unit record file.

The final content should be agreed by PSG Data Management and Dissemination, ESG Data Management, IMD Data Management and Statistical Consultancy and Training within MD.

5.1.2 Completing in the Quality Declaration

The Quality Declaration should be completed to allow an informed Quality Assessment to be completed. This is not to say, however, that the Quality Declaration needs to be a lengthy document which requires excessive effort to complete and maintain. Rather, in keeping with the concept of quality as being 'fit

for purpose', the Quality Declaration simply needs to be sufficient for users to be able to assess the appropriateness of a data collection for their own requirements.

For each item, only a short qualitative description (one to two paragraphs in most instances), a response to a list of choices or the provision of some quantitative information should be sufficient.

It is recommended that the person completing the Quality Declaration be familiar with the completion and use of the Data Collection Assessments.

For this purpose, **APPENDIX 1 Template - Quality Declaration** includes:

- a descriptive overview of each dimension of quality;
- the data characteristic;
- the questions to be answered by the Collection Manager;
- a definition / explanation of the data characteristic; and
- an explanation of how the information contained in the Quality Declaration might be used in a Quality Assessment.

5.2 Role of the ABS

The ABS has a key role in describing existing data sources, both as a major disseminator of statistics and as a leader for the National Statistical Service.

As a leader, the ABS has a responsibility for providing advice on the content of the Quality Declaration, reaching corporate agreement on a standard template to be used by all ABS data collections (except possibly derived collections as the Quality Declaration is better suited to describe data sources one at a time). As noted above, the final content of the Quality Declaration should be approved by PSG Data Management and Dissemination, ESG Data Management, IMD Data Management and Statistical Consultancy and Training within MD.

A further extension of this leadership role is to disseminate guidelines and templates to the broader National Statistical Service, so that other data providers are also able to document their data collections using Quality Declarations. This approach has already been initiated with the ABS Statistical Consultancy Unit providing advice to the Department of Education, Science and Training on the development of a Data Collection Assessment Framework. Also, as with defining data needs, these methods need to be documented with appropriate learning vehicles (e.g. reference documents, on-line learning, seminars or formal courses). Such strategies should be coordinated where possible with international efforts by government statistical agencies in the field of quality.

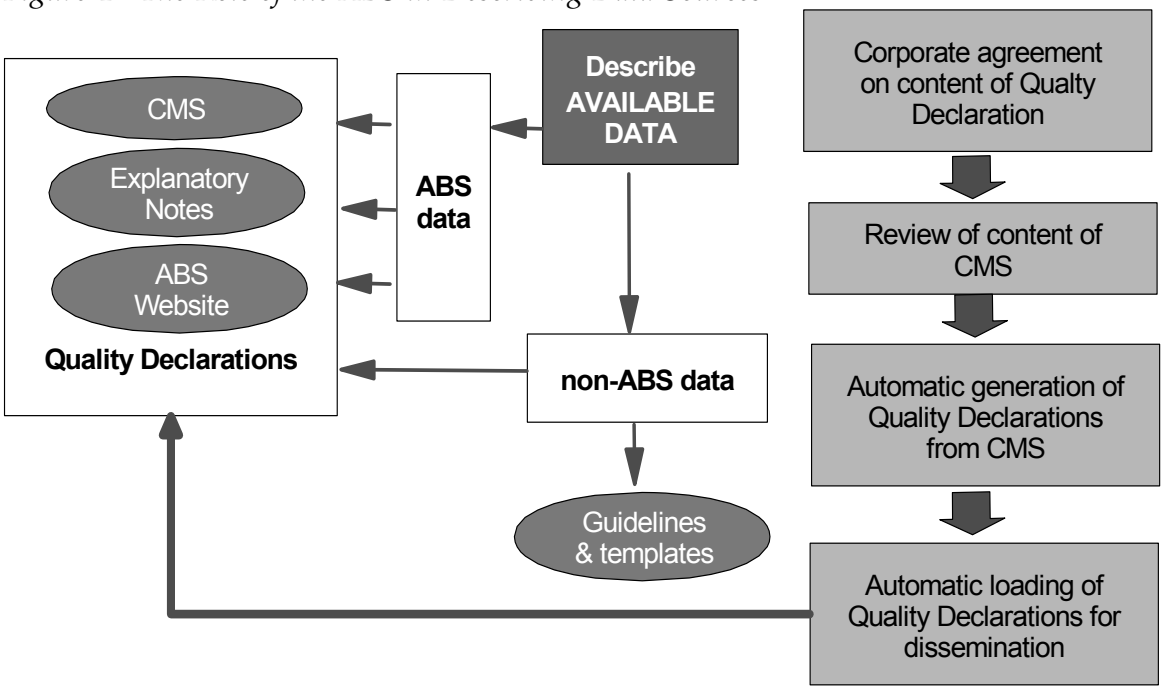
As a major disseminator of statistics, ABS also has a role in disseminating completed Quality Declarations through avenues such as the ABS Website, the Directory of Statistical Sources and ABS publications. In ABS publications, the Quality Declaration forms a natural base for Explanatory Notes, providing the Explanatory Notes with both a consistent structure across all ABS collections and a basic

minimum information to be included.

However, to be able to readily disseminate Quality Declarations the ABS must first agree on the content of the Quality Declarations. Then, as the corporate repository for metadata, the Collection Management System (CMS) will need to be reviewed in terms of its content and how well its fields will provide the necessary information required for the Quality Declarations. Next, it is desirable for the Quality Declarations to be generated automatically from the Collection Management System. This process has already been tested, with a successful prototype having already been developed.

Finally, it would also be desirable to facilitate automatic loading of the Quality Declarations for dissemination into ABS publications (through PPW) and to the ABS Website.

Figure 4 - The Role of the ABS in Describing Data Sources

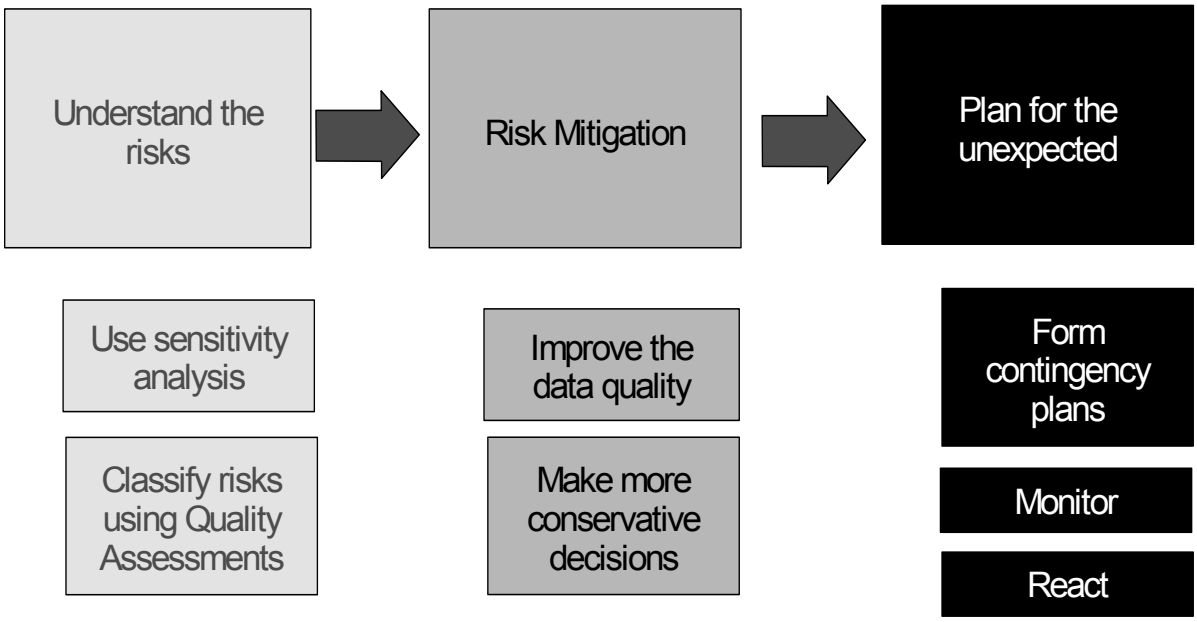


6. Risk Management

Risk management is the process by which information about the quality of the data is integrated into the decision-making process. If the data quality is poor, then the risks of making a poor decision using those data are greater. Conversely, if the data quality is high, then greater confidence can be placed in the information being used to make an informed decision. Thus, the concept behind making quality-informed decisions is to make decisions where the risks of using the data are appropriately managed.

For the purposes on facilitating quality-informed decisions, risk management has been subdivided into three stages which are broadly summarised in *Figure 5 - Risk Management* below:

Figure 5 - Risk Management



6.1 Understanding the risks

The first step in risk management is to understand the risks, and recognises two key steps: sensitivity analysis; and classifying risks using Quality Assessments.

The purpose of sensitivity analysis to identify the various levels of risk associated with using a specific data source for a given data need. This is achieved by examining each of the data characteristics identified in the Quality Declaration and trying to understand the potential impact on the underlying decision of using the data.

As a result of this analysis, each characteristic is classified according to the degree of match between the data need and the data source, ranging from *'the data collection significantly falls short of requirements'* to *'the data collection significantly exceeds requirements'*. This process is documented in a Quality Assessment form.

6.1.1 Sensitivity Analysis

Sensitivity analysis is best explained as the process of considering possible alternate scenarios and their potential impact on the underlying decision. In most cases, this will focus more specifically on considering how the data values might vary across a range of different scenarios and how the different values might lead to different decisions. This can be as simple as going through a process of asking two key questions:

- How different would the data need to be for me to make a different decision?
- How likely is it that the data would be that different as a result of the mismatch between my data need and the data source?

Typical examples would be the impact of sampling error, response rates and mismatches in scope and classifications between the data need and the data source. However, scenarios might also consider issues such as limitations of the type of analysis that can be performed (if there are problems with accessibility, for example).

These scenarios are often, by nature, subjective. For example, for response rates, different scenarios are created based on a series of subjective judgements on how non-respondents might differ from respondents and the likelihood of such a difference occurring. Where possible, additional information should be used to verify these assumptions (such as analyses of non-respondents).

The use of sensitivity analysis and how the different scenarios might impact on the underlying decision is discussed in greater detail in **APPENDIX 1 Template - Quality Declaration** which includes an explanation of how the information contained in the Quality Declaration might be used in a Quality Assessment.

6.1.2 Classify Risks Using Quality Assessments

The second part of understanding the risks is to classify the risks using a Quality Assessment template. The Quality Assessment follows the same format as the Quality Declaration by addressing each of the characteristics within each dimension of data quality, but requires a simple subjective assessment as to whether the specific characteristic for that data collection meets the user's specific requirements (with a short explanation). Thus, each characteristic are classified into one of five categories accoridng to the associated level of risk:

1. *the data collection significantly falls short of requirements;*
2. *the data collection is sufficient with some areas of reservations;*
3. *the data collection is sufficient for the requirements;*
4. *the data collection significantly exceeds requirements; and*
5. *there is insufficient information to judge the suitability of this characteristic*.

In addition, the Quality Assessment requires an indication of the overall suitability of the data collection for the user's requirements.

The purpose of the Quality Assessment is to identify how well the data collection *process* and data *outputs* meet the needs of the user. This includes identifying both limitations (which should impact on the way the data are used) and instances where the data collection exceeds the user requirements (indicating potential areas for savings with regards to that particular user need).

It is the role of the key users or clients of the data collections to complete Quality Assessments and is suitable for assessments from people both internal and external to the ABS.

These assessments can play a key role in assessing the appropriateness of a given collection in meeting the data need associated with a particular objective and can be

used as a basis for improvements to the data collection or making appropriate risk management strategies when the data needs are not perfectly met (as described in **6.2 Mitigating risks**). Thus, it is also in the best interest of the user to complete a Quality Assessment.

For each data collection there would be a separate Quality Assessment for each different objective (noting that this includes both single users with multiple objectives and multiple users each with a single, but different, objective).

A template has been developed to assist areas complete Quality Assessments and can be found at **APPENDIX 2 Template - Quality Assessment**.

6.1.2 Completing in the Quality Assessment

As the Quality Assessment helps identify how well the data meet the needs of the user, it needs to be filled out from the perspective of how the data quality of each characteristic impacts on the way the data can be used (using the information in the Quality Declarations).

These assessments are of the form:

With regards to this characteristic,

- ❑ *the data collection significantly falls short of requirements;*
- ❑ *the data collection is sufficient with some areas of reservations;*
- ❑ *the data collection is sufficient for the requirements;*
- ❑ *the data collection significantly exceeds requirements; or*
- ❑ *there is insufficient information to judge the suitability of this characteristic.*

Comments:

_____
_____
_____
_____
_____

The *Comments* field is used to explain and document why the particular assessment was made. The type of explanation should refer to the user's own requirements and note specific reasons why the characteristic might be considered deficient or exceeding their requirements (or what additional information would be required to be able to make an assessment).

Note, however, that an assessment is not required for all characteristics included in the Quality Declaration. An explanation of why certain characteristics have been excluded from the Quality Assessment has been provided in the template at **APPENDIX 1**.

In addition, a general summary of the Quality Assessment is provided at the start of

each Quality Assessment. This overall assessment takes the form:

In general this data collection:

- ❑ *significantly falls short of requirements for addressing the outcomes;*
- ❑ *is sufficient with some areas of reservations;*
- ❑ *is sufficient for addressing the outcomes; or*
- ❑ *significantly exceeds requirements for addressing the outcomes.*

The main shortcomings are in the characteristics of:

_____
_____
_____
_____
_____
_____
_____
_____

<<provide bullet point list of characteristics rated as significantly falling short of requirements>>.

The main strengths are in the characteristics of:

_____
_____
_____
_____
_____
_____
_____
_____

<<provide bullet point list of characteristics rated as significantly exceeding requirements>>.

Finally, having completed a Quality Assessment, it can be quite useful to transfer the results to a Quality Assessment Summary Table as depicted below in Table 2 below. In the table, each data characteristic is recorded in a cell corresponding to its appropriate dimension of quality and the Quality Assessment it received. For example, if the response rate fell short of the requirements, then it would be recorded in the row 'Accuracy' and under the column heading 'Falls short of requirements'. This table provides a quick overview of the identified risks for using the data collection for the specified data need.

Table 2 - Quality Assessment Summary Table

	Insufficient information	***Falls short of requirements***	***Sufficient with some areas of reservation***	***Sufficient***	***Significantly exceeds requirements***
Relevance					
Accuracy					
Timeliness					
Accessibility					
Interpretability					
Coherence					

6.2 Mitigating risks

In **6.1 Understanding the risks**, risks were identified by comparing the data need to the data source, as documented in a Quality Declaration. The potential impact of these risks were understood through the application of sensitivity analysis and the results were documented in a Quality Assessment.

The next step is to investigate options for reducing the level of risk. This framework has identified two such avenues:

- Improving the data quality; and
- Making more conservative decisions.

6.2.1 Improving the data quality

Improving the data quality primarily deals with looking for opportunities to improve the match between the data need and the data source. These opportunities can be divided into those which involve directly modifying the data collection and those which do not.

Directly modifying the data collection will generally only be an option where the user is identified and accepted as a major stakeholder. Even then, there is a need to balance the needs of one user or a group of users across the needs of all users, noting different decisions will lead to different data needs and therefore different ideal data sources. Also, the degree to which a data collection can be modified and the effort and resources required for such modifications will vary considerably for different data collections. Any decision to modify the data collection should be the result of a cost-benefit analysis with the cost of implementing changes assessed against the benefits

Modifications to a data collection may stem from any of the Data Collection Assessment ratings for characteristics:

- *The data collection significantly falls short of requirements.*

  The modification should be structured so that the characteristic in question will better meet requirements. For example, the scope of the data collection might be

changed to include a broader geographical scope, the sample size might be increased to meet specific user requirements, or the databases might be improved to enable easier access to the data.

Naturally, any such modifications would first need to be fully costed and a decision made as to whether the improvement is warranted. Note, however, that given the characteristic falls significantly short of requirements, it is likely that the characteristic is significantly compromising the quality of the data and potentially the decisions being made on the basis of these data.

- *The data collection is sufficient with some areas of reservations.*

  These modifications will be similar to those above, although the assessment has already identified that the data collection is already nearly sufficient to meet needs. As such, it is likely that efforts will be focused on the more crucial areas where characteristics have fallen significantly short of requirements. However, it may also be that the costs of modifying the data collection sufficiently is comparatively small.

- *The data collection is sufficient for the requirements.*

  While the data collection may meet the specific needs of the user, the user may have identified some potential areas for improvement, which may involve only marginal additional costs or allow for the data to be used more effectively across a wider range of purposes.

- *The data collection significantly exceeds requirements.*

  In these instances, it appears likely some savings might be achieved by reducing what is being offered. For example, the amount of editing might be reduced, it might be decided that a lesser degree of documentation would suffice, or it might be decided that there are insufficient benefits in preserving an outdated classification.

- *There is insufficient information to judge the suitability of this characteristic.*

  This reflects a need to improve the level of documentation. Without the information necessary to make an assessment of the characteristic, it is unknown whether the data will meet needs. In addition to exposing users to making incorrect inferences, it is not possible to identify where improvements or savings can be targeted.

There are also a number of options for reducing the level of risk which do not involve directly modifying the data collection:

- *apply a data collection to only part of the problem.*

  This option accepts that the data source is sufficient for answering part of the

question, but not all of it. For example, the scope of the data collection might only cover specific States but information is required for all States and Territories. Alternatively, only certain information might be accessible meaning that only part of the question can be addressed. In such cases, it might be preferable to only use the data to address only part of the question, particularly when the risks associated with making inferences about the areas not covered by the data source are high. This option combines well with the next option of accessing multiple data sources.

- *accessing multiple data sources.*

  Accessing multiple data sources provides two benefits. Firstly, it provides an opportunity to check for external validation of the data sources through looking for consistency in the data values across different data sources (in the context of their respective levels of quality).

  Secondly, accessing multiple data sources opens the door for using each data source for its respective strengths and using other data sources to cover the weaknesses. This might be as simple as using different data sources for different States and Territories when a national picture is desired. Alternatively, multiple data sources could, for example, be used in synthetic estimation to provide small area estimates. One example of this would be in producing small area estimates for people with disabilities where propensities for disabilities across various demographics can be estimated using survey results which can then be applied to small area data from the Census of Population and Housing.

  Another example would be to use additional data sources to test or refine the assumptions used in the sensitivity analysis. For example, a more frequent survey might be used to estimate trends as a basis for developing scenarios on how much data may have changed since the data collection was last run. Similarly, the confidence intervals for comparable estimates from different data sources could be compared to restrict the likely range of values the true population value might take.

- *deciding more information on the data collection is required.*

  This option specifically addresses the Quality Assessment rating of '*There is insufficient information to judge the suitability of this characteristic*' by suggesting that more research is done to provide sufficient information to assess the suitability of the characteristic in question. Without sufficient information, the risks associated with the data characteristic are high. For example, it might not be possible to know whether the risk of a high level of non-response bias is high as the response rate is not known.

- *deciding a new data source is required.*

  Finally, it may be decided that there is no data source or group of data sources which can be used to adequately feed into the underlying question - the risks

associated with existing data sources are too high and cannot be sufficiently mitigated. In this case, it might be necessary to develop a new data source.

If a decision is made to investigate options associated with developing a new data source, the same process of assessing the proposed data source against the specified data need using the data quality framework should be followed.

It should also be noted that often there are insufficient resources to address all aspects of quality perfectly, so compromises need to be made to achieve an 'affordable level of quality'. However, in making these compromises, two issues need to be considered:

- Where should compromises be made?
- Once compromises have been made, will the data still meet data requirements sufficiently?

6.2.2 Making more conservative decisions

Having identified the risks, it is important that the underlying decision takes these risks into account. In other words, these decisions should take into account the quality of the data as well as the values of the data. As such, making more conservative decisions is specifically aimed at the person making the decision which has generated the data need.

It is difficult to provide specific options here, as the options are dependent on the underlying decision and the corresponding areas of risk. However, it is important to understand that these options do exist. One example might be a decision on whether the allocated budget will support subsidising a new drug as part of the Pharmaceutical Benefits Scheme where more conservative decisions might include implementing localised trials, reducing subsidy levels at first, restricting eligibility criteria for receiving the subsidy (e.g. need a Seniors Card), delaying a decision pending more information or even deciding to use the money to subsidise a different drug where expected usage patterns are better understood.

6.3 Planning for the Unexpected

In the first step, the risks were identified. Then, in the second step, attempts were made to mitigate these risks. In this final step, plans are developed and put in place in case these risks are realised:

- *form contingency plans* in case the identified risks are realised;
- *monitor* what is happening in case the risks are realised; and
- *react* to the realised risks when they are identified, using the prepared contingency plans.

6.3.1 Form contingency plans

In understanding the risks, the Quality Assessment template was used to classify the

risks. The areas which were identified as higher risks (usually classified as significantly falling short of requirements) are the same areas where contingency plans are required (unless the risks were later mitigated sufficiently).

Contingency plans are simply strategies of what to do if certain risks are realised. For example, a low response rate for a survey generates a risk that the survey results are significantly influenced by non-response bias. As a result, inappropriate decisions might be made on the basis of the biased results. This is covered to some degree in sensitivity analysis, but a subjective judgement is still made on the likely degree of non-response bias. Thus, it is important to have a plan in place if later information suggests that earlier decisions were inappropriate. Continuing with the example of the low survey response rate, it would be wise to have a contingency plan in case later information did suggest that the survey had suffered from a higher than expected degree of non-response bias.

These contingency plans should relate to the underlying decision. Using the PBS example referenced earlier on basing funding a subsidy for a drug on existing usage levels, a contingency plan might be to reduce or drop the subsidy if usage levels turn out to be much higher than expected. Similarly, the eligibility criteria for receiving the subsidy could be restricted. These plans can be very similar to those considered early at the risk mitigation stage. However, instead of mitigating the risk immediately through making a more conservative decision, the decision might not fully take into account the associated risks. Rather, the risk mitigation option would only be implemented if further information suggested that the risks had been realised.

6.3.2 Monitor and React

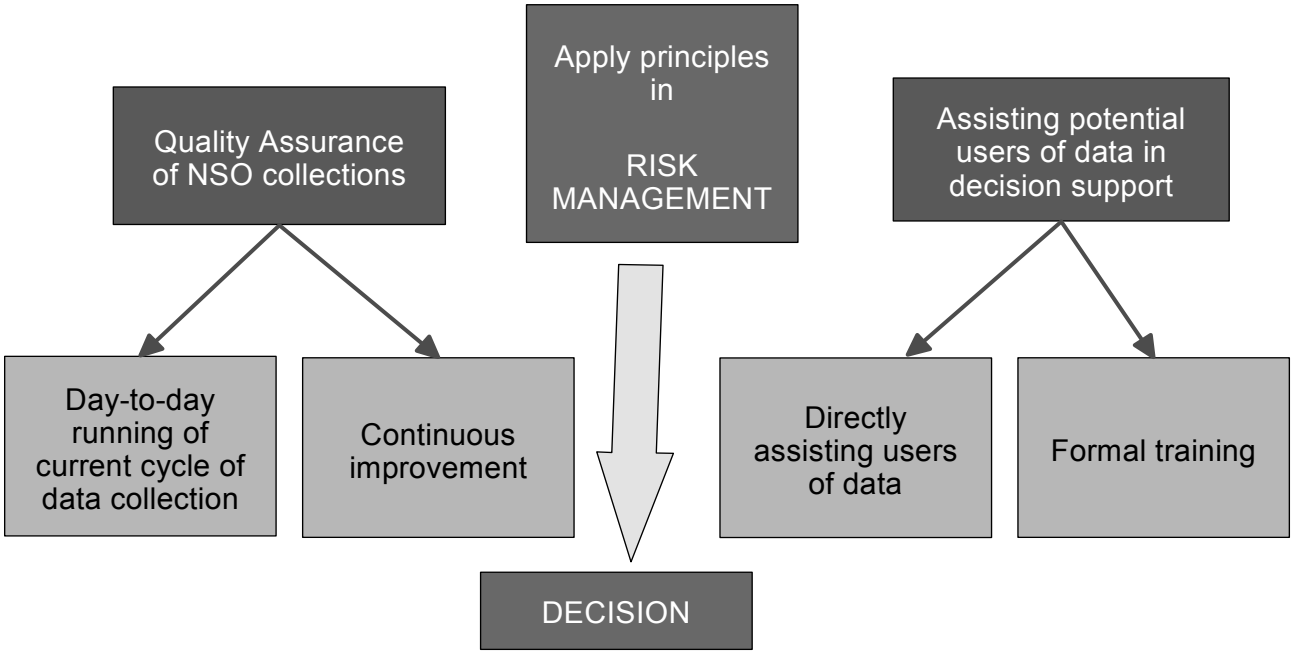
Monitoring is a key part of planning for the unexpected. Having formed contingency plans, it is important that the information is available which will trigger these contingency plans into action.

While it may be possible to continue to monitor data from a regular survey or an ongoing administrative collection, this will not always be possible. As such, it is important to also consider other ways to monitor the impact on the underlying decision. For example, monitoring budgets would assist in avoiding overspending budget allocations. Similarly, a decision to run specialised training programs for the unemployed would benefit from monitoring both participation levels in training programs, participant comments on the training and overall levels of unemployment.

6.4 Role of the ABS

The ABS has two specific roles in risk management, which can be broadly described as quality assurance of ABS collections and 'decision-support' for users of data.

Figure 6 - The Role of the ABS in Risk Management



Quality assurance of ABS collections is an area where the ABS has much experience. With the benefit of the decision cycle and the data quality framework, it is hoped that ABS efforts in this area can be further improved. Quality assurance includes two distinct aspects - the close day-to-day scrutiny of a current cycle as it progresses to survey clearance, ensuring that the data quality is sufficient to be published, and the broader review process which considers where efforts should be focused in the spirit of continuous improvement.

Through continued efforts of developing appropriate quality measures, ABS will be better placed to monitor the data quality and respond accordingly. However, just as data need to be relevant, accurate, timely, accessible, interpretable and coherent, so do the quality measures. Thus the quality measures must be relevant, feeding directly into our own decision cycle. Similarly, automated generation, loading and presentation systems for quality measures will make the quality measures more timely and accessible. Appropriately documenting and presenting the quality measures will aid in their interpretability. In keeping with those principles, it is important that, as the data collection progresses through data processing and estimation, these measures are readily available and continuously updated so that those monitoring the collection are able to respond quickly to problems with the data.

The data quality framework is also valuable for more focused methodological reviews of data and their associated collections. Areas for improvement can be identified using processes comparable to the Quality Assessment process described earlier, with efforts being focused at improving those errors where the greatest (affordable) gains in quality can be realised, or conversly how quality can be best maintained under reduced resourcing.

The second role available to the ABS is that of assisting the real users of the data in 'decision support'. Decision support can be described as assisting users to make their Quality Assessments and then apply the results of these Quality Assessments to their underlying decision through the adoption of appropriate risk management strategies. This can be achieved either through directly assisting the data user (e.g. user group consultations, outpostings, consultancies) or through formal education (e.g. training courses, seminars, on-line learning, reference materials).

APPENDIX 1 Template - Quality Declaration

The template presents the six dimensions of data quality (as described in the main report in **4.2**):

- Relevance;
- Accuracy;
- Timeliness;
- Accessibility;
- Interpretability; and
- Coherence.

For each dimension, the template gives a number of relevant data characteristics which need to be assessed. For example, for 'relevance', some of the characteristics are the scope, reporting unit, frame, classifications and concepts. A definition for each characteristic is provided in the third column. In the last column, there is an explanation of how that characteristic is important from a user assessment perspective. For example, if the 'scope' of the collection excludes people or groups that the user is interested in, the user will need to make some judgements about how these exclusions impact on their decision making capability.

Relevance

The ***relevance*** of statistical information reflects the degree to which it meets the real needs of clients. This is addressed in the Quality Assessment by:

- looking for mismatches in scope, classifications, concepts and data items between what the data collection provides and what the user requires; and
- understanding who the respondents are and how the information is collected.

Looking for mismatches is important because a mismatch tells us that the data collection is not measuring exactly what user wants to measure. As such, it is important to understand the potential impact of the mismatch on the decisions that the user wishes to make.

Understanding who the respondents are (e.g. universities) and how the information is collected (e.g. electronically via e-mail or the Web) is important because this assists in better understanding the limitations of the resulting data.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Scope	What is the population <u>actually</u> covered by the data collection?	Scope includes both the geography covered by the data collection (e.g Victoria) and any other rules used to identify whether a unit is included or not (e.g. exclude people 15 years or under, exclude non-residents) .	If the population covered by the data collection is different to the population you are interested in, then the following questions need to be asked: ● For parts of the population that you are interested in, but are <u>not</u> available in the data collection, are they likely to exhibit different characteristics? ● Can you subset out parts of the population that you are <u>not</u> interested in, but are included in the data collection? If not, are these additional units likely to exhibit different characteristics? In answering these questions, it is also important to remember the impact on totals as well as averages. For example, missing out on a part of the population is likely to mean that the totals will be too low.
Reporting Unit	Who collates and provides the data?	The reporting unit describes who actually provides the data. In some cases, the reporting unit will also be the unit of interest. However, this will not always be the case (e.g. universities might report course information on students). In those cases where the data are provided and collated by different people, please details of both.	The reporting unit is important as the information collected will generally be from the perspective of the reporting unit. For example, collecting information on fields of study from students and the institution they are studying at may well produce different results.
Frame	How is the list of potential respondents compiled? Are there any data quality issues associated with the frame (e.g. new units, defunct units, duplicates, age of frame)?	The results from a data collection are highly dependent on the list used to identify who should respond to the data collection. The quality of this list will have a strong impact on the quality of the data.	For example, a list prepared using the White Pages would exclude households without home telephones or with silent numbers. The issues here are similar to those identified for scope.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Classifications used	What are the key classifications used?	A classification is set of defined groupings or categories - based on common relationships - into which all members of statistical units can be divided or arranged. These groupings or categories can be ordered systematically, are mutually exclusive and exhaustive, and are based on one or more data items. Examples of classifications include: State, Industry; Highest Level of Educational Attainment; Age (in 5 year groupings); and Country. In those instances where classifications used correspond to industry, national or international standards, this should be indicated.	If the classifications used in the data collection do not match up with requirements, then it is important to consider the potential impact of this. For example, a difference in industry classifications may mean that you are unable to exactly measure the industries you are interested in. As with scope, it is important then to assess the likely impact of this mismatch.
Concepts used	Describe any key concepts addressed in the data collection.	A concept in the context of a data collection usually refers to an issue which is often difficult to measure directly (e.g. well-being, some economic concepts) or needs to be derived through several data items (e.g. unemployment, disability). Often the key concepts are the key issues which the primary user is seeking to measure in the data collection.	If the data collection is not measuring the exact concept you are interested in, it will be necessary to assume that the concept you are interested in would produce similar results to those in the data collection, had it been measured. The greater the difference in the concepts, the more tenuous this assumption becomes and the greater the danger that decisions will be made using data which are not conceptually relevant to your needs.
Key data items	What are the key data items collected?	A data item is a particular characteristic which is measured or observed. There are two main types of data items: ● Parametric data items are quantitative measures and have both an associated unit of quantity (e.g. $, hectares, hours) and an associated type (e.g. flow, stock, index, movement). ● Classificatory data items are described in terms of a category (e.g. industry, state, country of birth) rather than using a quantitative or numerical measure.	For the collection to be useful, it needs to collect the information you are interested in. Mismatches in data items will lead to similar problems as mismatches in concepts or classifications.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Mode of data collection	What mode of data collection is used?	The mode of the data collection describes the method used to collect data. Examples include: ● e-mail; ● web; ● Computer Assisted Telephone or Personal Interview; and ● Personal Interview.	The way the data are collected may lead to certain limitations in the data, often relating to the scope of the data collection or the type of information that can be collected using that mode (e.g. personal interviews may cause problems with sensitive questions, but allow the interviewer to better clarify issues with the respondent).
Intended audience and purpose	For what purpose(s) is the data collection run? Who is the primary intended audience for the data?	The intended audience are the primary users of the data collection. In most cases, the data collection will have been designed specifically to meet the needs of these users. The purpose of the data collection is defined as the primary use for which the data will be used by the intended audience.	While a Quality Assessment is not required for this characteristic, this assists in providing the user with an understanding of the broader context of the data collection.
Owner of data collection	Who is responsible for managing the data collection? Who is responsible for deciding on the data items collected?	The data collection manager is defined as the person or position responsible for the operation coordination and running of the operational aspects of the data collection. In addition, a person or group of people will be responsible for deciding which data items are collected or included. This may differ from the data collection manager.	Once again, a Quality Assessment is not required for this characteristic.
Authority	Under what (and whose) authority and/or legislation is the data collection run?	This provides information relating to expected response rates and the general context under which the respondent is required to provide the data. For example, the quality of information collected under an Act of Parliament for the provision of federal funds might be expected to differ from that collected from university administrative records provided on a purely volunteer basis.	Not assessed in Quality Assessment.

Accuracy

The ***accuracy*** of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. As such, it is important to consider issues of both sampling error and non-sampling error (where applicable).

Issues such as mismatches in scope or classifications may also be considered here, but they are addressed primarily under Relevance.

For the Quality Assessment, the user needs to consider whether the accuracy of the data collection will be sufficient to meet their needs. If not, they then need to consider the impact of using the data. This may mean that decisions will be made using data from the data collection, when the underlying information that they are interested in could be significantly different. In other words, the data may be misleading, resulting in poor decisions.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Level of sampling error (applicable to survey samples only)	What are the relative standard errors for key data items? Include relative standard errors of key data items also for key subpopulations.	Sampling error reflects uncertainty in the true population value because information was collected from only a sample of the population. This is often measured as the relative standard error (i.e. standard error of the estimate as a percentage of the estimate). This can be used to identify a range of values that the true value is expected to lie between (e.g. 95% confidence interval).	If the range of values is high, this can impact on the decisions based on the data. For example, if you knew that the unemployment rate was in the range between 0% and 20%, would this restrict the type of policy decisions that you would be comfortable making?

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Response rates	What is the response rate? What steps are in place to attempt to maximise the response rate?	The response rate is calculated by dividing the number of responding units by the number of units which were selected and were in scope of the data collection. Examples of methods used to maximise the response rate include (but are not restricted to): • use of primary approach letters; • interviewers well trainined in establishing a rapport with respondents or the design of respondent-friendly questionnaires; • informing respondents how the results of the data collection will benefit them; and • detailed call back strategies.	In most instances, an assumption is made that the non-respondents would have provided similar information to the respondents. However, the non-respondents may in fact be quite different to the respondents, so the data will be biased to reflect those units which have responded. For example, imagine a data collection on university students where all the overseas students failed to respond. Had the overseas students responded, different conclusions may have been reached. In interpreting the response rate, it is important to consider how your conclusions based on the data may have changed if the non-respondents had responded very differently to the respondents. This is often best handled using a sensitivity analysis approach (see *6.1.1*). In completing the Quality Assessment, first consider how much the data would be likely to change and then consider how that might impact on any resulting conclusions or decisions made. Understanding the steps for maximising the response rate should provide some insight into the potential for non-repsonse bias.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Adjustments to data	What methods are in place for edits and data validation? What data items have more than 10% of units with missing values or have been edited or imputed? For imputed data items, approximately what percentage of units have been altered on the basis of editing or imputation? Are the data subject to large revisions?	Editing is the process of checking data records to ensure that they contain valid entries and changing the records where they do not, whereas imputation is the process of estimating data for individual records which have not been completed. Data validation is a general term for methods used to check that the data appear correct.	The concerns with high levels of editing and imputation is similar to the concerns associated with high levels of non-response. That is, how much are our decisions being influenced by data which didn't come directly from the respondents but were estimated? Similarly, if the data are subject to large revisions, there is a high degree of uncertainty about what the final data will actually be. Consider how much the data might change due to revisions and whether the revised data would lead to different decisions.
Other data issues	Are there any other issues that might impact significantly on the accuracy of the data?	Other issues may also impact on how well the data being collected actually measures what it is supposed to measure. Examples include: ● different levels of data quality for different data items in administrative collections; ● sensitive information; and ● recall bias.	Other issues, such as those listed here, can also influence the data collection's ability to accurately measure what the user actually wants to measure. For example, the respondent may not be able to provide the information with any degree of uncertainty, as they cannot remember the details sufficiently or they are being asked to provide an opinion on something on which they feel they do not have sufficient information. This information also needs to be considered as part of the accuracy of the data.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Level of training	What is the level of training received by those involved in the collection design and operation (i.e. questionnaire design, systems used to collect information, systems for editing and processing the data, etc.)?	Poor training can cause significant problems with the ultimate quality of the data. For example, questions could be misleading or ambiguous so the respondent may not have interpreted the questions as was originally intended. Similarly, poor training for data processing could lead to errors being introduced at data entry.	In making a Quality Assessment, the user needs to consider whether the level of training is sufficient for the data collected. This will be related to the nature and complexity of both the data collection procedures and the data to be collected.
Comparability in data values with related data sources	How does the data collected compare with similar data sources?	Comparability in data values with other data sources offers some insight into whether the data seem to be measuring what the user is interested in (noting that the user's requirements may be sufficiently different to prevent the use of the other data sources).	For this characteristic, the Quality Assessment focuses on whether a possible lack of comparability between the data values from this data collection and other related sources is sufficient to cause some concern with the data collection.

Timeliness

The ***timeliness*** of statistical information refers to the delay between the reference point (or the end of the reference period) to which the information pertains, and the date on which the information becomes available.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Recency of data	What was the last reference period for the data collection? How often is the data collection usually run? When is the next data collection expected to occur?	The reference period refers to span of time to which the data refers. This may either refer to a single point in time or a span of time.	For this characteristic, the Quality Assessment is asking about the suitability of the timeliness of the data. If circumstances are likely to have changed significantly since the last time the data were collected (e.g. internet usage) and the data needs to reflect the current situation, there will be problems comparable to those experienced under the relevance and accuracy dimensions - the data may not be measuring what the user wants to measure which may lead to inappropriate decisions using that data. Thus, it may be concluded that the value of the data is limited given that the data are no longer relevant to the current situation.

Accessibility

The ***accessibility*** of statistical information refers to the ease with which it can be accessed by users. This is addressed in the Quality Assessment by considering:

- knowledge that the data exist;
- ease of accessing the data; and
- the security of the data. This aspect is for internal use only and should be removed from the DCO copy supplied to external clients.

This impacts on decisions regarding whether the data collection is an appropriate data source, with respect to ease of obtaining the data, its security and the impact on any dissemination of results.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Ease of getting data	What is the average time taken to fulfil a data request? What data are readily available on the Web? What publications are available, and where are those publications available? What is the associated pricing policy?	A data request will generally refer to a request for tabulated data. The complexity of the data request may vary for different requests so consider the average time required to meet a request of 'average' complexity. The pricing policy is the set of rules or guidelines for determining the cost for a user to purchase data.	Even having received permission to access, it might prove too difficult to get the data in a suitable form or it might take too long to get the data. Similarly, access to the data may prove to cost too much given your available resources.
Knowledge data exists	How are people internal to the department made available about the existence of the data? How are people internal to the department made available about the existence of the data?	Knowledge that the data exists is an important aspect of the accessibility of a data collection. This includes how information on the data collection is made available both internally (e.g. on the Collection Management System with the ABS or externally (e.g. on the Web or hardcopy publications in most libraries).	In the comments field, the user should indicate how they became aware of the data and how easy it was for them to locate the data. It is expected that ABS data are listed and documented on the Collection Management System.

Interpretability

The ***interpretability*** of statistical information reflects the availability of the supplementary information and metadata necessary to interpret and utilise it appropriately. Interpretability has been addressed in the Quality Assessment by asking the questions:

- Is there sufficient information to make an informed Quality Assessment on all characteristics?
- How easy is it to obtain more information about the data and data collection if required?

If there is insufficient information to understand properly how well the data meets the user's specific needs, then they are in danger of using inappropriate and/or misleading data to make important decisions.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Level of documentation for data collection	Has this Quality Declaration been 'signed off' by the area in charge of the data collection? What more detailed information is available regarding the data collection?	The Quality Declaration is the document as described in this appendix. The level of documentation should be aimed at providing sufficient information for someone without previous knowledge of the data collection to complete a Quality Assessment (without using the assessment of 'insufficient information'). More detailed information might be available through other sources, such as user guides, a website or other documentation maintained by the data collection manager.	The Quality Assessment for this characteristic makes an assessment as to whether the level of documentation in the Quality Declaration is sufficient. Insufficient information to make a Quality Assessment means that there is uncertainty regarding the data quality for that characteristic. As such, any decisions using the data which are affected by that characteristic will be based on data on dubious quality and may lead to inappropriate decisions being made. The Quality Assessment should indicate that the level of documentation is insufficient for those characteristics which have been rated as "*there is insufficient information to judge the suitability of this characteristic* ". The comments field should indicate which characteristics have received this assessment.
Internal accessibility of documentation for data collection	Is the Quality Declaration ready available within the department?	The Quality Declaration should be available for all potential users of the data within the department to access, in case they need to review available data sources for a given need. Ideally the Quality Declaration should be stored on a corporately endorsed (standard) storage medium for documentation on data collections.	In the comments field, the user should indicate how easy it was for them to locate the Quality Declaration.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
External accessibility of documentation for data collection	Is this Quality Declaration available to people outside the Department (Web or Other - please specify)? Is more detailed information available on the Web (Web or Other - please specify)? What level of documentation is provided in publications?	This characteristic refers to the availability of the Quality Declaration to people who do not work within the department. More specifically, these people include anyone who might be interested in understanding the quality of the respective data collection (e.g. academics, policy analysts in other departments). Possible methods for external accessibility would include the inclusion of the Quality Declaration on a Website or in publications released to the general public.	For this characteristic, the Quality Assessment comments on whether documentation will be made sufficiently available for those outside the department. For users within the department, this assessment draws on whether it is important that people outside the department are able to access the documentation (e.g. to support published data). This would also alleviate the degree to which the area managing the data collection needs to be called upon to answer questions about the data collection.

Coherence

The ***coherence*** of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time. This is captured in the Quality Assessment by focusing on changes over time to the data collection as any such changes will impact on any interpretation of how things may have changed over that period. For example, a perceived change in results between two time periods might simply reflect a change in definition. Thus, it is important to know when these definitions have changed and how much they have changed, and considering the potential impact of those definitional changes on the data.

Data Characteristic	***Questions to be answered by Collection Manager***	***Definition/Explanation***	***Relationship to Quality Assessment***
Consistency of classifications over time	List any changes in key classifications over time	To try and maintain or improve the general relevance of a classification, they are often reviewed and updated over time. Examples of classifications which are subject to review include: ● Statistical Local Areas (geographic); ● Collection Districts (geographic); ● Industry; and ● Countries.	The Quality Assessment should consider the changes in the key classifications in the specific context of the user's requirements. Some classifications may not be relevant to your needs or the changes may be minor compared to your needs. Alternatively, some changes may cause major problems in comparing data over time.
Consistency of concepts and methodology over time	List any changes in key concepts and methodology over time.	To try and maintain or improve the general relevance of a statistical concept, they are often reviewed and updated over time. For example, the concept of employment as measured may have changed over time. Similarly, other concepts such as innovation have evolved over time as more research is done in their respective fields. Similarly, changes in the collection methodology may impact on the resulting data. Examples might include changing the data collection methodology or the questionnaire.	The issues associated with this are the same as those listed above for assessing the consistency of classifications over time.

APPENDIX 2 Template - Quality Assessment

The data source being assessed is:

_____

The data need against which this data source is being assessed can be summarised as follows:

_____
_____
_____
_____
_____

In general this data collection:

- ❑ *significantly falls short of my requirements;*
- ❑ *is sufficient with some areas of reservations;*
- ❑ *is sufficient for addressing my data needs; or*
- ❑ *significantly exceeds my requirements.*

The main shortcomings are in the characteristics of:

_____
_____
_____
_____
_____
_____
_____
_____

<<provide bullet point list of characteristics rated as significantly falling short of requirements>>.

The main strengths are in the characteristics of:

_____
_____
_____
_____
_____
_____
_____
_____

<<provide bullet point list of characteristics rated as significantly exceeding requirements>>.

Relevance

Scope	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Reporting Unit	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Frame	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Classifications used	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Concepts used	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Key data items	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Mode of data collection	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Authority	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Accuracy

Level of sampling error	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Response rates	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Adjustments to data	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Other data issues	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Level of training	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Comparability in data values with related data sources	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Timeliness

Recency of data	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Accessibility

Ease of getting data	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Knowledge data exists	I*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Interpretability

Level of documentation for data collection	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Internal accessibility of documentation for data collection	I*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

External accessibility of documentation for data collection	I*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

Coherence

Consistency of classifications over time	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____
Consistency of concepts over time	*With regards to this characteristic,* ❑ *the data collection significantly falls short of requirements;* ❑ *the data collection is sufficient with some areas of reservations;* ❑ *the data collection is sufficient for the requirements;* ❑ *the data collection significantly exceeds requirements; or* ❑ *there is insufficient information to judge the suitability of this characteristic.* *Comments:* _____ _____ _____ _____ _____ _____ _____ _____

THE IMPORTANCE OF A QUALITY CULTURE

Dennis Trewin
Australian Bureau of Statistics

1. Introduction

Fellegi (1996) provides a strong argument that the trust in the national statistical agency is how most users judge the quality of its statistical products.

"Credibility plays a basic role in determining the value to users of the special commodity called statistical information. Indeed, few users can validate directly the data released by statistical offices. They must rely on the reputation of the provider of the information. Since information that is not believed is useless, it follows that the intrinsic value and useability of information depends directly on the credibility of the statistical system. That credibility could be challenged at any time on two primary grounds; because the statistics are based on inappropriate methodology, or because the office is suspected of political biases."

Trust will not happen unless the culture is right. Culture is a word with many meanings but I am interpreting culture as "the way we do things". Core values are important to this. They cannot be just statements hanging on the wall. They have to be understood. They have to be reflected in behaviours, particularly by leaders of organisations.

The Australian Bureau of Statistics (ABS) places great reliance on adherence to its core values. More than anything, they distinguish us with other survey providers. The core values are:

- Relevance - regular contact with those with policy influence, good statistical planning, which requires a keen understanding of the current and future needs for statistics, are essential, as is the need for statistics to be timely and relatable to other statistics.

- Integrity - our data, analysis and interpretation should always be objective and we should publish statistics from all collections. Our statistical system is open to scrutiny, based on sound statistical principles and practices.

- Access for all - our statistics are for the benefit of all Australians and we ensure that equal opportunity of access to statistics is enjoyed by all users.

- Professionalism - the integrity of our statistics is built on our professional and ethical standards. We exercise the highest professional standards in all aspects of ABS statistics.

- Trust of providers - we have a compact with respondents; they are to encourage us to provide accurate information and we ensure that the confidentiality of the data provided is strictly protected. We keep the load and intrusion on respondents to a minimum, consistent with meeting justified statistical requirements.

Adherence to core values is just one element of maintaining a quality culture. Part 2 discusses the steps the ABS uses to maintain a quality culture.

It is now widely recognised that quality is much more than accuracy (eg Brackstone(1999) and Carson (2000)). In Part 3, I will discuss the different dimensions of quality before identifying in Part 4 what I think are some of the major quality challenges for the Australian Bureau of Statistics (ABS) over the medium term. I believe that most of these will be shared by other national statistical organisations.

2. Towards a High Quality Statistical Service

Quality assurance is a responsibility of all staff in the ABS. There is no central "quality management" group although Methodology Division is encouraged to be our conscience on quality issues - a role it takes on with enthusiasm, sometimes to the annoyance of others. However, I see that as a good sign - they are provoking

debate on some of the more difficult quality issues. Support from senior management for this type of role is very important.

The key strategies to ensuring a high quality are described under six broad headings.

- A high degree of credibility for the ABS and its outputs.
- Maintaining the relevance of ABS outputs.
- Effective relationships with respondents.
- Processes that produce high quality outputs.
- Regular review and evaluation of statistical activities.
- Staff who are skilled and motivated to assure the quality of ABS outputs.

2.1 **A High Degree of Credibility**

Credibility is fundamental to the effective use of official statistics. Credibility arises from a system of statistics which provides an objective window upon the condition of a nation's economy and society.

The legislative framework within which the ABS operates is an important pre-condition for the integrity of Australia's official statistics. The Australian Statistician is guaranteed considerable independence by law. This helps ensure that the ABS is, and is seen to be, impartial and free from political interference. In particular, the independence of the Statistician supports his objectivity in determining the statistical work program. Although the legal authority is there, it still needs to be reflected in the way my senior staff and I behave.

Government statisticians must not just apply professionalism skills to their work; they must also be seen to adhere to high ethical standards, especially with respect to objectivity and integrity. We are frank and open when describing our statistical methods to users; we publish information about our performance - for example, in terms of both sampling and non-sampling errors, and revision histories for key series; we are willing and able to identify and address user concerns regarding quality; we are receptive to objective criticism and be prepared to respond quickly

even if the problem is one of perception rather than reality. Also, we promote good relationships with the media as they have a major influence on public opinion of the ABS and its outputs. Also, most Australians find out about official statistics through the media. We engage in other user education activities aimed at fostering intelligent use of official statistics.

The fact and perception of ABS objectivity are reinforced by our policies of pre-announcing publication dates for main economic indicators, allowing very limited pre-release of publications (the details of which are in the public domain), and making special data services available on an even handed basis to all.

2.2 **Maintaining the Relevance of ABS Outputs**

There is, of course, some tension between (on the one hand) being responsive to changing policy needs and (on the other) maintaining the continuity of a system of statistics that can objectively monitor performance. Senior staff of the ABS devote a great deal of attention to maintaining personal contact with key users, so our organisation can gather intelligence about policy issues and emerging areas of economic, social and environmental concern. This includes regular lunchtime meetings with the most senior staff of the government agencies responsible for policy. The Directors of our State offices have similar arrangements with State officials. That intelligence feeds into strategic planning and judgements when our national statistical programs are designed and reviewed.

The ABS has a range of other means for communicating with the users of statistics, to ensure that our products are relevant to their needs. For example, advisory groups representing users and experts in various fields provide valuable guidance to our statistical activities.

There may also be some tensions or trade-offs between the different aspects of quality. The ABS positions itself at the higher accuracy end of the information market, to protect the valuable ABS 'brand name'. But if, for example, there is an urgent demand for data in a new field, some aspects of quality may be traded off in order to achieve timeliness and relevance. Nevertheless, there is a "bar" below which we will not go. Because it is probable that the new statistics will be used to inform significant decisions or debate, the ABS makes very clear statements about the accuracy of the data and the uses for which they are suitable. On occasion, such new statistics may be differentiated from our other products by labelling them 'experimental' or releasing as an information or occasional paper. We regard this form of branding as very important.

2.3 **Effective Relationships with Respondents**

An official statistical agency must maintain good relations with respondents if it wants them to co-operate and provide high quality data. The ABS approach includes - explaining the importance of the data to government policy, business decisions and public debate; a policy of thoroughly testing all forms before they are used in an actual survey; obtaining the support of key stakeholders; minimising the load placed on respondents particularly by using administrative data where possible; and carefully protecting privacy and confidentiality.

The ABS monitors and manages the load it imposes on both households and businesses; we have developed 'respondent charters' for both groups. As well, a Statistical Clearing House has been set up within the ABS to coordinate surveys of businesses across government agencies (including the ABS), to reduce duplication and to ensure that statistics of reasonable quality are produced.

All ABS forms and collection methods are tested to ensure that the data we seek are available at reasonable cost, and the best available ways are used to collect them. For business surveys, our units model, classifications and data items, are designed to be as consistent as possible with the way businesses operate. This now corresponds closely with reporting for taxation purposes, making it easier to integrate survey data with data collected for taxation purposes. For household surveys, the extensive use of cognitive testing tools within the ABS, and the establishment of a questionnaire testing laboratory, have helped to improve quality and to reduce respondent load. Standards for form design and form evaluation are set out in manuals and are promoted and supported by experts in form design.

The ABS uses efficient survey designs to minimise sample sizes and hence total reporting load; we also control selection across collections to spread the load more equitably. To take advantage of the current reforms of the Australia taxation system, the ABS is seeking every opportunity to improve the efficiency of our sample designs, and to use taxation data as a substitute for the data now gathered through direct collections. As mentioned above we have started the process of simplifying the business unit structure used in our surveys to make it consistent with the structure used for taxation purposes.

For household surveys, the introduction of computer aided interviewing has helped to streamline interviewing procedures, reduce respondent load, and improve the quality of data collected.

2.4 **Processes that Produce High Quality Outputs**

The quality of ABS statistics is underwritten by the application of good statistical methods during all stages of a collection including the design stages. The ABS has a relatively large Methodology Division (about 120 staff) which reports directly to the Australian Statistician. The Division is responsible for ensuring that sound and defensible methods are applied to all collections and compilations. Also, the Methodological Advisory Committee, a group of academic experts, provides independent reviews of our statistical methods.

The ABS puts substantial effort into developing statistical standards, including concepts, data item definitions, classifications, and question modules. All ABS surveys must use these standards. The standards are supported by relevant data management facilities to ensure they are accessible and to make it easier to use standard rather than non-standard approaches.

Sample design and estimation methods are the responsibility of the Methodology Division. Our designs aim to minimise cost for adequate accuracy. Where possible, a "total survey design" is used - accuracy requirements are set according to the intended use of the data, and accuracy is measured in terms of both sampling and non-sampling errors. In our business surveys for example, total survey design guides the allocation of resources to the intensive follow up of non-respondents or the editing of questionnaires; the effort for reducing non-sampling errors is optimised according to the impact of errors on overall quality. The cost to data providers is also taken into consideration. The "total survey design" has to be approved by a senior ABS staff member before it is implemented.

In recent years, the ABS has made substantial progress by applying standardised best practice across surveys. For example, business surveys based on the business register now draw their frames at a common date each quarter, and uses a common estimation method to ensure all collections have a consistent and complete coverage. Standards rules are adopted for frame maintenance, field collection and estimation, and generalised processing facilities are available to support the use of these rules. Standard methods are used to allow for "new businesses" not yet included on the survey frame. The ABS is thereby able to ensure the coherence of estimates across different business surveys.

For household surveys, a master sample system has been adopted since the mid 1960's. The system is updated regularly after each five-yearly census, and has been the cornerstone for ensuring the accuracy of statistics collected from household surveys.

Achieving quality in surveys is easier when computer systems support current best practice. The ABS has invested in generalised tools for collection and processing. Generalised tools have been developed for all major processing steps of both business and household surveys, including sample frame management, data input and editing, imputation, estimation and aggregation.

The ABS embraces a rigorous continuous quality improvement approach wherever appropriate. The Australian Population Census is a classic example of raising quality through a strategy of measuring quality and involving all staff in examining and devising solutions for quality problems. This approach was applied very effectively at the data processing centre for the 1996 Census; the centre achieved significant savings, better quality and an improvement in timeliness. Continuous quality improvement is also applied to the coding of businesses on the business register, and to many other ABS processes.

At the output end of collections, each subject group is required to confront its data with other ABS data and with external information, to ensure the coherence of our statistics. Most of the key macroeconomic have to be "signed off" by the national accountants in meetings established especially for the purpose of clearing the statistics. The national accountants then have an obligation to use this data without further adjustment in the compilation of the accounts, enhancing consistency. More generally, confrontation of different data sources is undertaken by our national accountants Since 1998, the ABS has adopted an 'input-output approach' to compiling national accounts estimates, and the new methodology has led to more consistent accounts. In particular, the rigorous data confrontation and balancing process at detailed levels have helped to identify data deficiencies and have suggested methods for reducing discrepancies between estimates. Information about quality is fed back to the economic collection groups and is resulting in a more focused approach to improvements in the quality of source data.

One important quality improvement initiative that the ABS has pursued is the development of an Information Warehouse to manage and store all of our publishable data. By drawing together different datasets into a single database, the Warehouse enables our statisticians to confront statistics produced from different collections. Furthermore, as all forms of publication, be they paper based or electronic, are to be produced from a single data store, with the objective of ensuring that data released in different products, and at different times, are consistent.

Another important element of quality management is documentation. Good documentation supports review activity and facilities the dissemination of quality information to users, so they can assess the fitness of the data for the purposes they have in mind. As part of the Information Warehouse initiative, the ABS can now enforce standards for documentation of the metadata that describe concepts, definitions, classifications and quality.

A relevant and responsive statistical service must do more than provide data to clients. The ABS has also recently strengthened its analytical ability. A team of analysts has been set up to develop new measures of socioeconomic concepts, to explore relationships between variables and to prototype new analytical products. The expanded program of analysis work is also expected to deliver significant benefits in the form of insights into data gaps and quality concerns.

2.5 **Review and Evaluation of Statistical Activities**

Each ABS area is responsible for continuous quality review and improvement. For statistical collection areas, quality management is supported by sets of performance indicators. A standard set of measures has been developed to permit a comparison of quality across collections. Tools are now being developed to calculate these measures as part of our normal survey processes, and the Information Warehouse will allow us to store and display the measures. The key indicators are also included in the annual reports each Branch makes to the ABS Executive for review.

Quality measures are also of interest to the users of statistics. The Information Warehouse will improve users' access to information about quality issues. As well, the ABS places high priority on helping users understand the quality of data and its implications for them, and has adopted active education strategies to promote such understanding. As highlight in Lee and Allan (2001), there is much to do to improve user understanding of quality.

Each ABS household survey now includes an evaluation program which reviews the effectiveness and efficiency of all survey activities and assesses the extent to which the data are used by clients. The Statistical Clearing House conducts a review of each ABS business survey. These initiatives ensure that all collections are subjected to at least a basic evaluation, and brings to light opportunities for improvements to our quality and efficiency.

As well as making internal comparisons of performance across its own collection areas, the ABS has established a benchmarking network with overseas statistical agencies; the aim of the network is to share information about survey design, processes and costs. The benchmarking exercise is providing invaluable guidance to the ABS's efforts to improve its processes and outputs.

Finally, each subject matter area and support area in the ABS provides an annual report to the senior management forum, highlighting progress and problems in the area and providing performance measures.

2.6 **Skilled and Motivated Staff**

The ABS could not provide high quality information to its user community if it did not employ people who bring skills and energy to our statistical work. The staff are responsible for implementing the strategies discussed above. They must take a professional approach and be committed to the development of new methods, to continuous quality improvement, and to the open discussion of methods and quality issues.

Quality improvement and on-going statistical work compete for the time and energies of our staff. The ABS's approach is, as far as possible, to integrate quality work with on-going processes and systems. We also emphasise to staff that quality management is a corporate priority and ensure that tools and resources are made available to support it. In particular, the ABS is implementing a tighter approach to project management; this is being supported by manuals, systems and training.

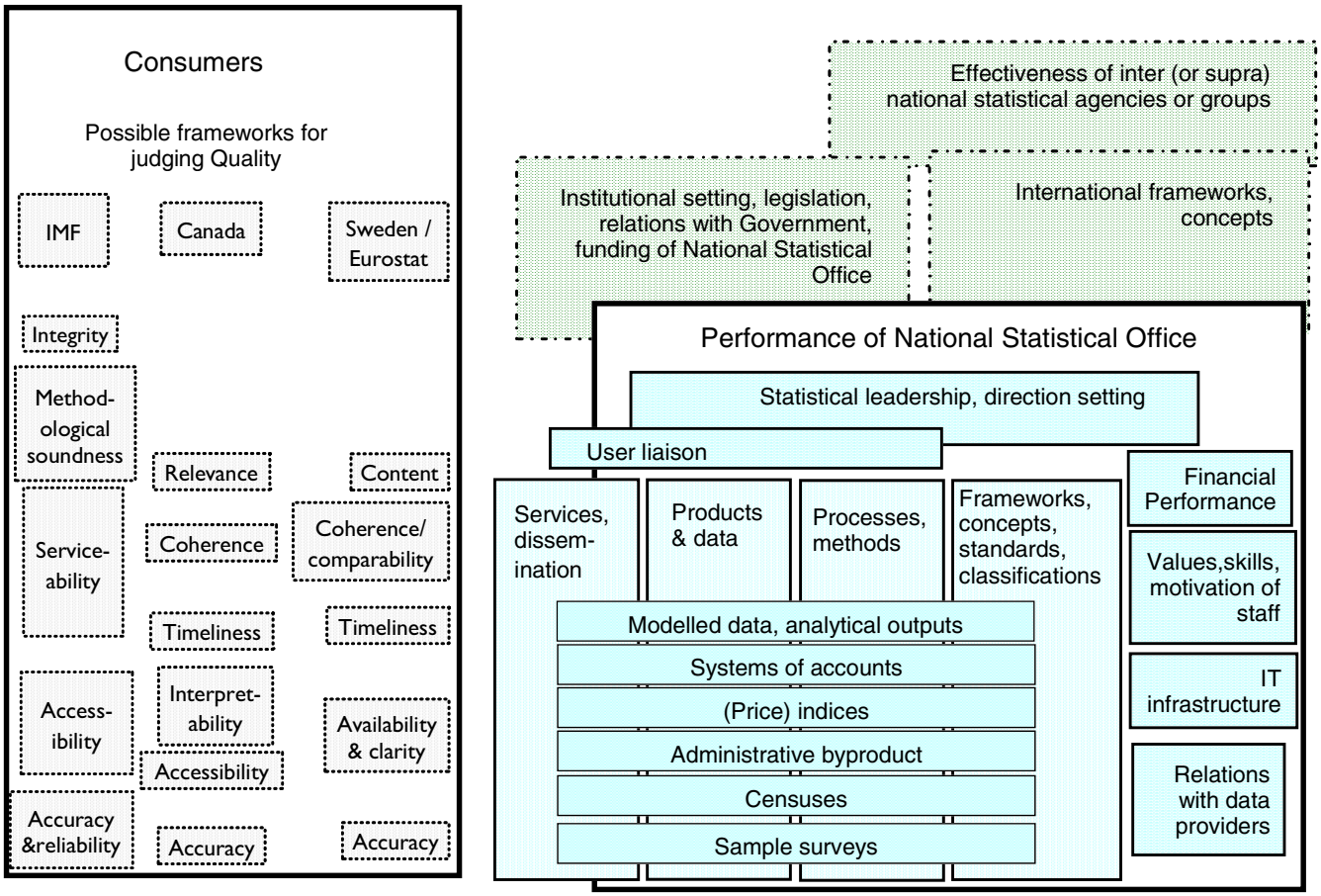
Statistical training plays an important role in maintaining and improving quality. The ABS is always searching for new, more effective, approaches to skills development. An important element of our performance management system is a focus on identifying and addressing individuals' development needs.

Relationships with other national and statistical agencies are also a very important element of the ABS's efforts to improving official statistics. The ABS is committed to using international standards; we take advantage of the wide range of expertise embodied in those standards. On the other hand, there is an obligation for us to make a positive contribution to the development of the standards. In doing so, we try to take account of the interests of the Asia/Pacific region as well as those of Australia. With ever increasing globalisation of economic activity and the pursuit of world wide social goals, the compatibility between Australian's statistics and those of other countries, is an important element of quality. The ABS maintains strong links with many overseas agencies. We are fortunate that there is a lot in common in the challenges we face and there are great benefits from sharing experiences with other statistical agencies.

3. Dimensions of Quality

Figure 1 is taken from Lee and Allen (2001). Among other things, it neatly summarises, on the left hand side, three existing frameworks for judging quality. There are some differences with the descriptors used but basically they are providing the same message - there is much more to quality than accuracy. This is now widely accepted although it was not so long ago that discussion of the quality of a statistic focussed on its accuracy and the sampling variability in particular.

Figure 1



There are also several messages in the right hand side of Figure 1.

(i) There are many different ways of compiling official statistics - from modelled data/analytical outputs to sample surveys. The quality challenges differ between the different means of compiling statistics and in Australia we are making greater use of administrative data, systems of accounts (linked to the national accounts) and models or analytical methods to produce statistical outputs, compared with five years ago.

(ii) There are several groups of activities associated with statistical outputs - from "frameworks, concepts, standards and classifications" through to "services/dissemination". Each is important in its own right and has its own quality challenge.

(iii) The Performance of a National Statistical Office is extremely important to its quality image as recognised in the opening quote of the paper. A number of the elements are specified in Figure 1. All are important. Indeed you cannot have a high performing statistical office unless you rate well against each of these elements.

(iv) There are other elements such as institutional settings which are also an important element.

My main purpose is describing the above is to emphasise that the list of quality challenges for a national statistical office is very large. All have to be tackled in some way - this would not be possible unless you have a quality culture ie attention to quality is the responsibility of all staff. There are many "moments of truth" to genuinely test the quality culture.

Psychologists say that it is not possible to grasp more than seven points at one time so I have limited myself in the remainder of this paper to identifying seven major quality challenges for the ABS.

(i) The increasing use of large, but imperfect, administrative and transactional data bases for compiling official statistics.

(ii) Increasing user expectations are raising the quality "bar".

(iii) Managing the tension between improving business processes (which often means removing those responsible for statistical outputs from direct involvement with input processes) and maintaining or improving the quality of statistical outputs.

(iv) Quality assurance on electronic outputs.

(v) The presentation of statistics on the internet, including the need to educate the user community on quality of official statistics.

(vi) Managing the transfer of knowledge and skills with an ageing senior management team, many of whom will retire over the next 5 years.

(vii) Use of international statistical standards to maintain comparability where the standard may not be the most appropriate for national statistics.

4.1 **Increasing Use of Administrative/Transactional Data Bases**

We have used administrative data bases for many years (eg vital registrations for births and deaths, customs for trade data) to compile office statistics. Others have been used to develop frameworks for statistical collections. The issues at hand are the increasing availability of these data bases, their under-utilisation for statistical purposes, and taking advantage of the potential to link across data bases and ABS data sets using a common identifier (ie the Australian Business Number for business statistics).

Examples of administrative data bases that are becoming available are extended personal and business income tax data bases, health insurance transactions, and details of those on income support.

Transactional data bases are becoming available, although not in readily accessible form. Data bases of particular interest to the ABS are scanner data bases from retail outlets and eftpos data bases.

There are some particular advantages in using administrative or transactional data bases:

- they reduce the compliance cost we impose on respondents

- they are often "censuses" and therefore provide scope for producing detailed

data sets (eg by geography)

- they often have a longitudinal element (eg tax data) to support this form of analysis

- they often contain an identifier which facilitates analysis across data sets (eg the Australian Business Number will facilitate analysis across business tax data sets, customs data, and ABS surveys).

There are negatives of course - for example, the definitions may not be consistent with statistical concepts; less attention may have been given to incoming quality; and they may be out of date. Managing privacy aspects is a particularly important element. Although are motives are entirely honourable, matching data bases is a sensitive issue and ignored at our peril. Many of our users, particularly those in the academic community, are not as sensitive to these concerns.

There is also the question of whether the ABS should produce the statistical outputs or the agency responsible for the data sets. A number of issues come into consideration - the importance of the outputs to the national statistical service, costs, the extent to which quality can be managed and the basic question of whether the administrative agency is prepared to give up custodianships. Only the most important data sets will be brought into the ABS for compiling official statistics; for the others, we will work with the administrative agency to help them deliver statistical outputs into the public domain.

What have been our key responses to this important quality management issue?

- We are developing protocols for the publication and management of data from administrative sources. Associated with this is the promotion and support of good statistical and data management practices.

- For each statistical field, we are preparing information development plans which identify those areas of greatest importance and set out specific activities which will lead to increased availability of non-ABS data, particularly quality management issues.

- A major investment project has been the greater utilisation of taxation data to provide cost-effective statistics.

- We are investigating methods for assuring the quality of the very large but imperfect data sets that are available through administrative and transactional data holdings.

4.2 **Increasing User Expectations**

User expectations on quality are changing - they are much higher than what they were as recently as 5-10 years ago. This trend is likely to continue. The increasing globalisation of financial markets will mean that key macroeconomic statistics have international, as well as national prominence.

There is a perception that statistics have become more volatile. In some cases they have because the underlying phenomenon has become more volatile. However, we don't believe statistical measurement methods are a significant contributing factor - in most cases development methods have led to improvements although the perception may be different. For example, the volatility in the key national accounts series is considerably less than what it was 10-15 years ago yet this is quite different to the perception of many users.

We also receive more criticism of inaccuracies in very detailed data (eg Population Census tables) than previously. Again, it is not that the quality is deteriorating - it is the expectation is higher.

We have to accept that "the bar is rising" and do what we can to improve quality to the expected level. That is not always possible of course so managing expectations is important. This can be done by:

- providing good explanations of the strengths and weaknesses of particular data sets;

- talking to key users whenever possible about the strengths and weaknesses of data series;

- responding to their informed criticism (seek partnerships in improving quality eg in our detailed foreign trade statistics we openly seek feedback from users on the quality of the statistics); and

- providing as much explanation as possible for statistics that might seem unusual or different to expectation.

4.3 **Improving Business Processes**

Like several statistical organisations, the ABS is looking at how it might use new technologies, and other elements such as increased access to taxation data, to improve the efficiency of its business statistics processes.

We are also investigating the business processes associated with household surveys, particularly as increased use is made of CAI. However, in this section I will concentrate on business statistics to describe this particular quality challenge.

A team has been set up to look at the possibilities - referred to as strategic scenarios. We are looking for revised business processes that will be in place for at least 10 years and will yield a significant return on the investments required to set it up. It is likely that we will:

- extend the responsibilities of the Business Register Unit to capture and store taxation data with a direct link to the Business Register through the Australian Business Number (ABN). The ABN is now allocated through the taxation registration scheme and is available with most business transaction data bases (It is also collected in ABS surveys.).The data is stored in a way that it can be used by the various ABS statistical areas to compile statistics directly from taxation data or in combination with ABS survey data;

- set up a new business statistics processing environment based around an input data warehouse, with the Australian Business Number as the link across the various data sets;

- there will be greater centralisation of a number of the functions associated with compiling business statistics.

We can see the positives in these developments - more efficient delivery of business statistics, enhanced use of taxation data and other administrative data, data bases that support a wider range of statistical analysis. However, it will reduce the level of contact that statistical output areas have with their input data sources. What impact will that have on quality? What strategies can we deploy to mitigate the impact? These are important questions that we will have to answer before making a final decision on how we are to manage our business statistics in the future.

4.4 **Quality Assurance on Electronic Outputs**

Great care is taken on the quality of our paper products. This has been built on many years of experience. Our record is good and the quality assurance processes well embedded in the way we go about our business. Yet, more and more of user community receive their data in electronic form only. They will make analyses based on these outputs often leading to important decisions being made. It is just as embarrassing to us to have errors in electronic outputs as to have them in paper outputs.

Our quality assurance procedures for electronic outputs are not as sophisticated, but they are evolving. The key responses have been as follows:

- Our data warehouse supports the storage of all the objects associated with the dissemination with a particular set of statistics, including data cubes and meta data.

- Statistical areas are asked to approve each object - they are individually developing their own techniques for quality assurance (but sharing ideas on best practice).

- A publishing system is being developed which supports the simultaneous release of all outputs. If they are delivered from the same set of objects, there is less chance of inconsistency between the outputs.

4.5 The Presentation of Statistics on the Internet

Ultimately the user can only make judgements about the fitness of a statistical output for their purposes. These vary of course and what might be fit for one purpose may not be for another. There is an obligation on us to provide a range of supporting information on data outputs, including that on quality, so that they can make their own judgements on fitness of use. There are a number of existing, well proven practices relating to declarations about the quality of statistics. These activities are now a routine part of existing dissemination practices. They include:

- Concepts, Sources and Methods publications which describe in detail the methods used to compile major statistical outputs. These are available on our web site as well as on other media.

- An assortment of Information and Working Papers, and feature articles in publications, which are used to draw attention to issues specific to particular outputs or changes that are being made to their compilation methods.

- A policy of "no surprises" when there are significant changes to the methods used for the compilation of statistical series. As well as Information Papers etc, if there are important changes to statistical series, we embark on a program of seminars and bilateral discussions with key users to explain the changes and the reasons for their changes.

- Material on methods is included in all our publications. The ordering and physical presentation of this information is according to agreed standards. These were developed following research undertaken for us by a communications consultant on how our users use the material in statistical publications.

- The analysis section of our publications includes material which explains, among other things, large or unusual movements in our statistical series. Often this will be based on information that is only available to ABS staff through their contact with respondents or their intimate knowledge of the methods used in compiling statistics. Our User Groups have advised that this is one of the most valuable forms of analysis that we can undertake.

We believe that our key users have a reasonable understanding of the quality of the statistics they use. However the increased reliance on electronic dissemination poses new challenges. In one sense this move provides a wonderful opportunity to present a range of information on quality that is easily accessible through a few well designed "clicks". But because information about the quality of the statistics is "not in your face" like it can be in hard copy publications it is easier for users to avoid the key messages that your are trying to convey. The real challenge for us is to develop methods for presenting quality in a way that is not easy for users to avoid the main messages we want to convey.

One means of doing this may be to provide separate messages which draw their attention to particular information you want them to receive on quality. These could be automatically activated as particular statistical series are accessed or could be delivered by a separate email message. Research is required into the most effective means.

Lee and Allen (2001) have described some of our research work to date on this issue . The work is still at the exploratory stage. Things that are being investigated are:

- Useability testing of how users prefer to access information on quality.

- Showing leadership and developing of user education programs on how to use information on quality. A trial version of the course will be available in late 2001.

- The development of four prototype tools to assist users understand the quality of particular statistics. The four prototype tools are "Quality Issue Summaries", "Quality Measures", "Data Accuracy" and "Integrated Access to Data and Metadata".

I have not gone into detail on these developments as Bill Allen will be giving a separate presentation on this work later in the Conference.

4.6 **Managing the Transfer of Knowledge and Skills**

Like several other national statistical organisations, many of the ABS management team, and other senior staff, are aged in their 50's. Some have retired in recent years. Others are expected to over the next few years. If managed correctly, this is a great opportunity to refresh the organisation through providing new blood to management positions. These will normally be younger staff who will bring new ideas and energy into the management team.

On the other hand, experience and know-how will be lost. Both sides of this equation need to be managed carefully. Our strategy is as follows.

- We have developed special programs for those staff with potential. Specifically, they undertake a leadership and management development program which has been specially customised for the ABS. Staff are chosen for these programs by senior managers. You cannot self select yourself to be a participant in the

program. Furthermore, after staff have completed the program they can be expected to be chosen for a special assignment or rotated to a new position. The underlying philosophy is that the best way of learning is to obtain a variety of work experiences. A very high proportion of recent promotions to senior management positions have been participants in these programs. So far this has helped us to adequately cover the gaps created by a larger number of retirements than in the past.

- We retain links with retired ABS staff through a variety of informal and formal means (eg social functions, including them on the distribution list for ABS News, etc). Their knowledge is accessible if required.

- A stronger emphasis on knowledge management, using the facilities of our groupware product (Lotus Notes), means that key parts of our work are well documented and easily accessible.

- The strong move to standardisation of methods and systems means there is less dependence on local knowledge.

- For some key positions (eg Director of National Accounts) there is shadowing of work prior to the retirement of the incumbent.

To date we have managed this transition well. We have been able to adequately fill vacant senior positions and at the same time refresh the organisation by promoting staff with fresh ideas. There is a need to remain adroit.

4.7 Use of International Standards

Our starting position is at where international standards exist we should use them. This has not always been the case. For example, although our industrial classification has been loosely based on ISIC, and a concordance developed with ISIC, the classification is largely homegrown reflecting the specific interests of Australia and New Zealand. We have agreed to use the 2007 version of ISIC, at least to the 3 digit level, with variations only where there are specific circumstances

which justify it.

There are often pressures on us to divert from international standards. Sometimes this is to make the Australian situation look better (eg government finance statistics). In other cases, such as with the ILO unemployment definition, the pressure is because the international definition does not seem to reflect the real situation in Australian circumstances. We resist these pressures but it is important that we have a well documented international standard as a reference point to justify our position. Nevertheless, diversions from the international standard are made on an exception basis - they need to be well documented with a clear explanation of the reason. In cases where there is a clear need to have information on a basis other than the international standard our position is that we should publish statistics on both basis. The headline figure would still reflect international standard as increasingly the Australian situation is being compared with that of other countries and it is important that it is done on a comparable basis.

There is a tension that needs to be managed but if we are serious about the importance of international comparisons it is imperative that international standard is the main guiding light in developing the concepts, sources and methods used in Australia. There is also an obligation on us to make a significant contribution to the development and revision of international standards.

5. **Conclusion**

We would all agree that attention to quality is a fundamental aspect of our operation. In this paper, I have attempted to show that there are many dimensions to quality. This same message is clear from the frameworks for quality that have been developed by other organisations, such as the IMF, Statistics Canada and Statistics Sweden. The consequence is that a quality organisation depends on the actions of all it's staff as all can have an impact on quality in one way or another. It cannot be left to a project team designated as having responsibility for quality. Therefore, it can only happen if there is a genuine quality culture within the organisation. In this paper I have attempted to describe how we achieve this within the ABS. Nevertheless, it is important to have someone who performs the role of the corporate

conscience on quality. Among other things they draw attention to the most important risks to quality or behaviours they see as contrary to our corporate objective. We have given this responsibility to Methodology Division and made the chief part of the Executive so that it is easier for key messages to be conveyed.

6. Acknowledgments

Section 2 of this paper was based on a paper prepared by Frank Yu (ABS) for the Ninth Meeting of East Asian Statistical Offices.

7. References

Brackstone G (1999), Managing Data Quality in a Statistical Agency, Survey Methodology, Vol 25, no2, pp129-149

Carson C (2000), Towards a Framework for Assessing Data Quality, The Proceedings of the Statistical Quality Seminar, Jeju Korea, Korean National Statistical Office and International Monetary Fund.

Fellegi I (1996), Characteristics of an Effective Statistical System, International Statistical Review, Vol 64, pp165-197

Lee G and Allen B, (2001), Educated Use of Information about Quality, Bulletin of the International Statistical Institute, 53rd Session, Seoul, Korea.

Overview of ABS Quality Program
November 2008

One of the ways the Australian Bureau of Statistics (ABS) upholds its credibility in the community is through dedicated adherence to maintaining the quality of all aspects of work and at the same time looking for ways to improve current practices.

The ABS Making Quality Visible (MQV) project is a way in which the ABS has actively taken steps to formalise the process of maintaining quality. The Methodology and Data Management Division (MDMD) has a dedicated resource managing this project, along with the support of managers throughout the division who actively promote quality initiatives within their work programs and other areas in the ABS.

This paper provides a brief overview of the quality initiatives that the ABS is currently undertaking along with a status update of progress.

Quality initiatives related to 'Processes'

Quality Gates (QGs)

Quality gates (QGs) are a tool that are designed to improve the early detection and treatment of errors within a processing cycle. They are check points at significant stages within the cycle where an assessment is made of the quality before proceeding to the next stage of the cycle. The assessments can be either of a quantitative or qualitative nature. Each quality gate includes tolerance levels, roles and responsibilities of areas involved in regards to the quality gate, along with the action required at each gate.

There has been wide acceptance of the concept of quality gates throughout the ABS and a number of areas are beginning to think about their processes and how quality gates may fit. Progress has been slow in the implementation of quality gates throughout the ABS. Part of the reason behind this is the initial development of quality gates requires time to flesh out the placement of the gates and acceptable range of results expected from the gates. However, quality gates will provide efficiency gains in the long term by identifying issues earlier than otherwise would have occurred.

Quality Infrastructure System (QIS)

The Quality Infrastructure System (QIS) is a repository which links to other corporately approved systems to allow the capture and storage of survey information to aide in the quality assurance of the entire end to end process.

Monitoring of particular aspects of a survey, whether it be response rate, relative standard errors, data items etcetera, is available within the cycle, at the end of a cycle and with comparisons over time, through the use of the QIS. This ability enables the implementation of quality assurance processes, such as quality gates, to be monitored more easily.

The automation of the loading of information into the repository means that survey areas, once loaded, can obtain up-to-date information throughout their process by logging into the QIS.

The QIS has the ability to display survey information in various output forms, such as tabular and graphical. With the addition of Notes 8, future developments into widgets on desktops, such as dashboard type information for Senior Managers, will be investigated.

Work is being undertaken this financial year to improve already existing functionality in QIS in order to promote the System more widely throughout the ABS as a key quality assurance tool.

Clearance documentation

Clearance documentation is management information that is presented at the end of the processing cycle before the release of a statistical product. This information focusses on the quality of the data for the current cycle and any changes to that quality over time. Issues that occurred during the cycle are presented in this document. For example unusual movements compared to the previous cycle will be highlighted in this document along with an explanation for why these unanticipated movements are correct. Explanations may include real world impacts that have affected the data.

The ABS has a standard Clearance Documentation template for Economics collections and work is underway on a similar standard template for Social collections.

Quality of time series

The Time Series Analysis (TSA) area of the ABS has progressed the development of training courses. The first course developed - Understanding Time Series in the ABS - has been piloted in central office (CO) and the New South Wales (NSW) office.

A second national statistical training institute (NSTI) training module which focuses on the quality assurance of seasonally adjusted and trend estimates produced during processing cycles is under development. This module is more process-oriented and will focus on how to check and interpret the log file warning messages generated from running concurrent seasonal adjustment each cycle. Improvements to these log files have also been made, including an overall summary of the cycle, and will be released into production later this year.

Work is also progressing on developing recommended content for clearance documentation relating to seasonally adjusted and trend estimates. This work, along with the NSTI training course and log file warning messages, is designed to increase the capability of client areas to quality assure their processes and to increase their understanding of the time series they are analysing.

Internally, TSA has also implemented quality gates for the Annual Seasonal Reanalyses (ASR) processes. These gates aim to ensure that all aspects of an ASR are completed thoroughly and consistently across collections.

Quality reviews

Quality reviews are a new concept within the ABS that are being piloted in 2008. The concept of these reviews was endorsed by senior managers of the ABS on the 30 October 2007. The concept for quality reviews is heavily modelled on a similar program of quality reviews in Statistics Sweden, and on the Gateway Review process administered by Department of Finance and Administration (DoFA) for large IT projects.

Quality reviews are an intensive review of a collection or a process, and are conducted over a short period of time (usually five working days). The quality review is conducted by a team of three independent reviewers from methodological, system and operational backgrounds

and the findings of the review are provided to the area for which the review is being conducted, so that they can improve their processes where possible.

The first pilot review was conducted for the confidentiality procedures for International Trade statistics in January 2008. The second review was conducted in May on the Retail Trade survey. A third review will be scheduled shortly, most likely for a household collection or some other area of social statistics.

The pilot tests have revealed some issues to be addressed in the pilot quality review process, such as: clearer definition of the scope of the review; provision of a generic timetable for the reviewers to assist in the planning of the five days; the level of confidentiality that is appropriate for the final review reports; and a way of achieving a list of available panel reviewers, especially for emergency reviews.

It is intended that an evaluation report of the pilot quality reviews and recommendations will be presented to senior managers within a month of the third and final pilot review being completed.

Internal Audits

Internal audits are one of the ways the ABS ensures that our internal practices are efficient and correct. The ABS is currently conducting a top down review of our systems to identify areas of weaknesses. Areas of possible weakness identified from this review will undergo a more vigorous audit called a Quality Review (discussed above).

Quality initiatives related to 'Capability'

Quality training

MDMD is developing a suite of three NSTI courses to develop skills and knowledge relating to the making quality visible vision. These are Quality Concepts and Frameworks; Managing Statistical Risk; and Reporting Quality in the ABS. It is expected that these three courses will replace the current *Data Quality in the ABS* course.

Quality Concepts and Frameworks explains why the ABS has such a vested interest in maintaining quality and examines the different tools that have been developed to help with this direction. The course was piloted on the 31st of October 2008 in Central Office, with three attendees from Statistics New Zealand. Future presentations of this course will occur as demand dictates.

Managing Statistical Risk will discuss in detail tools which can be used by areas in the ABS to help with identifying statistical issues in the end to end process, and how to deal with any quality incidents once identified. Quality gates are one of the tools that will be discussed in detail along with documentation that is available for quality management. The course is scheduled for release in early 2009.

Reporting Quality in the ABS examines both internal and external data needs in the context of the seven dimensions of the ABS Data Quality Framework. The course primarily helps participants gain a better understanding of the creation of the content of quality declarations (discussed further below) which use the ABS Data Quality Framework. As quality initiatives evolve this course will be updated to reflect the current needs of staff.

Quality Manual

This manual is a central repository for all staff in the ABS. It contains information on a wide range of topics relating to management of the end to end cycle. It is an evolving document as more chapters will be added as developed.

The quality manual contains information on:

- the ABS Data Quality Framework;
- the Quality Declaration User Guide (a how to, and content document);
- Quality Gates;
- Quality Incident Response Plan;
- Statistics Sourced from Administrative Data; and
- A Template for quality assurance for change processes.

The ABS Data Quality Framework was developed from two international frameworks - Statistics Canada and Eurostat, to come up with the seven dimensions of data quality. These dimensions are Institutional Environment, Relevance, Timeliness, Accuracy, Coherence, Interpretability and Accessibility. These seven dimensions are used to help define the content of Quality Declarations (discussed below). Further information on the content of this framework can be found in the ABS Annual Report which is on the ABS website.

The Quality Gates section of the Quality Manual contains information on what a quality gate is, how it fits in the end to end process and how to develop them.

The Quality Incident Response Plan (QIRP) is a plan for resolving serious doubts about key statistical results. The QIRP approach is designed to provide a quick and rigorous high level response to an identified serious statistical problem and is based upon the formation of a dedicated ad hoc multidisciplinary team to investigate and resolve the issue. The process is pragmatic and focuses in quickly on the events and issues that really matter. This manual contains information on the practicalities of a QIRP, what to do and when, and recognising a QIRP in the first instance.

The Statistics Sourced from Administrative Data chapter looks at the management of the quality of statistical outputs produced from administrative data from the initial acquisition phase to final release of data.

The Template for quality assurance for change processes provides easy to follow instructions for monitoring any known change factors in a process. The template goes through the steps of identifying the known changes and then creating quality assurance procedures to manage these changes. This includes identifying risks, measures and people responsible for the monitoring.

Quality Assistant

The Quality Assistant is a database that provides links to useful information regarding quality initiatives in the ABS. It contains links to manuals, international papers and ABS specific papers. All staff are able to access the Quality Assistant from their desktops.

Quality Framework for Assessing Model-based Estimates

A quality framework for assessing model-based estimates was presented to the Methodology Advisory Committee (MAC) in June 2006. This framework contains the

following dimensions: Relevance, Accuracy, Plausibility, Consistency, Robustness, Interpretability, Transparency, Endogeneity, Sustainability and Cost Efficiency. Depending on the assumptions underlying the model the importance of each of these dimensions will vary when assessing the quality of model-based estimates. It was noted in the MAC paper that the criteria set out in the framework might not be exhaustive and that depending on specific circumstances other factors may need to be taken into account in the quality assessment. Further work on the development of this framework will continue based on advice received from the MAC.

Quality initiatives related to 'Outputs'

Quality Declarations (QDs)

Quality declarations are short, sharp statements describing the quality or fitness-for-purpose of a statistical product that are designed to operate effectively in a web-based environment. They are succinct and concise pieces of information designed to convey some of the quality aspects of a particular collection using the seven dimensions of the ABS data quality framework. The presentation of information is layered so that casual users obtain a brief headline indication of quality, whilst users who are interested in more detail can easily access it. The implementation of QDs was endorsed in May 2007 and the roll-out to began in 2008 to the ABS website.

The introduction of QDs to specific ABS products increases in 2009 to include compendiums and confidentialised unit record file (CURF) releases.

Quality Page

Discussion has been held recently on the idea of developing a 'Quality' theme page on the ABS website which would provide information to users about the quality initiatives that the ABS uses in order to provide high quality data.

Endeavours to release the ABS Data Quality Framework plus other informative documentation on quality assurance practices has started, with releases to the ABS website expected from early 2009.