

Innovations in Official Statistics

Australian Bureau of Statistics

Background paper presented to the
United National Statistic Division seminar
New York, Friday 20 February 2009

1. INNOVATIONS IN OFFICIAL STATISTICS

Introduction

This paper highlights innovations undertaken by the Australia Bureau of Statistics in the areas of data dissemination, and data collection and management. The drivers, the key motivations for the approaches taken, and the resulting actions put into practice for innovations in these areas are highlighted. The two source papers are provided as attachments for further reading and greater detail.

2. INNOVATION IN DATA DISSEMINATION

[See Attachment 1: Data Communication - Emerging International Trends and Practices of the Australian Bureau of Statistics *S M Tam and Regina Kraayenbrink* (2006)]

Summary

2. The main goal of statistical releases is communicating statistical stories and the associated metadata, in a manner that allows users to understand the statistical messages. In addition, the contexts, caveats and limitation of the statistical data must be effectively communicated to enable informed judgements about fitness for the intended use of the statistics.

3. Selecting the right statistical content that meets the information needs of users is one of the key factors of effective communication. In addition, how the statistician writes, structures, and presents the stories and data will largely determine how effective the communication will be. ABS research into cognitive psychology suggests that choosing layouts, presentational techniques and organisation cues to minimise users' cognitive load assists effective on-line communication of statistics.

4. Increasingly the ABS is utilising digital dissemination methods, with this paper focussing on dissemination via the Internet. However, some of the communication strategies apply to all means of dissemination.

5. Strategies being pursued by the ABS to achieve the communication vision include:

- adopting a layering approach to present the information from simple concepts to more complex information;
- implementing effective writing techniques that address the behaviours of the web audience;
- making the stories less dense, more visible and cross linked on the internet; and
- using contextual linking; not only will this minimise cognitive load of reading through large metadata documents, but will also ensure that the relevant metadata are provided at the same time with the statistical data when users are reading or accessing statistical information on-line.

Drivers

6. The common theme and challenges faced by National Statistical Offices (NSOs) today are focused on how to make the World Wide Web an effective medium for the on-line communication of statistics. Like many other NSOs, the Australian Bureau of Statistics (ABS) is positioning itself to use the internet as the principal channel for data communication.

7. There are a number of strategies in place to fulfil the ABS goal of efficient and effective web publishing. These strategies are:

- improving communication of statistics to facilitate user discovery of information, assessment of its fitness-for-purpose and understanding of the statistical messages in the data;
- broadcasting and proactive dissemination of information such as through Real Simple Syndication and Email Notification;
- improving self help;
- writing once and publishing many times to improve the efficiency and consistency of released information; and
- improving overall usability and accessibility of the web site.

8. Given its wide spread use and convenience of 24 hour/7 day access, ABS has been positioning itself to use the internet as the principal channel for dissemination. The ABS objectives for electronic dissemination are to:

- increase the users and uses of statistics for informed decision making;
- increase user understanding of the content, caveats, contexts and limitation of statistics, i.e. the fitness for purpose of the statistics; and
- improve cost effectiveness by maximising the opportunities available from the internet.

Actions

9. To complete the transition from a predominately paper-based statistical publishing regime to a web-based publishing regime, the ABS' web vision includes:

- a greater uptake of electronic dissemination by subject matter areas in a way that **improves communication**;
- providing more **self-help facilities** for users; and
- continuous improvements to the **usability and accessibility** of the ABS web site.

Improving Communication of Statistics

10. To achieve this, the aim is to provide more relevant, understandable and interesting content, and effective presentation of information on the web site by utilising opportunities available from web technology. Using the outcomes of cognitive research, the ABS has developed the following strategies for improving the communication of statistics on the ABS web site:

- using a **layered approach** for the presentation of information;
- developing basic guidelines for **presenting/writing for the web**;
- **contextual linking** of metadata with statistical data;
- using the concept of web magazines to ensure that **statistical stories** are visible to web surfers; and
- using **data visualisation techniques** to make statistical stories interesting, easy to read and to allow users to discover patterns in the statistical data.

11. The **layered approach** requires setting up information in layers, with the simplest information presented first, and the most complex information last, suitably hyperlinked to allow easy navigation from one layer to the next, and vice versa

12. For **presenting/writing for the web**, research suggests that many web site visitors consume information on-screen. This requires:

- presenting information in digestible but self contained chunks - ABS calls these "information nuggets" - with the nuggets suitably hyperlinked to allow interested users to pursue further detail;
- reducing the "density" of the information;
- reducing "propositional complexity" (i.e. the number of propositions in a sentence);
- presenting information in an easily scannable form;
- presenting materials on the web site so as to support visual perception; and
- writing content in a manner to support language comprehension. Information presented for on-screen reading should not be too dense, or verbose, and should also be amenable to scanning for key words by web users.

13. **Contextual linking** is important to ensuring that statistical data will be used in the right contexts and are fit for purpose. Facilities are being developed on the ABS web site to provide contextual links from data within publications, time series spreadsheets and data cubes to metadata. The web pages will, as required, contain links to more detailed data or metadata.

14. A (better) way to present information is to place important and interesting information first in clear visible positions on the web page e.g. the links to the key '**stories**' within the publication. This provides the required stimuli to draw readers' attention to these "stories" and entice them to read them by clicking the links.

15. If a picture is worth a thousand words, an animated and interactive picture could be worth even more. **Data visualisation** techniques include Sparklines (miniature sized graphs) and Dash Boards (multiple graphs presented on a single page). One application of the Dash Board currently being pursued by the ABS is to provide an alternative view of the Key National Indicators table (which includes key statistics about Australia's economy) into a dash board of graphical indicators.

16. **Email notification** and **Real Simple Syndication (RSS)** are current tools for "pushing" ABS data to subscribers. RSS is used to push Statistical Headline News whilst email notification is used to push hotlinks to publications. For the latter, the plan is to refine the tool to provide better options for users to tailor the email notification to their needs i.e. select targeted individual publications rather than targeted groups of publications. Other "push" technologies that may be available for effective broadcasting of ABS data are the opportunities available from Personal Digital Assistants (which increasingly have built-in wireless access to the internet) and podcasting.

17. **Podcasting** is an audio or video visualisation tool. Whilst the use of podcasting for statistical communication is yet to be widely explored, the National Agricultural statistics feature of the US Department of Agriculture is the first statistical organisation that provides podcasts of its broadcast reports. These podcasts are provided in audio or video content, rather than just text.

18. The **layered approach** is fundamental to the ABS broadcasting strategy. "Tourists" who have limited knowledge of the types of statistical information available from the ABS web site, can browse the Statistical Headline News to look for interesting leads that may entice them to read more. On the other hand, experienced users, "harvesters"/"miners", can bookmark the relevant web page, thereby bypassing the common navigation paths and reducing the number of clicks required. Note that expert users of a particular field of statistics may well be a "tourist" in another field.

Improving Self Help Facilities

19. Self help facilities need to be easy to use. Ideally, self help facilities should provide users with the flexibility to select ("slice and dice") and format tabular data on-line, before having to decide what to purchase/download.

20. The strategy for achieving this is to make more detailed, but confidentialised, data cubes available on the web site, supported by on-line data catalogues, metadata search and manipulation tools.

21. Other areas of self help to be pursued are:

- the provision of **spatial data services** (e.g. Web mapping services);
- the extension of the **self managed subscription system** to allow for ordering of print-on-demand publications; and
- the development of a **self help knowledge database**. With similarities to an on-line national information service, the self help knowledge database will provide advice on frequently asked, or encountered, questions relating to statistics.

22. The ABS has also aspired to provide **Web Services** - a facility to automatically provide up-to-date statistics to statistical/econometric models used by "harvesters". Two prototypes have been developed to illustrate the concept for a coding service and a table service.

Improving Usability and Accessibility

23. The User Centred Design Unit (UCDU) is a team within the ABS that are responsible for improving the overall design, usability and accessibility of the ABS web site. A Cognitive Laboratory is available on ABS premises for conducting usability tests. The ABS web site has been designed to meet Australian Government Accessibility Standards, and these standards are outlined in the World Wide Web Consortium –Web Accessibility Initiative –Web Content Accessibility Guidelines (WCAG).

24. By conducting regular usability and accessibility testing, the ABS is able to build relationships with our external clients, make informed design decisions and move closer towards achieving the communication objective.

3. INNOVATION IN DATA COLLECTION AND MANAGEMENT

[See Attachment 2: Reducing Respondent Burden – the Australian Bureau of Statistics Experience *Peter Harper et al(Economic Statistics Group)*]

Summary

25. The Australian Bureau of Statistics (ABS) has had considerable success in recent years in reducing respondent burden, while at the same time expanding significantly on the range of statistics available.

26. Most of these reductions have been achieved through the use of administrative data and through the use of a broad range of smarter statistical methodologies, including methodologies aimed at improving relationships with respondents. A number of internal organisational, methodological and technological changes have also reduced respondent burden across a broad spectrum of surveys or have increased the potential to do so in the near future.

27. The ABS will continue to place a high priority reducing respondent burden, and expects to make still further significant gains in this respect over the coming years as the full potential of recent initiatives are realised, with the goal being to never collect any data from a provider if it is available from another source or can be safely inferred from one.

Drivers

28. Changes in the legal and political landscape in Australia have provided some significant opportunities to reduce respondent burden. In most of these cases, the ABS has been very successful in influencing policy and legislation makers to take explicit account of national statistical opportunities and reduce respondent burden. These initiatives have addressed both actual and perceived burden, the latter often seen as the more strategic.

29. Much of the recent focus in Australia in reducing respondent burden has been on business surveys. This is consistent with government policy, the ABS Corporate Plan, and the recommendations of the Australian Government's 1996 Small Business Deregulation Task Force. The development and adoption of a business reporting standard similar to the model in operation in the Netherlands (the Dutch Taxonomy Project) is currently under development.

30. New technologies and methodologies will improve the degree of tailoring and customisation of contact that can be supported in dealing with providers and allow tailoring to the way providers prefer to communicate, working towards multiple communication channels. There are external drivers for expanding the ABS' capabilities in this area including meeting the ABS' obligations to a revised eGovernment strategy for Commonwealth agencies.

31. While the ABS does not compile estimates of the total load on households, the load imposed by each survey is carefully planned and monitored to keep it to the minimum necessary to produce quality statistics. As part of the 2006 Census of Population and Housing, the ABS offered Australians the opportunity to complete their Census form online, with nine percent of estimated dwellings submitting their Census forms via the eCensus.

Actions

32. The ABS actively collaborated with the Small Business Deregulation Task Force to develop a range of recommendations which would reduce respondent load significantly, including:

- setting a **target** of a 20% **reduction in the burden** imposed by ABS collections;
- establishment of a **Statistical Clearing House** where all statistical collections conducted by Commonwealth agencies would be subject to intensive review in terms of their need, relevance, methodology and respondent burden; and
- the creation of a **Business Surveys Charter** which set out the rights and obligations of businesses selected in ABS surveys. *The ABS also developed a Household Surveys Charter.*

Targeting the reduction of respondent burden

33. The ABS has far exceeded the targeted reduction in respondent burden. In part, this has been due to changes to the Australian taxation system with the ABS very well placed to take advantage of the opportunities for increasing **use of ATO administrative data** in business survey infrastructure and business survey operations. Specific administrative data now used include:

- Business Income Tax data (BIT);
- Business Activity Statement (BAS);
- Australian Business Register (ABR);
- Personal Income Administrative Data.

34. When considering the use of administrative data within statistical processes, the ABS has undertaken a comprehensive **risk assessment** for each of the individual strategies and developed contingency plans accordingly.

35. ABS internal organisational, methodological and technological changes include the **Business Statistics Innovation Program** (BSIP), a highly ambitious program begun in mid-2002 and involving the most significant organisational change in ABS business statistics areas for over thirty years. The fundamental shift was from subject matter based organisational units responsible for all phases of the statistical cycle within that field of subject matter to more functionally specialised units responsible for a particular phase(s) of the cycle across a number of subject matter areas.

36. An Economic Statistics Data Centre (ESDC) now has responsibility for nearly all interactions with providers of business data. It is responsible for **management of providers** in all business-based collections, all business based frames and all functions prior to input editing as well as being an advocate for providers, and providing quality statistical survey frames and collection management.

37. Developments of new technological and methodological facilities needed to perform these functions include:

- Total Approach Management (TAM) in respect of **relationships with providers**. The aim is improving standard correspondence with providers resulting in more cooperation and hence lower perceived respondent burden by providers of business data.
- Better **targeting of non-respondents** during the response chasing phases of collections, including centralising coordination of the handling of written complaints from business providers, improving turnaround and consistency, all of which are factors impacting on the perception of burden by individual providers.
- Provider Integration Management System (PIMS) for **tracking of all contacts** between the ABS and a business;
- State-of-the-art **call centre telephony** to better handle incoming calls from business providers. Many represent important opportunities for managing respondent burden interactively with the provider.
- Input Data Warehouse (IDW) from which all input data relating to businesses, including administrative data, is accessed, edited and analysed by ABS staff, to **more effectively exploit existing data**.
- ABS Survey Facilities, which are designed for **making greater use** of auxiliary information for estimation and imputation. This is a logical development flowing from the implementation of the IDW is to use increasingly sophisticated techniques for modelling, imputing and estimating the data needed from the data already collected, in preference to seeking more information from providers.
- Improved **tailoring and customisation** of contact in dealing with providers, so as to support multiple communication channels with providers for all surveys, using whichever communications medium is most appropriate to the task at hand.
- Use of consistent management information across surveys and to harness **economies of scale**.

Establishing a Statistical Clearing House

38. The Statistical Clearing House has now been operating for nine years and has reviewed 1,016 surveys (368 ABS, 648 non-ABS) collections. Of these collections, 35 were not approved, while many others have changed their methodologies as they have gone through the review process with the result often being significantly reduced respondent burden.

39. The ABS is now looking to expand the Clearing House to household survey collections.

Surveys Charter

40. An important element of the ABS Business Surveys Charter is the provision for small businesses to seek an exemption from a survey if they have responded to it for three or more years and their exclusion does not jeopardise the quality of results. Of the requests for exemption that have been rejected, the most common reason is the business in question is not small and therefore their exclusion will have a detrimental effect on the quality of results. Nearly all of the rejected applicants have accepted the reasons given and continue to be cooperative respondents.

41. The ABS has also developed a Household Surveys Charter. A significant proportion of social and population statistics is already compiled from administrative data sources however sections of the population where significant steps are being taken to reduce respondent burden include the Indigenous population. Concerns about maintaining privacy and ensuring appropriate use of the data seem more paramount than reducing respondent burden.

ATTACHMENT 1

Innovation in Data Dissemination

Data Communication - Emerging International Trends and Practices of the Australian Bureau of Statistics

S M Tam^a and Regina Kraayenbrink^{b,}*

^a *First Assistant Statistician, Information Management and Census Division, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, Australia. Email : Siu-Ming.Tam@abs.gov.au*

^b *User Centred Design Unit, Information Management and Census Division, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, Australia. Email : Regina.Kraayenbrink@abs.gov.au*

* Corresponding author. Tel.: +6 12 6252 7939; Fax: +6 12 6252 7102; Email: Regina.Kraayenbrink@abs.gov.au

Data Communication - Emerging International Trends and Practices of the Australian Bureau of Statistics

S M Tam^a and Regina Kraayenbrink^{b,*}

^a First Assistant Statistician, Information Management and Census Division, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, Australia. Email : Siu-Ming.Tam@abs.gov.au

^b User Centred Design Unit, Information Management and Census Division, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, Australia. Email : Regina.Kraayenbrink@abs.gov.au

Abstract

Results of a survey recently conducted by Statistics Canada suggest that many National Statistical Offices (NSOs) are in different stages of progress in migrating from a paper-based publishing regime to a web-based publishing regime. The common theme and challenges faced by NSOs today are focused on how to make the World Wide Web an effective medium for the on-line communication of statistics.

Like many other NSOs, the Australian Bureau of Statistics (ABS) is positioning itself to use the internet as the principal channel for data communication. There are a number of strategies in place to fulfil the ABS goal of efficient and effective web publishing. These strategies are:

- improving communication of statistics to facilitate user discovery of information, assessment of its fitness-for-purpose and understanding of the statistical messages in the data;
- broadcasting and proactive dissemination of information such as through Real Simple Syndication and Email Notification;
- improving self help;
- writing once and publishing many times to improve the efficiency and consistency of released information; and
- improving overall usability and accessibility of the web site.

In this paper, the strategies and techniques used by the ABS to improve statistical communication are discussed.

Keywords: Internet; web; publishing; dissemination; metadata; communication, cognitive methods.

1. Introduction

A survey was conducted by Statistics Canada [11] on the twenty two National Statistical Offices (NSOs), mainly from Europe and North America, that participated in the September 2005 International Conference on Marketing and Output Databases. The survey covered practices in data dissemination and a number of interesting trends have been identified. The more noteworthy trends are as follows:

- almost all NSOs are in different stages of migrating from a paper-based publishing regime to an electronic publishing regime;
- the World Wide Web is predominately used as the tool for electronic publishing;
- about two thirds of the responding NSOs have developed a strategic vision to guide the development of their web sites;
- the top five web site concerns nominated by NSOs were, in order of importance: development of web site content and related standards;

* Corresponding author. Tel.: +6 12 6252 7939; Fax: +6 12 6252 7102; Email: Regina.Kraayenbrink@abs.gov.au

implementation of information management and data presentation tools; on-line presentation of metadata; facilitation of electronic publishing and measuring satisfaction of web site visitors;

- the top five web site concerns raised by web site visitors were, in order of importance: ease of web site navigation; effectiveness of on-line search capabilities; availability of regional level data; documented statistical methods used in analyses and effectiveness of on-line information retrieval;
- over half of the NSOs predicted an increasing pace with the following developments in the coming years: streamlining and improving web site navigation; development of topic based web site structure; and improving the search capabilities of web sites;
- only about one third of the NSOs have developed publishing standards for web site information; and
- about 80% of NSOs have made the transition to providing all published information on the web site free of charge - about half of those who have done so had also reported rising servicing costs due to increased access to information.

Whilst the findings of the survey covered many different aspects of publishing, dissemination and communication of statistics, the common theme seems to be making the NSO web sites an effective medium for disseminating statistical information. No doubt the NSOs will conduct their own investigation and research into how to publish effectively on the internet, however, it is believed that a lot can be learnt from the newspaper industry. After all, the business of news companies is in disseminating news effectively using the mass media and many, if not all, have made a very successful transition from paper to web publishing. One reason for the success of news web sites is that, when they first used the internet, they did not simply reproduce a PDF of the newspaper on the web site. Rather, news stories were produced specifically for the web making use of the power and capability of the internet to hotlink and bundle related stories together.

Whilst some producers of statistics may think that their job is done by publishing the information on the web site, this paper argues that the job of a statistician is not done until the messages in the statistical release, together with the limitations of the information, are communicated to the users. Encouraging understanding of statistics will assist users to make informed judgements about fitness of purpose of the statistics for relevant decision making. The way statisticians write, structure and present the statistical stories, statistical data and metadata will largely determine how effective the information is communicated to the users.

In this paper, we discuss the strategies and techniques used by the ABS to improve the on-line communication of statistics [14].

2. ABS Directions for the Electronic Dissemination of Statistics

2.1. Current Situation

The ABS web site currently consists of over 365,000 web pages and more than 120,000 downloadable files. During 2005-06, nearly 80 million web pages were viewed and over 650,000 downloads of e-publications were made. The ABS is consistently ranked eighth (behind the Australian Taxation Office, Centrelink etc.) as the most frequently accessed Australian Government web site.

Since 2000, all ABS publications, spreadsheets, data cubes, research and information papers dating back to at least 1998, have been available from the ABS web site. The web site also includes other statistical support material e.g. Statistical Concepts Library, Directory of Statistical Sources and extensive school curriculum materials for teachers and students. Clients can subscribe to an email notification service which provides details of ABS daily releases in subject areas nominated by the client. Furthermore, the ABS has introduced a service, Real Simple Syndication (RSS), which allows subscribers to access ABS statistical headline news from their desk top. The ABS also provides on-line facilities for purchasing ABS products e.g. to order a paper publication.

To assist in meeting the ABS' community service obligations, since December 2005, access to all statistical information on the ABS web site has been made free of charge. For many years, the ABS has also run the Library Extension Program (LEP) to provide free access to selected statistical publications on paper. In the last couple of years, this service has been extended to make all electronic publications free to the general community via some 500 plus libraries throughout Australia. Users who do not have access to the Internet at their homes can access ABS statistics via the LEP, or by calling the National Information and Referral Service contact centre.

Unit record data from ABS household based collections are available to authorised users to conduct statistical research. The ABS has also taken strategic steps towards making this data more readily available to the tertiary education sector for teaching and research purposes. An agreement struck with the Australian Vice Chancellors Committee now provides every student, lecturer and university researcher with access, subject to approval to the Australian Statistician, to ABS Confidentialised Unit Record Files (CURFs).

Traditionally, the ABS only released CURFs via CD-ROMs. In 2004, the ABS launched its Remote Access Data Laboratory (RADL) whereby clients can (batch) submit programs via the internet to the ABS. These programs are automatically run against a CURF within the ABS environment, with the results, in most cases, automatically returned to the user. There are also checks in place to ensure no identifiable information is available in the results.

2.2. Aims of Electronic Dissemination

Given its wide spread use and convenience of 24 hour/7 day access, ABS has been positioning itself to use the internet as the principal channel for dissemination. The ABS objectives for electronic dissemination are to:

- increase the users and uses of statistics for informed decision making;
- increase user understanding of the content, caveats, contexts and limitation of statistics, i.e. the fitness for purpose of the statistics; and
- improve cost effectiveness by maximising the opportunities available from the internet.

A major challenge for the ABS is to complete the transition from a paper-based statistical publishing regime to a web-based publishing regime. The strategic directions to achieve the ABS' web vision include:

- a greater uptake of electronic dissemination by subject matter areas in a way that improves communication;

- continuous improvements to the usability and accessibility of the ABS web site, and
- providing more self-help facilities for users.

3. Strategic Directions

3.1. Improving Communication of Statistics

The ABS' goal is to increase use of statistics and to improve understanding of the content, caveats, contexts and limitations of the data. By improving communication, the ABS aims to assist user assessment of fitness for purpose of the data. To achieve this we are working to provide more relevant, understandable and interesting content. We are also working on effective presentation of information on the web site by utilising opportunities available from web technology.

The ABS has recently reviewed literature on cognitive psychology. Based on this research, principles and guidelines for content preparation and on-screen presentation [7] have been developed that address the three fundamental cognitive processes associated with comprehension - perception, attention and learning. For a more detailed discussion of the application of cognitive psychology to improve the on-line communication of statistics, see Appendix 1.

Using the outcomes of our research, we have developed the following strategies for improving the communication of statistics on the ABS web site:

- using a layered approach for the presentation of information;
- developing basic guidelines for presenting/writing for the web;
- contextual linking of metadata with statistical data;
- using the concept of web magazines to ensure that statistical stories are visible to web surfers; and
- using data visualisation techniques to make statistical stories interesting, easy to read and to allow users to discover patterns in the statistical data.

3.1.1. Adopting a layered approach

Our research on cognitive psychology suggests that using a layered approach is pivotal in reducing cognitive load, improving communication and supporting a diverse range of on-line users with different levels of statistical sophistication. The approach requires setting up information in layers, with the simplest information presented first, and the most complex information last, suitably hyperlinked to allow easy navigation from one layer to the next, and vice versa. See Fig. 1.

Through the rollout of the new design for the Statistics section of the ABS web site in January 2006, the web site has already been designed with the following layers of information:

- Statistical News Headline;
- Simple story of the data (Main Features);
- Detailed story (Detailed Publication) and Detailed statistics (Spreadsheets, Data Cubes and Time series); and
- Related data.

However, consistency in populating the layers across the diverse subject matter fields within the ABS still remains an area for development.

Fig. 1. Layering of Information



3.1.2. Presenting/Writing for the Web

Our research suggests that many web site visitors consume information on-screen. For effective communication of information, one of the important issues to consider is to ensure that the mind is able to handle the cognitive load. This requires, amongst other things, information to be presented in digestible but self contained chunks - ABS calls these "information nuggets" - with the nuggets suitably hyperlinked to allow interested users to pursue further detail; reduction in the "density" of the information, reduction of "propositional complexity" (i.e. the number of propositions in a sentence); and presenting information in an easily scannable form.

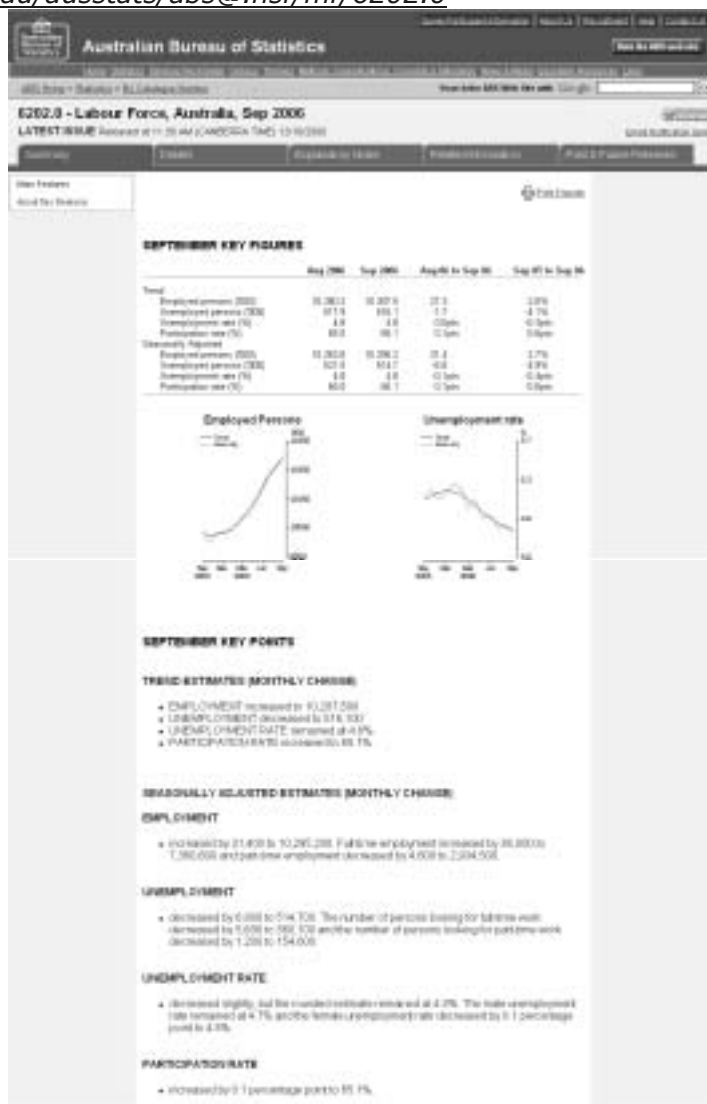
In addition, the presentation of the materials on the web site has to support visual perception (i.e. the ability to attach meaning to objects presented on the web site) and the content written in a manner to support language comprehension (i.e. the ability to connect to, and interpret, the written language). Information presented for

on-screen reading should not be too dense, or verbose, and should also be amenable to scanning for key words by web users. Our directions here are to develop good practice guides and training courses for presenting/writing for the web to achieve effective communication.

Fig. 2. Effective writing for the web

The summary content for ABS product 'Labour Force, Australia (6202.0)' presents a key figures table, simple graphs, dot points and short sentences. See:

<http://www.abs.gov.au/ausstats/abs@.nsf/mf/6202.0>



3.1.3. Contextual Linkage of Metadata

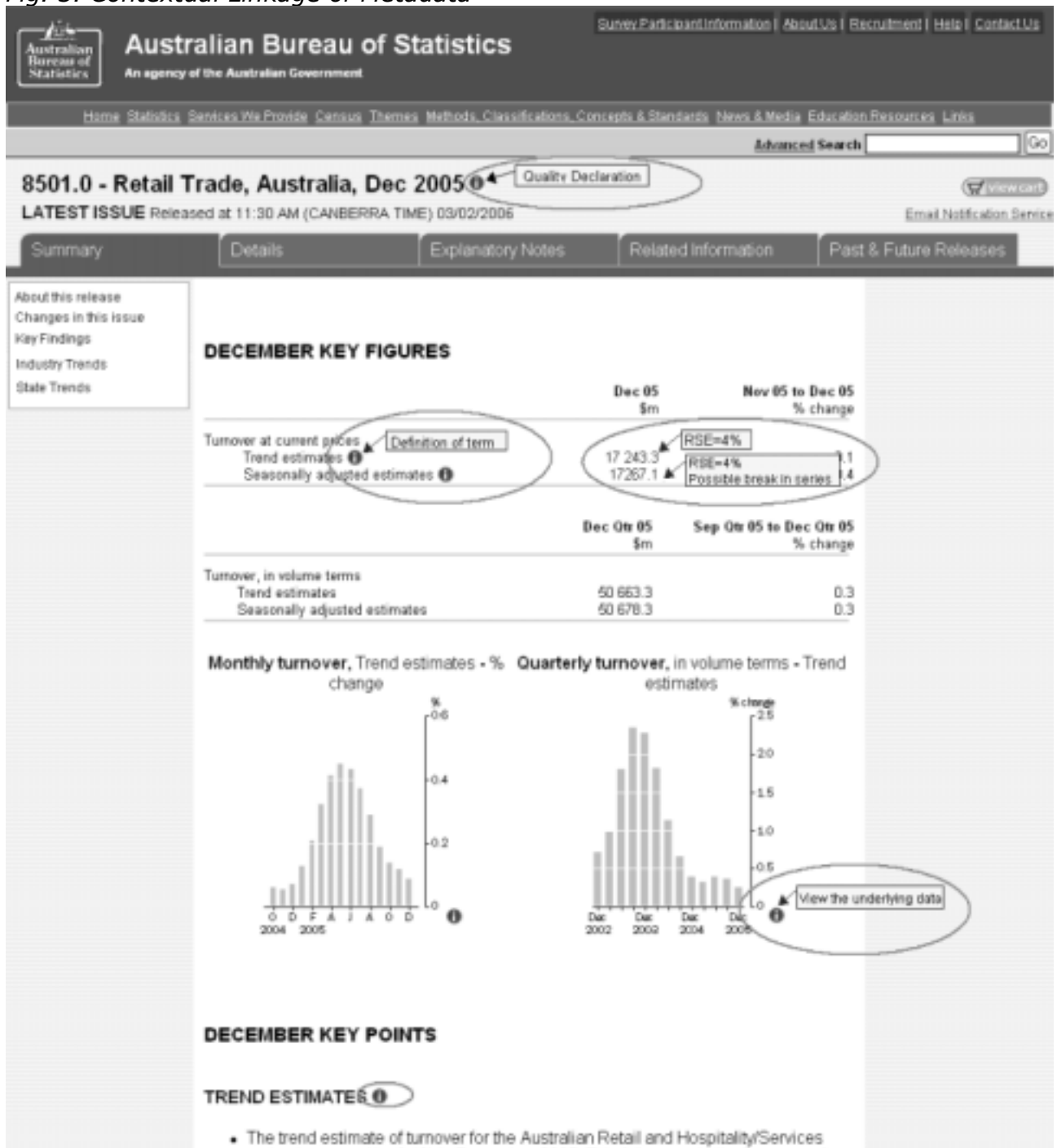
The current ABS approach for disseminating ABS metadata on the web site (e.g. Directory of Statistical Sources) is based on a "big bang" approach. We currently give users a lot of information and let them help themselves to the parts they need to know. Our cognitive research suggests that this is not an effective approach.

A more effective approach for communicating metadata is to adopt an approach similar to the one used for communicating statistical data - follow a layered approach

and tailor to the needs of the statistical user - whilst ensuring contextual linking between statistical data and metadata. In doing so we aim to minimise cognitive load and improve communication of statistical caveats. Contextual linking is important to ensuring that statistical data will be used in the right contexts and are fit for purpose.

We are developing facilities on the ABS web site to provide contextual links from data within publications, time series spreadsheets and data cubes to metadata. At the collection level, the web pages will provide information on the "Quality Declaration". At the statistical product level, there will be direct links from statistical terms to definitions, statistical classifications, questions used in the collection etc.; or links from statistical data to provide information about statistical errors i.e. quality statements. The web pages will, as required, contain links to more detailed data or metadata. An example of this is shown in Figure 3 below. A prototype for this has been built and is available at <http://www.abs.gov.au/about/ePublication>.

Fig. 3. Contextual Linkage of Metadata



3.1.4. Publishing Web Magazines

With the ability to hyperlink information on the web site, there are new and innovative ways for effective communication of ABS information. By merely reproducing our paper summary statistical releases electronically, there is a risk that users may omit reading important information, and experience a high cognitive load. This is because "context shrinks when reading on screen". Refer to Appendix 1 for further details.

An example to illustrate this point is provided by the Main Features of the publication 8501.0 Retail Trade, Australia (see Fig. 4). The Main Features step through "Key

Figures", "Key Points", "Notes" and then provide some further key points under the headings "Industry Trends" and "State Trends". With no upfront indication that Industry and State trends analyses follow, the "Notes" section provides a "presentational cue" to the readers that the statistical stories are coming to an end, and could result in the Industry and State trend analyses being missed.

Fig. 4. Potentially misleading presentational cue



A better way to present the information is to place important and interesting information first in clear visible positions on the web page e.g. the links to the key 'stories' within the publication. This provides the required stimuli to draw readers' attention to these "stories" and entice them to read them by clicking the links. An example of how this may be presented is provided in Fig. 5. Similar to the treatment by newspapers, the three stories are "bundled" together to provide an organisational cue that they come from the same statistical collection, and are conceptually related. In this case, the three "stories" present different dimensions of the analysis.

Fig. 5. Mock-up of multiple stories with headlines on the ABS homepage

The screenshot shows the Australian Bureau of Statistics website. The header includes the ABS logo and navigation links like 'Home', 'Statistics', 'Services We Provide', etc. The main content area is divided into sections: 'Product Releases' and 'Media Releases'. A red circle highlights a specific release titled 'Retail Trade continues strong growth - Retail Trade, Australia (cat. no. 85010)'. The release text states: 'Household good retailing continued its moderate hand growth' and 'NSW trend has been in decline for the last four months'. Other sections include 'Statistics by region' with a map of Australia, '2005 Year Book Australia Available Now', and 'ABS 2005 Year Book launched today'.

This method of presentation can be extended, of course, to not only summary publications, but detailed publications and compendium publications as well. An example of the first ABS example of the Web Magazine concept can be found in the SA Stats product (See: <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1345.4>).

3.1.5. Using Data Visualisation Techniques

If a picture is worth a thousand words, an animated and interactive picture could be worth even more. With the advent of the internet, animation and interactivity is increasingly being used to illustrate and tell interesting statistical stories. The animated population pyramid and thematic statistical maps, supported by application software developed by the Germany and UK Statistical Offices, are being used in many NSO web sites. See the animated population pyramid of Australia's age structure:

<http://www.abs.gov.au/websitedbs/d3310114.nsf/home/population+pyramid+previ>
[w](http://www.abs.gov.au/websitedbs/d3310114.nsf/home/population+pyramid+previ)

One particularly promising data visualisation tool for telling engaging statistical stories is GapMinder, the brain child of Professor Hans Rosling from Karolinska Institute. For further details, see: www.gapminder.org. The key to telling such stories is the ability to provide succinct statistical messages from the statistical data, freed from the qualifications on data that the statistician customarily provides in paper publications which would more effectively be contextually linked.

Other data visualisation techniques available include Sparklines (miniature sized graphs) and Dash

Boards (multiple graphs presented on a single page). For further details, see Edward Tufte's discussion about Sparklines and DMReview article titled: Data Visualization: Intelligent Dashboard Design (http://www.dmreview.com/article_sub.cfm?articleId=1035522). One application of the Dash Board currently being pursued by the ABS is to provide an alternative view of the Key National Indicators table (which includes key statistics about Australia's economy) into a dash board of graphical indicators.

Podcasting is an audio or video visualisation tool. Whilst the use of podcasting for statistical communication is yet to be widely explored, the National Agricultural statistics feature of the US Department of Agriculture is the first statistical organisation we are aware of that provides podcasts of its broadcast reports. See: <http://www.nass.usda.gov/Help/Podcasts/index.asp>). These podcasts are provided in audio or video content, rather than just text.

3.2. Broadcasting of ABS Data

Broadcasting in this context is defined as the proactive ("push") dissemination of information using the web site to suit a diverse range of user interests in a manner that facilitates communication. To do this effectively, we must ensure the information provided on the ABS web site is relevant to the diverse range of web users e.g. "visitors", "harvesters" and "miners". In many cases, reaching the potential audience will require some "pushing" of the relevant data/information.

Email notification and RSS are current tools for "pushing" ABS data to subscribers. RSS is used to push Statistical Headline News whilst email notification is used to push hotlinks to publications. For the latter, we plan to refine the tool to provide better options for users to tailor the email notification to their needs i.e. select targeted individual publications rather than targeted groups of publications. We are also keeping an eye on other "push" technologies that may be available for effective broadcasting of ABS data, such as the opportunities available from Personal Digital Assistants (which increasingly have built-in wireless access to the internet) and podcasting.

The layered approach is fundamental to the ABS broadcasting strategy. "Tourists" who have limited knowledge of the types of statistical information available from the ABS web site, can browse the Statistical Headline News to look for interesting leads that may entice them to read more. On the other hand, experienced users, "harvesters"/"miners", can bookmark the relevant web page, thereby bypassing the common navigation paths and reducing the number of clicks required. Note that expert users of a particular field of statistics may well be a "tourist" in another field.

3.3. Improving Self Help

Whilst cost recovered services like information consultancies will continue to be provided, our strategic direction for ABS services is to increase focus on the high value adding and complex information consultancy, leaving the simpler data extraction tasks to "self help". Self help facilities need to be easy to use. Ideally, self help facilities should provide users with the flexibility to select ("slice and dice") and format tabular data on-line, before having to decide what to purchase/download.

The strategy for achieving this is to make more detailed, but confidentialised, data cubes available on the web site, supported by on-line data catalogues, metadata search and manipulation tools.

Other areas of self help that we will be pursuing are:

- the provision of spatial data services (e.g. Web mapping services) and geography services (e.g. aggregating Mesh Blocks, i.e. spatial units that have a minimum size of 30 dwellings, to higher level Australian Standard Geographic Classification (ASGC) units, or non-ASGC boundaries);
- the extension of our self managed subscription system to allow for ordering of print-on-demand publications; and
- the development of a self help knowledge database. With similarities to an on-line national information service, the self help knowledge database will provide advice on frequently asked, or encountered, questions relating to statistics.

The ABS has aspired to provide Web Services - a facility to automatically provide up-to-date statistics to statistical/econometric models used by "harvesters". Two prototypes have been developed to illustrate the concept for a coding service and a table service.

3.4. Writing Once - Publishing Many Times

Writing once and publishing many times is an important direction for statistical publishing. It improves efficiency and ensures that statistics on the same socioeconomic and environmental phenomena published by the ABS are consistent. The latter is an important issue as increasingly the same statistics are published through different dissemination channels such as paper publications, CD-ROM, and web site, and in different formats such as Excel or ASCII. The key to mitigating against the risk of publishing inconsistent data is to use the same input source for outputting in different channels/formats.

Over the past 10 years, the ABS has developed and established good data management principles and practices. Supporting infrastructure, including the ABS Information Warehouse, provides the source of statistics, and the Publication Production Workbench supports the production of source documents in HTML, PDF and Excel formats. The work by an international group to develop an international standard protocol for exchanging data and converting data to multiple output formats i.e. Standard Data and Metadata Exchange (SDMX) for time series data, will provide another effective world wide tool for data exchange.

3.5. Improving Usability and Accessibility

The User Centred Design Unit (UCDU) is a team within the ABS that are responsible for improving the overall design, usability and accessibility of the ABS web site. Usability testing is conducted on a regular basis in order to ensure redeveloped web pages and new designs are meeting the needs of ABS' diverse range of users. The UCDU are working towards following an iterative design process whereby we encourage quick design phases, rapid usability testing, modification of the design based on testing and then repetition of these steps if required.

A Cognitive Laboratory is available on ABS premises for conducting usability tests. Usability testing generally involves observation of an individual usability participant performing information retrieval tasks. As part of the post usability session analysis,

key usability issues are identified, task completion times are analysed and other qualitative feedback is taken into consideration. Internal stakeholders of any redesign project are also invited to observe the usability sessions which has proved to be extremely successful in improving understanding of how users interact with the ABS web site as well as gaining stakeholder support for design decisions.

The ABS web site has been designed to meet Australian Government Accessibility Standards, which aim to ensure the web site can be accessed by anyone (including people with disabilities, people in rural or remote communities and people with older web browsers). These standards are outlined in the World Wide Web Consortium – Web Accessibility Initiative – Web Content Accessibility Guidelines (WCAG). See: [Web Content Accessibility Guidelines \(WCAG\)](#).

By conducting regular usability and accessibility testing, the ABS is able to build relationships with our external clients, make informed design decisions and move closer towards achieving the communication objective.

4. Conclusions

This paper argues that the main goal of statistical releases is communicating statistical stories and the associated metadata, in a manner that allows users to understand the statistical messages. In addition, the contexts, caveats and limitation of the statistical data must be effectively communicated to enable informed judgements about fitness for the intended use of the statistics.

Obviously selecting the right statistical content that meets the information needs of users is one of the key factors of effective communication. However, in addition, how the statistician writes, structures and presents the stories and data will largely determine how effective the communication will be. ABS research into cognitive psychology suggests that choosing layouts, presentational techniques and organisation cues to minimise users' cognitive load will assist effective on-line communication of statistics.

Strategies being pursued by the ABS to achieve the communication vision include: adopting a layering approach to present the information from simple concepts to more complex information; implementing effective writing techniques that address the behaviours of the web audience; and making the stories less dense, more visible and cross linked on the internet. Another very important strategy for communicating metadata is contextual linking. Not only will this minimise cognitive load of reading through large metadata documents, but will also ensure that the relevant metadata are provided at the same time with the statistical data when users are reading or accessing statistical information on-line.

5. Acknowledgments

We would like to thank our colleagues in the ABS and other National Statistical Offices, particularly David Pullinger from the Office of National Statistics (UK), for the numerous discussions on web publishing. Ideas from these discussions have led to the development of this paper. An earlier version of the paper was delivered at a one day workshop held in New York and hosted by the United Nations Statistics Division in March 2006.

6. References

- [1] R. Atkinson and R. Shiffrin, Human memory: A proposed system and its control processes, in: The Psychology of Learning and Motivation: Advances in Research and Theory, (Vol. 2). New York: Academic Press, 1968.
- [2] J. Best, Cognitive Psychology: Fourth Edition. West Publishing Company, US, 1995.
- [3] D.E. Broadbent, Perception and Communication, London: Pergamon Press, 1958.
- [4] D. Chang, L. Dooley and J. Tuovinen, Gestalt Theory in Visual Screen Design, A New Look at an Old Subject. Australian Computer Society Inc, Monash University, Australia, 2002.
- [5] S. Haviland and H. Clark, What's new? Acquiring new information as a process in comprehension, Journal of Verbal Learning and Verbal Behaviour 13 (1974), 512–521.
- [6] W. Kintsch and J. Keenan, Reading rate and retention as a function of the number of propositions in the base structure of sentences, Cognitive Psychology 5 (1973), 257–274.
- [7] R. Kraayenbrink, Applying Cognitive Psychology to Data Communication, Unpublished ABS Research Report, 2004.
- [8] Macquarie University, Macquarie Dictionary: Revised Third Edition. The Macquarie Library Pty Ltd, NSW.
- [9] G.A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information, Psychological Review 63 (1956), 81–97.
- [10] C. Reigeluth, Elaborating the elaboration theory, Educational Technology Research & Development 40(3) (1992), 80–86.
- [11] D. Roy, Consultation with National Statistical Offices: Web Presence Strategic Vision, International Marketing and Output Database Conference, The Hague, 2005.
- [12] H. Simon, Models of Man – Social and Rational, New York, John Wiley and Sons, 1957.
- [13] J. Sklar, Designing Web Sites. Massachusetts: Course Technology, 2001.
- [14] S.M. Tam, ABS Directions for Electronic Dissemination, Unpublished ABS Paper, 2005.
- [15] S.M. Tam, On-Line Communication of Statistics – Back to Basics, International Marketing and Output Database Conference, the Hague, 2005.

7. Appendix 1. Applying Cognitive Psychology to Data Communication

What is Data Communication?

Communication is defined, for the purpose of this paper, as the sharing of information between statisticians and users in a way that maximises understanding. As a general definition, communication is “the imparting or interchange of thoughts, opinions, or information by speech, writing or signs” [8].

Why Data Communication?

Principally by improving data communication, we aim to maximise use of information to fulfil the mission of national statistical offices; and to minimise misuse of statistics by ensuring the contexts, caveats, and other limitation of the information is understood.

How to Communicate?

To maximise communication, we will need to understand how the mind comprehends information. ABS research into cognitive psychology suggests there are three key cognitive processes involved in comprehension, namely perception, attention and learning.

Whilst the cognitive psychology theory is of general applicability, a number of issues need to be kept in mind for web communication, as follows:

Web surfers are users, not readers

Internet users tend to scan and skim read the material on-line to look for the information they are after. The term "satisficing" was coined by cognitive psychologists to describe the behaviour that people will stop looking once they come across something they think satisfy their needs [12]. Accordingly, we must design our statistical releases to aid on-line users of the information e.g. concise writing, meaningful headings and information in dot points to assist scanning.

Context shrinks when accessing information on screen

Unlike holding a hardcopy publication, web users can get lost within a web site and will be looking for presentational cues to tell them where they are within the web site (or web page). Presentational cues could be as misleading if used incorrectly e.g. metadata information between statistical stories could be misread as cues signalling the end of the publication, and content authors have to be mindful of these types of pitfall.

Cognitive load and overload

The mind can only handle limited "chunks" of information at one time and the literature suggests between five and nine chunks of information (seven plus minus two) where a chunk is any meaningful unit (Miller, 1956). If there is more information available than the mind can handle or the reader perceives certain information as irrelevant for their purposes, filtering occurs. To avoid the filtering out of important information, content needs to be presented with cognitive load in mind.

Cognitive Processes for Comprehension [7,15]

Perception

Perception is about attaching meaning to sensory information e.g. symbols, icons and information on the web site.

So how does the mind perceive? The cognitive psychology theory suggests that using presentational cues, the mind attempts to recognise the sensory information, by transforming it into an internal pattern and comparing it with other patterns stored in long term memory. If it matches an existing pattern, then meaning is attached. If it resembles a memorised pattern, then the meaning is guessed. If not learning takes place once the mind discovers the meaning of this new piece of information, sometimes through trial and error for skilled based learning.

For visual objects, how the materials are organised to assist the mind to carry out pattern recognition is the key to maximise perception. Strategies to deal with this include the use of Gestalt Laws (see Guideline 2.3 below) and minimisation of cognitive load. As an example, the Gestalt Law of Proximity suggests that

items/objects placed close together are perceived as being related conceptually. An application of this law will be to "bundle" related statistical stories together, as they do in electronic newspapers, to provide a visual relationship between the stories.

Additional consideration is required for perceiving texts. Clearly a non-native speaker of e.g. Chinese would not be able to understand a Chinese passage which underpins the importance of language familiarity. Providing the context for texts is also key to reducing ambiguity (eg the word "bank" can mean the edge of a river, or a financial institution). In addition, reducing the number of propositions (i.e. ideas within a sentence) in a sentence will also help to improve communication [6].

Furthermore, the presentation of information in layers, ranging from simple to complex concepts, is supported by the Elaboration Theory [10] and Given-New strategy [5] for layering out information.

Attention

Attention is about focusing the mind on a limited number of stimuli and in web design different techniques are used to focus users' minds on the key information on the screen.

As the capacity for sensory and working memory is limited, the Filter Theory [3] states that when there is cognitive overload, the mind will only allow some information to go through and block out the rest.

The issue for web designers is how to focus the mind of web users on the elements on the screen with which we want to draw attention to. A few strategies can be suggested:

- Minimise cognitive load e.g. ensure the information is not dense, and provided in 5-9 chunks of information; the information is easily scannable etc.;
- Ensure that the important information is clearly visible and organised effectively e.g. breaking up long stories into smaller stories and bundling them together using presentational cues suggested by the Gestalt Laws; and
- Use alerting techniques e.g. use rotating headlines, animation etc. to draw attention to important information. See the Statbox on the home page at <http://www.abs.gov.au> for an example of alerting technique.

Learning

Learning is about acquiring skills, e.g. to master the web site to find information, or acquiring knowledge, e.g. statistics on contemporary social/economic conditions.

Learning is achieved via encoding the information into long term memory. The information processing activity of comparing patterns recognised by the sensory organs with memorised patterns is important for the mind to determine what to be encoded. If it is already a memorised pattern, no new information is acquired and so no learning takes place. Otherwise, learning takes place once it has been worked out what the new piece of information means.

For acquiring skills to master a web site, publishing standards and consistency are the key. Surely one does not want to confuse the mind by continually changing the layouts, patterns and symbols on a web site. Even if there is a proven need to change

the web design, it should be done sparingly and record of the changes should be maintained.

Another useful strategy is to adopt international well recognised designs and signs on the web site, e.g. adopting a three column (Sklar) design for web pages[13], or the icon "i" for finding out more information.

When there is a change in the design of the web site or any of its associated elements, the mind will have to discover, more often than not, through trial and error, the new meaning, and through which new learning takes place. Of course, learning can be reinforced by practice and once the mind has mastered the skills, automatic processing of the information can take place with little or no cognitive load (Best, 1995).

For acquiring knowledge on statistics, we want the learning to be insightful i.e. a good understanding the contexts, caveats and the limitations behind the statistics for determining fitness of use. Contextual linking of metadata to statistical data is the key here, as otherwise there is an imposition on the mind to link different pieces of information together. The corresponding load imposed will generally be a deterrence to users except for those who are really determined to find out the story behind the statistics.

ABS Guidelines for the Development of Web-based Products and Services

Based on our cognitive psychology research to date, and learning from newspaper web publishing, ABS has formulated the following guidelines [7]:

1. Structure - Organisation of Content

Guideline 1.1. Grouping of information into meaningful units

- Group conceptually related items together
- Limit the number of items per group
- Limit the "chunks" of information to 5-9, where a chunk is any meaningful unit

Guideline 1.2. Present content so that users can orient themselves and comprehend new information on a page

- Provide informative page titles at the top of each page
- Use a breadcrumb trail on all pages
- Include an introduction or introductory section announcing the topic

Guideline 1.3. Use organisational cues to make text visually accessible and easily scannable

- Provide headings and subordinate headings
- Provide introductions
- Provide overview (preview) and topic sentences
- Provide lists or tables
- Provide explanatory link labels
- Provide site map
- Use dot points
- Avoid distracting readers with unnecessary cues

Guideline 1.4. Create order within and across grouped content

- Place known information before new information in sentences, paragraphs or pages
- Use deductive organisation - i.e. place important information near the top of paragraphs and pages

2. Presentation - Visual appearance of content

Guideline 2.1. Strive for consistency in everything

- Use the same relative positioning of repeated elements on different screens
- Use identical physical appearance of repeated elements

Guideline 2.2. Minimise the amount of information per page

- Provide a balance between positive and negative elements on the screen
- Draw attention to important elements by surrounding them with greater amounts of white space (e.g. headings)
- Use white space to differentiate between different groups of information
- Use short pages instead of long pages in general

Guideline 2.3. Encourage users to perceive the intended relationships between elements on the screen by using Gestalt organisation principles in presenting screen layout (Wertheimer, 1924 cited Chang, Dooley, and Tuovinen, 2002)

- Proximity : Put related elements close together on the screen. An easily perceivable space between unrelated elements is also important
- Similarity : Use similar look and feel for related elements (including text and icons), use distinct and easily perceivable differences for unrelated elements
- Figure ground : Ensure the figure is easily distinguishable from the background
- Simplicity : Minimise the amount of information presented on the screen. Avoid complexity
- Balance/Symmetry : Ensure the visual screen design appears visually balanced
- Focal Point : Ensure every visual presentation has a focal point, called the centre of interest or point of emphasis. This focal point catches the viewer's attention and persuades the viewer to follow the visual message further
- Isomorphic Correspondence : Recognise that people perceive and interpret the objects in different ways, based on individual experiences. Design based on common conventions
- Law of Unity/Harmony : Ensure there is harmony and unity between the visual elements on the screen

Guideline 2.4. Ensure that objects afford the actions that are intended to be performed on them

- Follow conventional usage, both in the choice of images and the allowable interactions
- Follow a coherent conceptual mode. Once part of the interface is learned, the same principles apply to other parts

Guideline 2.5. Use spatial and temporal cues to direct attention

- Position important elements on the prominent positions of the screen

- Use colour to draw attention to elements
- Use alerting techniques (such as animation and sound)

Guideline 2.6. Create colour harmony

- Avoid excessive use of colour (i.e. colour pollution)
- Exaggerate lightness differences between foreground and background colours.
- Use dark colours from the bottom half of the hue circle with light colours from the top half of the hue circle
- Avoid contrasting hues from adjacent sections from the hue circle, especially if colours do not contrast sharply in lightness

Guideline 2.7. Use typography to maximise readability and legibility

- Avoid mixing of font styles (i.e. No more than two font faces; no more than two font colour and no more than three font sizes - regular, large and small)
- Use a suitable font size
- Use sans serif fonts such as Arial, Helvetica or Verdana
- Use sufficient line, letter and paragraph spacing
- Follow common conventions (e.g. standard link colours for visited and unvisited links)
- Provide sufficient contrast between text and background colours

3. Content - Selection of content

Guideline 3.1. Select content that is relevant to the audience

- Choose information of interest to audience
- Avoid unnecessary provision of information
- Provide layers of information (from simple to complex) to suit audience types

Guideline 3.2. Use words that can be easily understood by the target audience

- Use concrete and meaningful words
- Use words that frequently appear in language
- Use short words (with fewer syllables)
- Use link labels that clearly explain the linked content
- Avoid jargon. When acronyms or abbreviations have been used, explain in the first instance and/or link to a 'glossary of terms' page

Guideline 3.3. Structure of sentences so that can be easily understood by the target audience

- Avoid too many links embedded within sentences
- Avoid ambiguities in language

Guideline 3.4. State ideas concisely

- Use concise wording
- Use short sentences
- Reduce proportional complexity within sentences (i.e. decrease the number of ideas conveyed within sentences)
- Omit unnecessary information

Siu-Ming Tam is the First Assistant Statistician in charge of the Information Management and Census Division at the Australian Bureau of Statistics. Siu-Ming's responsibilities include population censuses, information services, data management, statistical geography and library.

Siu-Ming has a PhD in statistics from the Australian National University. He is currently a member of the Steering Committee on Data Dissemination established by the United Nations Statistics Division, an elected member of the International Statistical Institute, an Accredited Statistician of the Statistical Society of Australia and a past Vice President of the International Association for Official Statistics.

Regina Kraayenbrink is a web designer and usability professional in the User Centred Design Unit in the Information Management and Census Division at the Australian Bureau of Statistics. Regina's main role is to improve the usability, design and accessibility of the ABS web site. Regina has a degree in Visual Communications at the University of South Australia and a Bachelor of Commence (Management) at Adelaide University.

ATTACHMENT 2

Innovation in Data Collection and Management

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE

**Paper prepared for APEX 2 meeting, Daejeon, Republic of
Korea, September 2006**

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE

Paper prepared for APEX 2 meeting, Daejeon, Republic of Korea, September 2006

INTRODUCTION

The Australian Bureau of Statistics (ABS) has had considerable success in recent years in reducing respondent burden, while at the same time expanding significantly on the range of statistics available. Most of these reductions have been achieved through the use of administrative data and through the use of a broad range of smarter statistical methodologies, including methodologies aimed at improving relationships with respondents. The ABS expects to make still further significant gains in this respect over the coming years as the full potential of recent initiatives becomes realised.

2 Much of the recent focus in Australia in reducing respondent burden has been on business surveys. This is consistent with government policy, the ABS Corporate Plan and the recommendations of the Australian Government's 1996 Small Business Deregulation Task Force. Since 1995-96, the total amount of time taken by businesses in providing data to the ABS has been substantially reduced. The average total load on all businesses for the past five years has been around 440,000 hours, or 33% lower than it was in 1995-96. For small businesses, the average total load over the past five years has been around 187,000 hours, or 42% lower than it was in 1995-96. Figure 1 shows provider load over the past ten years.

Figure 1: PROVIDER LOAD IMPOSED ON BUSINESSES BY THE ABS (hours '000)

some significant opportunities to reduce respondent burden. In most of these cases, the ABS has been very successful in influencing policy and legislation makers to take explicit account of national statistical opportunities and reduce respondent burden.

7 In 1996, the Commonwealth Government established the Small Business Deregulation Task Force to develop recommendations which would reduce the government reporting burden of businesses, small businesses in particular. The Task Force looked at all types of government 'red tape' and commissioned an independent survey which concluded that the burden imposed by statistical collections amounted to around one percent of the overall burden. Notwithstanding this small proportion, the ABS actively collaborated with the Task Force to develop a range of recommendations which would reduce this load significantly. The most significant of these recommendations included:

- setting a target of a 20% reduction in the burden imposed by ABS collections;
- establishment of a Statistical Clearing House where all statistical collections conducted by Commonwealth agencies would be subject to intensive review in terms of their need, relevance, methodology and respondent burden; and
- the creation of a Business Surveys Charter which set out the rights and obligations of businesses selected in ABS surveys.

8 The Government accepted all the recommendations the Task Force made in respect of statistical collections. As noted above, the ABS has far exceeded the targeted reduction in respondent burden. The Statistical Clearing House has now been operating for nine years and has reviewed 1,016 surveys (368 ABS, 648 non-ABS) collections. Of these collections, 35 were not approved, while many others have changed their methodologies as they have gone through the review process with the result often being significantly reduced respondent burden.

9 The Business Surveys Charter has been widely promulgated by the ABS. It is available from the ABS website and a copy of it (in brochure format) is sent to all businesses included in ABS surveys. An important element of the Charter is the provision for small businesses to seek an exemption from a survey if they have responded to it for three or more years and their exclusion does not jeopardise the quality of results. Of the requests for exemption that have been rejected, the most common reason is the business in question is not small and therefore their exclusion will have a detrimental effect on the quality of results. Nearly all of the rejected applicants have accepted the reasons given and continue to be cooperative respondents.

10 The ABS has also developed a Household Surveys Charter. As for the Business Surveys Charter, this is also available on the ABS website, and also in a range of languages.

11 In 2000, the Commonwealth Government introduced wide ranging changes to the

**REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE⁴
REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING
RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT
BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE
AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN
BUREAU OF STATISTICS EXPERIENCE**

Reducing Respondent Burden – The ABS Experience

September 2006

Australian taxation system. For many years before this, the ABS had been accessing business taxation unit record data maintained by the Australian Taxation Office (ATO) for a range of specific purposes. These purposes included maintenance of the ABS Business Register and economic statistics used in the compilation of the national accounts. These purposes were specifically allowed for in the main legislation under which the ATO operated, the *Income Tax Assessment Act 1964*. The ABS therefore had an established legislative and strategic relationship with the ATO when significant changes to the Australian taxation system were introduced in 2000. This relationship meant the ABS was very well placed to take advantage of the opportunities, presented by these changes, for increasing its use of ATO administrative data in business survey infrastructure and business survey operations.

12 The major elements of the tax system changes were the introduction of; a Goods and Services Tax (GST); an Australian Business Register (ABR); the allocation of an Australian Business Number (ABN) to registered businesses; consolidated Pay As You Go payments by businesses for various tax obligations and the introduction of an associated Business Activity Statement (BAS). Currently the ABS receives the following taxation data on a regular basis; monthly listings of the Australian Business Register (ABR); annual Business Income Tax data (BIT); and monthly Business Activity Statement data (BAS). The ABS currently utilises taxation data to some extent for the following purposes; benchmarking, supplementation, data substitution, register maintenance, imputation and improved stratification. All of these uses reduce respondent burden to varying degrees. Many of the examples given later in this paper relate to the ABS use of administrative data available from the ATO.

13 On 12 October 2005 the Prime Minister and the Treasurer announced the appointment of a Taskforce to identify practical options for alleviating the compliance burden on business from Government regulation. The Taskforce was chaired by Mr Gary Banks, Chairman of the Productivity Commission, and included business and small business representation. The Taskforce delivered its report 'Rethinking Regulation: Report of the Taskforce on Reducing Regulatory Burdens on Business' ('the Report') to the Government on 31 January 2006. The Report, guided by the views of stakeholders representing industry, small business, consumers and Australian Government, made 178 recommendations on actions to reduce red tape across a wide range of policy areas. The Report also made recommendations on actions to improve regulation making processes and regulatory gate keeping.

14 While the impact of statistical reporting was not a key component of the Report, a number of recommendations will have implications, potentially quite significant, for the ABS over the long term. The main one, Recommendation 6.3, calls for the development and adoption of a business reporting standard similar to the model in operation in the Netherlands (the Dutch Taxonomy Project). At this stage, the Government has formed a steering committee led by the Department of the Treasury to evaluate the costs and benefits associated with standard business reporting. The ABS is a member of the steering committee, and an associated working group, which will make a submission to Government late in 2006.

15 The pressure to reduce respondent burden of households is not as great in Australia as it has been recently for businesses. This is possibly because a significant proportion of social and population statistics is already compiled from administrative data sources. Rather the most significant concerns recently in these fields of statistics, expressed from respondents' perspective, relate to the increasing need for and potential of linking microdata across various sources. That is, concerns about maintaining privacy and ensuring appropriate use of the data seem more paramount than reducing respondent burden. There are, nevertheless, sections of the population where significant steps are being taken to reduce respondent burden, for example, for the Indigenous population.

STRATEGIC INITIATIVES

16 In addition to the externally driven opportunities mentioned above, the ABS has initiated a number of internal organisational, methodological and technological changes which have already reduced respondent burden across a broad spectrum of surveys or have increased the potential to do so in the near future. Most notable among these initiatives has been the Business Statistics Innovation Program (BSIP).

17 Implementation of BSIP began in mid-2002 and involved a transitional phase of three years. This was a highly ambitious program with many objectives including extending the ABS statistical leadership capabilities, improving data quality, improving the efficiency of business survey operations and increasing opportunities for ABS staff. It also included the objective of "improving provider relations, improved reporting mechanisms and reduced provider load". BSIP has provided a good platform for future changes. The following paragraphs concentrate on the achievements of BSIP in respect of improved provider management.

18 BSIP involved the most significant organisational change in ABS business statistics areas for over thirty years. The duties of nearly all of the 540 staff involved in business statistics were changed. While many of these changes were complex and required extensive consultation and communication on changed expectations, the fundamental shift was from subject matter based organisational units responsible for all phases of the statistical cycle within that field of subject matter to more functionally specialised units responsible for a particular phase(s) of the cycle across a number of subject matter areas. This saw the establishment of a new unit entitled the Economic Statistics Data Centre (ESDC) which has responsibility for nearly all interactions with providers of business data. Generally speaking, the ESDC is responsible for management of providers in all business-based collections, all business based frames and all functions prior to input editing. The term 'frame' used in this context applies to registers, common frames and survey frames. The mission of the Economic Statistics Data Centre is to be an advocate for providers, and provide quality statistical survey frames and collection management inputs involving:

- business register functions;
- large business profiling and key provider management;
- survey frames creation and maintenance, sample selection and design and data repair functions;

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE
*REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING
RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT
BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE
AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN
BUREAU OF STATISTICS EXPERIENCE*

*Reducing Respondent Burden – The ABS Experience
September 2006*

- administrative data acquisition;
- forms development, testing and production, including electronic forms;
- despatch and collection of all forms, including induction and basic reminders and complaints handling;
- intensive follow up of non-respondents; and
- Computer Assisted Telephone Interviewing (CATI) operations.

19 The new functionally specialised units which came into existence under BSIP have responsibility for leading the development of new technological and methodological facilities needed to perform their functions. They are very ably assisted in this respect by ABS technological and methodological experts who now have fewer clients to deal with for each stage of the statistical cycle. One of the new methodologies the ESDC has pursued is Total Approach Management (TAM) in respect of its relationships with providers. ABS methodologists have investigated what this approach might entail in the Australian context, taking account of the experiences of other national statistical agencies. This research has already led to the trial and successful implementation of improved standard correspondence with providers (covering letters, reminders, etc) resulting in more cooperation and hence lower perceived respondent burden by providers of business data.

20 Another methodological improvement which has improved relationships with providers has been the introduction of better targeting of non-respondents during the response chasing phases of collections. This helps to ensure that the pressure put on non-respondents to comply with ABS requests is more commensurate with their individual statistical significance, eg based on their relative contribution to estimates. Similarly, centralising coordination of the handling of written complaints from business providers has resulted in improved turnaround of these complaints and more consistency in their treatment, all of which are factors impacting on the perception of burden by individual providers.

21 These and other methodological improvements have been greatly facilitated by the introduction of new technologies under BSIP. For example, the Provider Integration Management System (PIMS) enables the tracking of all contacts between the ABS and a business provider, be it a form sent, a form returned, a telephone conversation or some written correspondence. Many business providers have commented that this is what they expect from an organisation such as the ABS where previously they had to tell different parts of the organisation the same piece of information (eg change of address) many times over. This was a source of irritation which significantly added to the perception of respondent burden.

22 The ABS has also recently installed state-of-the-art call centre telephony to better handle incoming calls from business providers. The ABS makes/receives over 300,000 phone calls a year from business providers. These calls often relate to queries the provider has about completing an ABS form and therefore represent important opportunities for managing respondent burden interactively with the provider. This new technology helps to ensure these opportunities are not missed, eg due to busy lines or inadequate staffing.

23 The ABS has recently developed an Input Data Warehouse (IDW) from which all input data relating to businesses, including administrative data, is accessed, edited and analysed by ABS staff. Improved access to all the data available for individual businesses has improved the analyses necessary to make methodological improvements, including extending the use of administrative data, which are likely to further reduce respondent burden in the future. A perennial theme for agencies attempting to reduce provider load is to more effectively exploit existing data sources. For the ABS, the two key opportunities are data already provided for another survey and data from administrative data sources. The ABS Input Data Warehouse, together with improved identifiers for business units such as the ABN, is a key enabler for us to achieve this in the next few years. Ultimately, our aim is to use it as the repository for all our business microdata, regardless of source. In the case of data directly collected from survey respondents, it will facilitate the sharing of data across collection boundaries, something we have had limited success with to date. In the case of administrative data, it will provide a single, efficient and safe repository for data that has been traditionally scattered across collection stove pipes in large, sometimes poorly managed, datasets.

24 It is worth noting that the ABS has had in place for many years now a sample rotation methodology for economic surveys utilising the ABS Business Register, which controls the number of continuous collection periods that a small and medium sized business can be selected in for a particular survey. Also, when surveys are initially set up or redesigned, the overlap of business between collections is also controlled to minimise the number of surveys in which small and medium sized businesses can be selected.

25 The ABS is also developing a new suite of facilities for estimation and imputation, called ABS Survey Facilities, which are designed for making greater use of auxiliary information. More specifically, these methodological facilities and the use of administrative data are likely to lead to lower respondent burden and higher quality estimates through a combination of:

- (a) using more up to date information from administrative sources for stratification purposes to improve sample designs;
- (b) using administrative data as auxiliary benchmarks in Generalised Regression Estimation;
- (c) using synthetic/indirect estimation methods based on models to generate statistical output for small domains;
- (d) using data collected over time to improve estimation efficiency;
- (e) improving sample efficiency by linking sample design to outcomes to be achieved rather than desired RSE's for key outputs;
- (f) improving sample efficiency through using administrative data to improve frame quality; and
- (g) improving efficiency of time series analysis in extracting seasonally adjusted and trend estimates.

26 The abovementioned new technologies and methodologies will progressively

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCES
REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING
RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT
BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE
AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN
BUREAU OF STATISTICS EXPERIENCE Reducing Respondent Burden – The ABS Experience

September 2006

improve the degree of tailoring and customisation of contact we can support in dealing with providers. Based on the data we collect in PIMS and IDW in particular, we should be able to build up an increasingly accurate picture of how to best handle segments of our provider population (eg small businesses in emerging industries, farmers, medium businesses) and tailor our approach accordingly. There already exists some distinction in our strategies for large, medium and small businesses, eg the role of Large Business Units in profiling large businesses. Moreover, especially for those providers we deal with relatively frequently, increasingly we should be able to tailor our approach to the needs of an individual business - eg when we contact the business, special reporting arrangements, what form of data collection to use, and so on.

27 While it has been possible to do some of this for years, under the pre-BSIP model, it has been difficult to get consistent management information across surveys and to harness economies of scale. This has been a major obstacle to introducing electronic data reporting, for example. Ultimately, we hope to be able to capture information about the way providers prefer to communicate with us - eg, phone, mail, fax, Internet, etc - as a by-product of our initial contact with them, and apply this consistently across survey boundaries.

28 The degree of tailoring that will be supported will almost certainly increase over time, as the benefits to the ABS and providers grow. However, there are many examples which lead us to proceed cautiously. For example, experience in Australia across organisations that have deployed major online data collection systems (eg the tax office, banks, administrative agencies) shows that there are significant up front and ongoing costs for each extra channel of communication that needs to be provided, limiting the opportunities to make gains if take up rates are small.

29 Ultimately we would like to be able to support multiple communication channels with providers for all surveys, using whichever communications medium is most appropriate to the task at hand. Traditionally, we have relied most heavily on paper based communication, followed by telephone, fax, electronic, face to face and then a variety of other methods. Although this is hardly a new issue, most of our gains in optimising our communication channels have yet to be made. While all channels are used to some extent now, no single collection can readily support all of them. Issues such as provider preferences, security, support costs and vehicle bias often militate against applying a comprehensive solution.

30 The time is now right to pick up the pace towards multiple communication channels in the ABS for a number of reasons. Not only has the creation of functionally specialised units for business statistics collections as part of BSIP provided the ABS with a centralised capability to ensure that a consistent approach is adopted across all business collections, but we have also been able to observe and learn from the efforts of other statistical agencies and we can leverage off their experiences. Further, there is a steady increase in demand from businesses for EDR and we expect that this will pick up following the 2006 Australian Census of Population and Housing's use of an electronic collection instrument (eCensus) (see below).

31 There are also a range of external drivers for expanding the ABS' capabilities in this area including meeting the ABS' obligations to a revised eGovernment strategy for Commonwealth agencies. The urgency for ABS action on this front has intensified following the Prime Minister's April 2006 Rethinking Regulation Taskforce (mentioned in paragraphs 13 and 14 above). The whole of government initiatives will complement proposed early work in the ABS. It may also result in some changes to our long term strategy, as success with some whole of government infrastructure would present new opportunities to progress electronic data reporting beyond existing proposals or at least progress them earlier than first planned.

32 For the first time, as part of the 2006 Census of Population and Housing, the ABS offered Australians the opportunity to complete their Census form online. Due to the high-volume of traffic expected (including an expected significant load spike on Census night) and the short-term nature of the project, development of this system, known as the eCensus, was contracted to IBM Australia. At the time of writing over 760,000 Australian households, or nine percent of estimated dwellings, had submitted their Census forms via the eCensus. This includes approximately 40 percent of total responses, or 313,000 submissions, received between 6pm and midnight on Census night. No significant technical or security issues occurred during the six week period that the eCensus had been operational up to the end of August 2006.

33 A logical development flowing from the implementation of the IDW is to use increasingly sophisticated techniques for modelling, imputing and estimating the data we need from the data we have, in preference to seeking more information from providers. Better approaches to dealing with partial non-response, late returns, imputation from administrative data, etc, are examples. The Statistics Netherlands notion of a "virtual census" captures much of what we would hope to achieve in this regard, with the goal being that we never collect any data from a provider if it is available from another source or can be safely inferred from one.

SOME SPECIFIC EXAMPLES OF THE USE OF ADMINISTRATIVE DATA

Use of BIT data

34 BIT data have been used extensively for many years to improve the quality of the National Accounts and the results of industry surveys which provide source data for the National Accounts. Most notably, the annual Economic Activity Survey (EAS) draws heavily on BIT data to significantly supplement its directly collected sample resulting in more reliable and fine level industry estimates for Input Output purposes. Summary results, including a detailed description of the methodology used to combine BIT and directly collected data are contained in the ABS publication *Australian Industry* (cat. no. 8155.0).

35 Recent changes the secrecy provisions of the legislation enabling the ABS to make use of business taxation data allow the use of the data for general statistical purposes covered by the Census and Statistics Act 1905. Previously the use of BIT data was restricted primarily to uses related to the compilation of the National Accounts. These

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE¹⁰
REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING
RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT
BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE
AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN
BUREAU OF STATISTICS EXPERIENCE Reducing Respondent Burden – The ABS Experience

September 2006

changes have enabled the ABS to make much greater use of the very rich BIT unit record dataset than had previously been permitted. Some of the applications of BIT data now being considered include:

- reducing provider load in business surveys that are not data sources for the National Accounts;
- building the Business Longitudinal Database;
- compiling small area statistics for smaller businesses;
- providing alternative views of industry based statistics; and
- providing other fine-level cross-classifications of business-related data which is not feasible from ABS sample surveys.

36 Our uses of BIT data have had to take account of some of the shortcomings of that data. For example, BIT data does not allow multi-location businesses to be identified or disaggregated and this implies that many businesses will be coded to their main state of operation only. In addition, the BIT data file generally excludes non-profit institutions, non-trading trusts and funds (including superannuation funds, deposit funds and pooled superannuation trusts). The BIT file for a particular reference year also includes business taxation returns lodged for the financial year rather than a complete picture of all business operating. Because of these shortcomings, the ABS is continuing to investigate ways of improving the quality of estimates from the BIT data source.

Use of BAS data

37 BAS data have been used primarily to improve the efficiency of ABS surveys (most recently to improve stratification and estimation benchmarks for the Monthly Retail Business Survey); we plan to expand use in this regard as we learn more about the data and improve on its quality. For example, the ABS is currently investigating the most appropriate use of BAS data in the Quarterly Business Indicators Survey, one of our flagship quarterly collections. The expansion of the use of BAS data will not only reduce reporting demands on businesses, but should enable ABS to provide a range of new statistical information in future, including expanded state data. This is likely to improve quality and reduce provider load for existing collections, rather than provide a data source in its own right for users to access. Like BIT data, there are some fundamental constraints regarding the comprehensiveness and timeliness of BAS data which make it impractical to use as a primary source for direct statistical output. However, these constraints should not preclude its use in more sophisticated ways (as described in para 25).

Use of ABR data

38 The ABS published counts of businesses from the ABR in respect of October 2000, June 2001 and June 2004. It is not possible to accurately reflect the states in which multi-state businesses operate. However, an estimate of the number of locations is derived using the information provided at the time of ABR registration. This estimate will under-estimate the total number of locations belonging to multi-state businesses,

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE¹¹
REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE

BUREAU OF STATISTICS EXPERIENCE Reducing Respondent Burden – The ABS Experience

September 2006

as only one location is assumed in each state in which each such business operates. The ABS may, in future, consider developing an alternative methodology for identifying and allocating as appropriate, the operations of multi-state businesses to relevant states.

39 A further limitation of the ABR data is that the address information provided is not necessarily the location address of the business. In many cases it is the address of an accountant/tax agent or an alternative business address as nominated by the person registering the business. This complicates the derivation of estimates below State/territory level. Estimates of the number of businesses by postcode are available in the current release but, in addition to the problem with the ABR address information described above, postcode can be a very coarse indicator of location. Business counts classified by ANZSIC Division are currently available at the postcode level through cat. no. 8136.0.55.001. At the state level, business counts classified by ANZSIC Class are available through cat. no. 8138.0.55.001. The next release of business counts in cat. no. 8165.0.55.001 will incorporate counts of business entries and exits by selected classifications.

40 The first experimental output of Business Entries and Exits (BEES) was released in June 2005 as 8160.0.55.001 - Experimental Estimates, Entries and Exits of Business Entities, Australia, 2001-02 to 2003-04. This data was produced using ABR registration data received by the ABS on a monthly basis. The initial publication contains numbers of business entries and exits by state and by ANZSIC Division, for the years 2001-02, 2002-03 and 2003-04. Since that release, further developments have occurred in both the conceptual model and data source used for the statistics. Updated statistics will be released in December 2006 in Counts of Australian Businesses, including Entries and Exits (cat no 8165.0).

41 Work is underway on the development of a Business Longitudinal Database (BLD) using the ABR as the population frame, and a combination of administrative and ABS survey data. The first production version of the BLD will be available in December 2006, containing a combination of characteristics and activity information sourced from an ABS survey and financial data from administrative datasets for a panel of 3,000 small and medium businesses. An integrated business characteristics collection vehicle is currently being developed to both populate the BLD dataset and provide data for estimates currently produced in the Innovation and Business Use of Information Technology surveys. The longitudinal component of the survey will see each business kept in sample for 5 years, with a new panel of 3,000 businesses commencing annually. This initiative will assist in linking data within the ABS and blending both activity and financial information to ensure providers are not unnecessarily burdened with duplicative requests for data.

Personal Income Administrative Data

42 All the above examples relate to the use of administrative data about businesses. The ABS also makes extensive use of administrative data relating to individuals in an effort to reduce the burden that would otherwise be imposed. An important distinction

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE¹²
REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING
RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT
BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE
AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN
BUREAU OF STATISTICS EXPERIENCE Reducing Respondent Burden – The ABS Experience

September 2006

between this personal data and the business data referred to above is that the personal data is generally only available from these agencies in aggregated form, because of privacy considerations, whereas the business data referred to above is made available to the ABS as unit record (microdata) files. In Australia, there is far more sensitivity attached to the sharing of personal data than there is for business data, but the Australian Treasury has just announced a wide ranging review of taxation secrecy and disclosure provisions designed to simplify and standardise the various tax law secrecy provisions. Changes emanating from this review will hopefully allow ABS greater access to personal income tax data for statistical purposes.

SUMMARY OF ISSUES ASSOCIATED WITH USING ADMINISTRATIVE DATA

43 The brief descriptions in the previous section of the ABS use of some administrative datasets illustrate some of the major issues encountered. This section elaborates on these and outlines the broad strategies the ABS is employing to address them.

44 When considering the use of administrative data within statistical processes, the ABS has undertaken a comprehensive risk assessment for each of the individual strategies and developed contingency plans accordingly. For example, one of the key mitigation strategies for the use of taxation data within ABS is the maintenance of a strong working relationship between the two agencies. Over a considerable period of time, the ABS and ATO have expended considerable effort in developing and maintaining a close working relationship. This is supported by a number of forums ranging from senior management level through to a series of working groups attended by operational staff. This close working relationship has enabled the ABS the opportunity to better understand the relevant data items and even to influence the range of data items collected by the ATO and their supporting quality maintenance strategies. The statistical impacts of any revised methodologies introduced as the result of use of administrative data have been monitored and managed through a process of closely measuring the impacts, backcasting/bridging affected series, bridging of series and conducting parallel survey estimation. Strategies implemented by the ABS will continue to be reviewed and assessed on an ongoing basis in response to the continued volatility of data quality, and the potential for changes to priorities and processes within the ATO.

45 When administrative data sets are obtained from the ABS by users there is often a perception that the ABS has assumed an ownership role, including a responsibility to ensure the quality of the data. Some users assume the ABS closely monitors administrative data quality before it is disseminated by using detailed output editing techniques, despite the caveats we publish regarding quality of the dataset. The ABS does issue caveats to users which explain all data quality issues associated with any data set, but this often does not overcome the perception that the data is "pure" just because the ABS has delivered it.

46 The ABS is committed to providing users of our data with sufficient documentation to make appropriate use of our outputs. In addition, the ABS has developed Memoranda of Understanding (MOU) with agencies which provide data to the ABS from administrative processes. These MOUs often articulate long and short term directions for improving the quality of the data.

47 The ABS is committed to fully utilising administrative data sets to supplement and improve the quality of existing surveys. Where we achieve this, we can deliver improvements to the level of detail available in ABS publications or reductions in the load placed on our providers.

48 ABS policy requires that metadata about data which have been sourced from another agency and published by the ABS should be documented, and that the documentation should be released to the users of the data. This information is generally available through the technical notes at the back of publications and via the ABS web site.

49 For example, overseas arrivals and departures (OAD) data are derived from information available to the Department of Immigration and Multicultural and Indigenous Affairs (DIMIA). These data are published by the ABS in *Short-term Visitor Arrival Estimates, Australia* (cat. no. 3401.0.55.001) and *Overseas Arrivals and Departures, Australia* (cat. no. 3401). Comprehensive information on the quality and source of these statistics is set out in the respective Notes and Explanatory Notes. This information includes updates on issues affecting data quality, changes and improvements to the data, scope, estimation methods, corrections and imputations, and seasonal adjustment and trend processes.

50 The ABS is committed to improving the coherence of this documentation and is currently working on several projects to develop frameworks to assess quality of administrative datasets through all stages of the statistical process. However, implementation of these frameworks will be gradual, introduced on a case by case basis.

51 Finally, the ABS is currently taking the lead in developing a framework for describing data for use in the National Data Network (NDN). The NDN is a cross government initiative to facilitate acquiring, sharing and integrating of data (for further information, please see <<http://www.nationaldatanetwork.org>>). The framework for describing data will further provide greater consistency in the documentation of quality both in government statistics.

CONCLUSION

52 Over the past decade or so, the ABS has taken many steps to reduce respondent burden, reflecting opportunities both internally and externally driven. These initiatives have addressed both actual and perceived burden, the latter often seen as the more strategically important. The ABS will continue to place a high priority reducing

REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE¹⁴
**REDUCING RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING
RESPONDENT BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT
BURDEN - THE AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE
AUSTRALIAN BUREAU OF STATISTICS EXPERIENCE REDUCING RESPONDENT BURDEN - THE AUSTRALIAN
BUREAU OF STATISTICS EXPERIENCE**Reducing Respondent Burden – The ABS Experience

respondent burden, by use of smarter methodologies and increased use of administrative data, notwithstanding the issues and challenges these changes present. The investment the ABS has made in these strategies to date has already paid substantial dividends and this is expected to be the case into the future.

Economic Statistics Group
Australian Bureau of Statistics
September 2006