Measuring compliance effectiveness

Evaluating effectiveness
OUR COMMITMENT TO YOU
We are committed to providing you with accurate, consistent and clear information to help you understand your rights and entitlements and meet your obligations.

If you follow our information in this publication and it turns out to be incorrect, or it is misleading and you make a mistake as a result, we must still apply the law correctly. If that means you owe us money, we must ask you to pay it but we will not charge you a penalty. Also, if you acted reasonably and in good faith we will not charge you interest.

If you make an honest mistake in trying to follow our information in this publication and you owe us money as a result, we will not charge you a penalty. However, we will ask you to pay the money, and we may also charge you interest. If correcting the mistake means we owe you money, we will pay it to you. We will also pay you any interest you are entitled to.

If you feel that this publication does not fully cover your circumstances, or you are unsure how it applies to you, you can seek further assistance from us.

We regularly revise our publications to take account of any changes to the law, so make sure that you have the latest information. If you are unsure, you can check for more recent information on our website at www.ato.gov.au or contact us.

This publication was current at June 2012.
## CONTENTS

### INTRODUCTION
- Understanding compliance effectiveness 2
- Evaluating compliance effectiveness 3

### DEVELOPING YOUR SUITE OF INDICATORS
- Indicators 7
- Identifying potential indicators 7
- Validating your indicators 9
- Measuring your indicators 13
- Identifying your data requirements 16

### EVALUATING EFFECTIVENESS
- Gathering your data 19
- Analysing your indicators 19
- Determining your effectiveness 21

### REPORTING ON EFFECTIVENESS
- The executive summary 25
- The detailed report 26

### APPENDIXES
- Appendix 1 28
- Data collection methods and sources 28
- Appendix 2 31
- References 31
- Appendix 3 32
- Glossary 32

### MORE INFORMATION
inside back cover
INTRODUCTION

Evaluating the extent to which our compliance strategies have resulted in positive and sustainable changes in compliance behaviour and community confidence is critical to knowing whether or not we have been effective in achieving the ATO intent as outlined in our Strategic Statement.

We have developed a compliance effectiveness framework and methodology which helps us measure the impact of our compliance strategies on voluntary compliance and community confidence. This guide is part of a suite of publications designed to help people understand the framework and apply the methodology, as well as develop effective compliance strategies.

This booklet provides guidance to risk managers and evaluators in applying phases 3 and 4 of the compliance effectiveness methodology. It should be read in conjunction with Measuring compliance effectiveness – Applying our methodology (NAT 72342).

This booklet focuses on:
- developing a suite of indicators
- evaluating effectiveness
- reporting on effectiveness.

For more detailed information about the different measurement methods and analysis techniques and an extensive range of external reference material, refer to Literature review – Measuring compliance effectiveness 2007 (NAT 71078).

UNDERSTANDING COMPLIANCE EFFECTIVENESS

Compliance effectiveness is the extent to which the actual outcomes of our compliance strategies align with our desired outcomes. Did our strategies have the effect or impact on compliance behaviour and/or community confidence that we wanted them to?

The compliance effectiveness methodology is the tool we use to plan the right mix of strategies, measure our effectiveness and identify opportunities for continuous improvement. We use it for both planning and evaluation purposes. What we learn about our effectiveness informs our future planning and helps re-shape our compliance improvement strategies.

There are four iterative phases in our methodology (figure 1):
- **Phase 1** is about understanding and articulating the risk and making sure it aligns with the ATO intent.
- **Phase 2** is about clearly expressing our desired outcomes and what would be different if we are successful. It also involves identifying the right mix of treatment strategies which treat the drivers of compliance behaviour and not just the behaviour.
- **Phase 3** involves identifying potential indicators and validating them to ensure they are viable and useful. The suite of indicators should be capable of providing a defensible picture of our effectiveness.
- **Phase 4** deals with evaluating and reporting on the extent to which we have been effective in changing compliance behaviour, building community confidence, or both, over the immediate, intermediate and long term.

Phases 1 and 2 focus on planning to be effective while phases 3 and 4 focus on evaluating the extent to which we have been effective.
INTRODUCTION

When should an evaluation occur?
Identifying when an evaluation should occur is based on when you expect to see an effect.
In deciding when to undertake an evaluation consider the following questions:
- When do you expect to start seeing the change?
- When do you think the full effects will be seen?
- How often will you need to do an evaluation to be able to understand whether the change has been sustained?

EVALUATING COMPLIANCE EFFECTIVENESS

When you evaluate compliance effectiveness you assess the extent to which your compliance strategies have resulted in positive and sustainable changes in compliance behaviour with one or more of the key compliance obligations (registration, lodgment, reporting and payment) and community confidence.

The methodology uses indicators to provide insights into the effectiveness of our strategies because it is not always possible, practical or cost effective to use more definitive measures.

A suite of indicators is used to identify whether or not the success goals, and ultimately the desired outcomes, have been met. This is done because no single indicator will provide sufficient information to enable you to assess the overall effectiveness of your compliance strategies in achieving your desired outcomes.

When evaluating compliance effectiveness you need to answer the following key questions:
- Is there a change in compliance behaviour and/or community confidence?
- If there is a change, did we cause it (attribution)?

To understand this, you need to explore behaviour at:
- the broad or macro level to identify whether there are any observable changes
- the strategy or micro level to determine whether those changes can be attributed to our compliance strategies.
FIGURE 1: Compliance effectiveness methodology

**Phase 1**
Articulate risk  
Align with ATO intent

- How does the ATO intent translate into the context you’re working in?
- What are the behaviours and drivers of the risk?
- What is the compliance risk to achieving the intent?
- Who’s involved in the risk?

**Phase 2**
Define outcomes  
Develop strategies

- What outcomes are you seeking to achieve by addressing the risk?
- What strategies will you use to deliver these outcomes?
- How do you define success in terms of more specific goals?
- Who are your target groups?

**CHECKPOINTS**
You should ensure that:
- the risk aligns to the ATO intent
- the risk adequately reflects the behaviours and drivers
- the risk is refined, concise and not open to interpretation.

**CHECKPOINTS**
You should ensure that:
- the desired outcomes align to the ATO intent and the risk you have identified
- the overall desired outcomes are adequately covered by your specific (success) goals
- the strategies target participants of the risk and drivers of the behaviour
- the strategies address any unintended consequences.
**INTRODUCTION**

**Phase 3**
**Design indicators**

- Which indicators are feasible?
- What are the potential indicators?
- Which of these indicators will paint a defensible picture?
- What is the description of the indicators and their purpose?

**Phase 4**
**Validate indicators**
**Determine extent of effectiveness**

- How will the evaluation be conducted?
- What are the indicators telling us?
- What data will be required and where will we get it?
- What’s our baseline?

**CHECKPOINTS**

You should ensure that:
- each success goal can be measured and has a corresponding indicator (or suite of indicators)
- potential indicators are aligned with the success goals, desired outcomes and the ATO intent
- indicators are viable and provide a balanced picture of performance
- both qualitative and quantitative information is used.

You should ensure that:
- data exists, is available, or can be acquired at a reasonable cost
- we can explain how effective our compliance strategies have been – that is, the extent to which actual outcomes align with desired outcomes.
The answers to the questions posed in phases 1 and 2 of the methodology should have been agreed with the risk owner and endorsed by the relevant risk management committee (or equivalent decision making body) before developing your indicators.
Before you can determine whether your strategies have had the desired effect you must:

- identify potential indicators
- validate those indicators
- determine a suite of indicators that will help you paint a defensible picture
- decide how you will measure your indicators.

You need to undertake these activities soon after the risk statement, desired outcomes and strategies have been endorsed by your risk management committee (RMC) or equivalent. Working through these steps will allow you to put the processes and tools in place to ensure that the relevant data can be collected and analysed.

There is usually a considerable time lag between setting up and actually doing the evaluation because the strategies will take time to have an effect.

If you are evaluating effectiveness retrospectively you will be in a position to undertake these activities almost immediately after completing phases 1 and 2 of the methodology.

Bring together people who have the relevant business knowledge and those familiar with the business data. Consider involving:

- the risk manager
- intelligence experts
- strategy experts
- risk analysts
- business and data analysts.

Identifying the indicators
Identify appropriate indicators by looking at each success goal and thinking about the changes you would expect to see if the strategies were effective.

Think about how those changes could be measured. Ask yourself these questions:

- How would you know if there was an improvement in voluntary compliance?
- What would tell you if there was a change in community confidence?
- How did you know there was a problem in the first place?

Example of a potential indicator

<table>
<thead>
<tr>
<th>Success goal</th>
<th>Potential indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustained improvement in voluntary compliance with reporting obligations.</td>
<td>Change in the effective tax rate.</td>
</tr>
</tbody>
</table>

At this stage, don’t be too concerned with how difficult it might be to measure the indicator. Avoid choosing only indicators that are easy to measure. You may find that even though an indicator can’t be measured at the moment it may be useful at a later date when more data becomes available.

Use a combination of indicators that focus on both the broad and the detailed level to help you to identify whether there has been an observable change in compliance behaviour and community confidence. This can also help you to understand whether your compliance strategies are responsible for the change.

When identifying your indicators consider whether they are capable of showing:

- **sustainability** – identify indicators that can demonstrate your progress and whether a positive change is maintained or declines over time. Select indicators that can show change over the immediate, intermediate and long terms.
### Documenting the discussion

Document the discussion about each indicator including:

- what the indicator is supposed to do
- the expected result if the strategies are successful
- the external factors that may affect the usefulness of the indicator.

This will help to ensure that what the indicator actually measures is in line with the original intent. It also ensures that these considerations are available to others who may work with the indicators at a later date.

---

### ALIGNMENT

When identifying potential indicators make sure there is a clear link between the indicators and your success goals.

---

### Characteristics of indicators

When developing your potential indicators make sure they are:

- **indirect or flow on effects** – include indicators that can show whether the strategies have had an effect on the wider population beyond the part of the population directly targeted by your strategies.
- **unintended consequences** – consider indicators that will show whether any unintended consequences have emerged as a result of your strategies.

Identifying potential indicators often results in a list of 15 to 25 indicators. While this number would be impractical and costly to track, the discussion between people from diverse backgrounds and capabilities is invaluable in identifying indicators that might otherwise be overlooked.

### Table 1: Aligning indicators to success goals

<table>
<thead>
<tr>
<th>Success goals</th>
<th>Potential indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>The community is less tolerant of non-compliant behaviour</td>
<td>Change in the:</td>
</tr>
<tr>
<td></td>
<td>number of reports received on the Tax Evasion Referral Centre</td>
</tr>
<tr>
<td></td>
<td>responses to questions relating to non-compliant behaviour in our community, business</td>
</tr>
<tr>
<td></td>
<td>and tax practitioner perceptions surveys</td>
</tr>
<tr>
<td></td>
<td>tone of media comment relating to the non-compliant behaviour.</td>
</tr>
<tr>
<td>Taxpayers are more compliant with their lodgment obligations</td>
<td>Change in the:</td>
</tr>
<tr>
<td></td>
<td>percentage of the population that lodge their tax returns on time</td>
</tr>
<tr>
<td></td>
<td>percentage of taxpayers that lodge their activity statements by the due date.</td>
</tr>
</tbody>
</table>

---

### Aligning indicators to success goals

The purpose of evaluation is to understand the extent to which the success goals and ultimately the desired outcomes have been achieved. Consequently you must ensure that your indicators are directly linked to what you are trying to achieve.

Table 1 provides examples of aligning indicators to success goals. It also shows how the indicators can work together to help understand the extent to which the success goals have been achieved.

---

### Characteristics of indicators (cont.)

When developing your potential indicators make sure they are:

- **able to be measured** – indicators should be based on a combination of quantitative and qualitative data.
  - Quantitative data is information that can be counted. For example, 78% of individual taxpayers use a tax agent to lodge their income tax return.
  - Qualitative data is descriptive data that is difficult to measure in numeric terms. For example, the taxpayer reported the reason his claim was incorrect was because he did not understand his obligations.
  - Quantitative data can tell you how much of something has happened while qualitative data is often extremely useful in explaining why it happened. This may also help you learn more about your risk.
- **expressed in a neutral form** – the purpose of an indicator is to identify change. It should make no assumptions about the direction the change should take.
  - For example, express the indicator as a change in the ratio of work related expenses to salary and wages rather than as a reduction in the ratio of work related expenses to salary and wages.
**VALIDATING YOUR INDICATORS**

Once you have identified the potential suite of indicators you need to explore each indicator to understand whether it will provide the information that you are expecting. This process is called validation.

Validate your indicators by applying a range of tests including those in the SMART test:

- **Specific**
- **Measurable**
- **Achievable**
- **Relevant**
- **Timed**.

This test explores whether your indicators are specific, measurable, achievable, relevant and timed.

You should also explore whether:

- your indicators can identify whether a change can be attributed to your strategies
- the data can be compared over time
- the results are able to be substantiated
- your indicators are capable of presenting an unbiased picture of the extent to which your success goals have been achieved.

**The smart test**

- **Specific**
  Indicators should have a clear, unambiguous description so that everybody has a common understanding of the information the indicator is expected to deliver.

  Express the indicators in a way that will help you to identify whether compliance behaviour has changed as a result of your strategies. A simple way to do this is to describe the indicator as ‘a change or a difference in the rate or ratio or number of …’

**EXAMPLE**

A change in the percentage of activity statements lodged by the due date by sole traders with a turnover of less than $1 million.

Comparing the compliance behaviour before you implemented your strategies to the behaviour afterwards allows you to understand whether there is a change in behaviour that correlates with your strategies.

- **Measurable**
  You need to look at how you are going to measure changes from one period to the next. Consider the following questions when deciding which measurement method to use:
  - Can you express the indicator in a numeric form?
  - Will you compare the data against some form of benchmark?
  - Is a calculation such as a ratio involved? What are the components of the calculation?
  - Will you measure the whole population of interest or restrict your analysis to a sample?

**EXAMPLE**

A change in average rental income.

Calculated by dividing total dollar value of rental income by total number of taxpayers reporting rental income.

For more information on measurement methods and data collection, see pages 12–17.

- **Achievable**
  Indicators should be both practical and reasonable.
  You need to make sure that the data for each indicator is readily available or can be obtained at a reasonable cost.

**EXAMPLE**

A change in net tax payable to total income ratio.

This indicator is based on data that is readily available from our Data Warehouse.
MEASURING COMPLIANCE EFFECTIVENESS

01 DEVELOPING YOUR SUITE OF INDICATORS

- Relevant
  Indicators should be able to measure the expected or desired changes. They should also have a direct link with your success goals and desired outcomes.
  Don’t just choose an indicator simply because it’s easy to measure. Make sure it contributes to the evidence you’re gathering to help you understand the extent to which your success goals have been achieved.

- Example
  Change in the number of individual income tax returns lodged compared to the number of people identified in the Australian Bureau of Statistics’ (ABS) Australian labour force statistics.
  These statistics are generally presented as data which spans a calendar year rather than an income year. There is a notable time lag between collection and presentation.
  When comparing this ABS data against ATO data, ensure the correct months are represented in the data set for the income year rather than calendar year and that they correspond to the year in which lodgment behaviour is being assessed.

  Additional validation tests
  When validating your indicators you should also make sure that they are:
  - attributable – indicators need to measure something that your strategies can reasonably be expected to influence.
  - you need to think about whether a change in behaviour is the result of your strategies or whether that change was influenced by some other factors.
  - bear in mind that:
    - there is a time lag before the data becomes available
    - the data is available on a regular or infrequent basis
    - the data is available for all of the relevant time period.
  
  Where other factors are likely to have an impact on the usability of the indicator, record both the issue and how you expect it to affect the results.

  Other factors that could cause a change in behaviour include:
  - other ATO strategies
  - court decisions on tax matters
  - changes in tax rates
  - economic fluctuations
  - globalisation
  - natural disasters
  - media coverage of tax related issues
  - unemployment
  - taxpayer literacy and numeracy levels.

- Example
  Success goal
  Individual taxpayers voluntarily comply with their income tax lodgment obligations.

  Indicator
  A change in the percentage of individual returns lodged by the due date.
  An increase in this indicator will provide evidence of improved voluntary compliance with lodgment obligations.

  You also need to understand what the indicator should show in the immediate and longer term. For example, a successful strategy may show a significant improvement in the short term that levels off in the future – an immediate effect with a sustained change.

  Timed
  You need to understand whether your strategies affect voluntary compliance and community confidence over time. Select indicators that will identify change and show your progress over the immediate, intermediate and long term. Indicators also need to be based on data that can be produced regularly enough to track progress and quickly enough for it to be useful, with only a short time between the period the data covers and when it becomes available. Consider whether:
  - there is a time lag before the data becomes available
  - the data is available on a regular or infrequent basis
  - the data is available for all of the relevant time period.

  There is often a time lag between the collection and availability of external data. You need to carefully examine your data sources to understand how timing issues will affect the usability of your indicator.
EXAMPLE: Weak and strong indicator

Success goal
The community displays a reduced tolerance for participation in the cash economy.

Weak indicator
Change in the frequency and tone of media comment relating to the cash economy.

This indicator is weak because it is likely to be too broad to identify any real behavioural change.

Measuring something more specific can often strengthen the indicator.

Strong indicator
Change in responses to specific questions in our surveys (including community, business and tax practitioner surveys).

Accepting or rejecting indicators
The validation process enables you to make informed decisions as to whether an indicator should be:

- accepted – if you plan to accept an indicator that has validation issues consider the impact these flaws will have on the strength of the indicator. Also consider the impact of these flaws on its value in the overall suite of indicators.
- rejected – if you plan to reject an indicator consider the impact this will have on the usefulness of the overall suite of indicators. If its absence would have a detrimental effect then consider whether there are any alternatives to fill the gap.
- retained for future use – generally, indicators that don’t meet all of the validation requirements should be discarded. But if the flaw is only temporary, such as where a time-based indicator hasn’t been in existence long enough to be able to track progress, then it should be retained for consideration in any future evaluations.

Painting a defensible picture
After you have completed the validation process you need to determine whether your suite of indicators will provide a defensible picture.

Ask yourself whether, overall, the suite of indicators is likely to provide:

- sufficient evidence to justify your conclusions regarding the extent of your effectiveness
- a picture that is strong enough to withstand scrutiny and challenge.

For more information about external factors, refer to pages 46–48 of our Literature Review – Measuring compliance effectiveness (NAT 71078).

comparable – continuity is necessary in order to make comparisons in trend analysis over time. Data can be subject to continual revision. For example, new data may be received over time that may have an impact on your original benchmark population. While it is not always possible to avoid differences in the underlying data, recognition of those differences and the influence they have on the indicator analysis should be documented.

able to be substantiated – the results and conclusions drawn during the analysis of an indicator should be capable of being substantiated by an independent authority. In other words, an independent reviewer with relevant qualifications should be able to come to the same conclusions.

- Unbiased – the information used to measure effectiveness should be able to be impartially collected, analysed and evaluated.

For example, when data is reported can result in a bias. Taxpayers who are due a refund tend to lodge earlier than those who will end up having a tax debt. The timing of the collection and analysis may produce different results.

In the spirit of continuous improvement, it is essential to not only embrace our successes but to learn from our experiences. The results must present a true picture of the extent to which we have achieved our success goals.

Understanding strength
Indicators are not used in isolation to understand the extent to which a success goal has been achieved. When considered together as part of a suite, they should provide a defensible picture of the outcomes.

Weak indicators often perform a useful function in the overall suite and should not be discounted on their strength alone. However, too many weak indicators may make it difficult to present a defensible picture. You can attempt to overcome this by exploring ways to strengthen or replace some of the indicators.

comparable – continuity is necessary in order to make comparisons in trend analysis over time.

Data can be subject to continual revision. For example, new data may be received over time that may have an impact on your original benchmark population. While it is not always possible to avoid differences in the underlying data, recognition of those differences and the influence they have on the indicator analysis should be documented.

able to be substantiated – the results and conclusions drawn during the analysis of an indicator should be capable of being substantiated by an independent authority. In other words, an independent reviewer with relevant qualifications should be able to come to the same conclusions.

- Unbiased – the information used to measure effectiveness should be able to be impartially collected, analysed and evaluated.

For example, when data is reported can result in a bias. Taxpayers who are due a refund tend to lodge earlier than those who will end up having a tax debt. The timing of the collection and analysis may produce different results.

In the spirit of continuous improvement, it is essential to not only embrace our successes but to learn from our experiences. The results must present a true picture of the extent to which we have achieved our success goals.

Understanding strength
Indicators are not used in isolation to understand the extent to which a success goal has been achieved. When considered together as part of a suite, they should provide a defensible picture of the outcomes.

Weak indicators often perform a useful function in the overall suite and should not be discounted on their strength alone. However, too many weak indicators may make it difficult to present a defensible picture. You can attempt to overcome this by exploring ways to strengthen or replace some of the indicators.
Consider whether your indicators will show that:
- there has been a change in behaviour of the risk population
- there has been a change in community confidence
- the change has been sustained
- your strategies have contributed to the change
- the effect can be seen more broadly than those targeted by your strategies
- there were any unintended consequences as a result of your strategies.

If not you should go back to the beginning of the indicator identification process and identify more potential indicators. Remember to validate any new indicators.

Seeking endorsement
Once you have finalised your suite of indicators you should seek endorsement from the relevant decision making body. Approval should generally be obtained from your business line’s RMC or equivalent.

The validated suite of indicators should be endorsed by your RMC (or equivalent) before continuing further.

MEASURING YOUR INDICATORS
Once you have decided on your indicators you then need to determine how you will measure them. What method will you use to identify whether:
- a change in behaviour or community confidence has occurred
- the change can be attributed to your strategies.

Identifying a change
Before you can understand whether there has been a change in the behaviour or community confidence, you need to know what the compliance behaviour or community confidence looked like before your strategies were implemented.

Baseline
A baseline is a reference point against which other things can be compared or measured. It is used to identify whether there has been a change over time.

A baseline measurement should reflect the starting point prior to the strategies being implemented. It should then be used for comparison purposes to determine whether there has been a change over time.

Comparing against a single data point, however, is not always a reliable method of identifying a change. While you may see a difference between the original measurement and a later observation, there is no way to understand whether this change is caused by your strategies or whether it simply represents natural fluctuations in the data being observed.

Observing historical trends in the data before your strategies are implemented will help you to identify whether a change has occurred.

Figure 2 looks at on-time lodgment behaviour over a three-year period. Note the general direction evident in the data.

Identifying trends
A trend is the general direction in which something tends to move.

Trend analysis involves the collection and presentation of information in order to identify patterns.

Trend data is usually presented as a set of data points measured over uniformly-spaced, successive time periods.

A trend will be evident where the data shows a general tendency to move in the same direction over time. By drawing a line through the historical data to highlight the trend and extending that line over future observation points, you get a sense of where the subsequent observations should be if the trend is maintained and there is no change.

A trend line does not necessarily need to be straight. Sometimes exponential growth or decline may be observed.

Figure 3 shows a difference between the trend line and the new data point. This indicates that something has affected the behaviour.
Developing Your Suite of Indicators

Analysing trends
Several methods can be used to understand whether there are changes in trends. These include:

- **Pre and post studies**
  Pre and post studies assess changes as a result of an intervention. Measures are taken both before and after the intervention in an effort to understand whether there has been a change.

  For example, tracking behaviour over time, both before and after the compliance strategies are implemented, can help you to understand whether there has been a change in behaviour that corresponds with the implementation of the strategies. But it will not tell you whether the strategies caused that change.

- **Longitudinal studies**
  A longitudinal study involves repeated observations of a specific set of participants over a period of time. It allows you to measure key variables at different points in time. The same data is collected from the same participants at regular intervals. Longitudinal studies are particularly useful for understanding the long-term effects of your strategies.

  When choosing your sample for a longitudinal study you must consider the effect of churn on the sample size. Churn is the number of people who move in and out of a particular group at any given time. If your population has a high churn rate you are likely to lose a significant number of participants along the way. Increasing your sample size to allow for natural attrition of the participants over time is one way to overcome this.

- **Measuring against standards**
  A standard is a reference point against which compliance behaviour or community confidence can be measured. It sets a required level of quality or performance that indicates quality or success.

  For example, measuring behaviour against a standard allows you to see whether behaviour is coming back into line with acceptable standards (see figure 4).
Benchmarks are particularly useful when examining behavioural change at the broader level but may become less useful when drilling down to a specific sub-population.

Inferential statistics
Inferential statistics are a group of methodologies that allow you to draw inferences about a population based on the behaviour of a sample.

When an appropriate benchmark is not available you can use inferential statistics to identify whether or not there is a causal relationship between the behaviour being observed and your strategies.

They use sampling methodologies to identify target and control groups from within a specified population (see figure 6).

EXAMPLE: External benchmark
Australian resident population data from the ABS could be a useful benchmark to understand changes in the number of resident individuals who are registered in the tax system.

Benchmarks are often based on external data, such as ABS statistics, and are comparable across the broader population base.

Several different types of comparisons can be used. Some are suitable for high level measurements (for example, benchmarks) while others are more suitable for drilling down to a sub-population level (for example, target and control group studies).

Benchmarks
A benchmark is a reference point for comparison purposes. It allows you to isolate the effects of the strategies from the effects of external influences such as economic fluctuations.

For example, comparing the behaviour of a risk population to an appropriate benchmark is useful in understanding whether the observed behaviour is in line with expectations. If your strategies have had an effect, you would expect to see a change in the relationship between the benchmark and the compliance behaviour of the risk population.

Attribution
Identifying a change in compliance behaviour or community confidence does not necessarily mean that your compliance strategies caused that change.

Using a comparison technique enables you to separate the effects of your strategies from other factors that could be causing the change. This will help you to identify whether the observed change is a result of something you have done or a reflection of changes in the environment that are outside your control.

Figure 5 shows that both the benchmark and the population of interest were tracking together until the start of the treatment period. The measure taken at the end of the treatment period shows a distinct difference in the behaviour pattern of the two groups.
DEVELOPING YOUR SUITE OF INDICATORS

EXAMPLE: Randomised control trial

Taxpayers are randomly selected from the population and separated into a target group and a control group.

The target group is used to test the effectiveness of an advisory letter on their compliance behaviour.

The control group is isolated from the treatment.

The difference in behaviour between the groups is then used as an indicator of the effectiveness of the strategy.

Theoretically, random selection should result in the selection of samples that are representative of the population. However, where a population has identifiable sub-groups that are likely to behave differently, the accuracy of the study can be further enhanced by randomly selecting cases from each sub-group. This eliminates the possibility of a sub-group being over-represented within the sample.

You should seek expert assistance when setting up these types of comparisons to ensure that you don’t build unintended flaws into your indicator analysis.

You should consider:

- appropriate sample sizes
  - Having too small a sample will limit the confidence that can be attached to the results.

- proper selection techniques for group membership

When using a randomised control trial you need to select your target and control groups before you implement your strategies to ensure that the control group can be isolated from the treatment.

- matched groups
  - When it isn’t possible to use a randomised control trial you can use a matched group instead. This is a comparison group that is similar to the treatment group. Groups are matched according to similarities in their characteristics.

  While the results from a matched study are adversely affected by unknown differences between the groups, well-matched groups can approach the rigor of a randomised control trial.

Seeking expert assistance

Assistance in understanding the broader environment or identifying appropriate external benchmarks can be obtained through internal specialist areas.
Another useful resource is the ABS/ATO Information Management team located in the Office of the Chief Knowledge Officer. This unit can provide assistance in identifying, obtaining and interpreting information from the ABS. Assistance in setting up your comparison studies can be obtained from the relevant business and data analysis specialists within your business line.

IDENTIFYING YOUR DATA REQUIREMENTS

**Who’s responsible?**
You should allocate ownership of the data gathering and indicator analysis tasks. Making your expectations clear from the outset will allow you to identify and deal with potential issues that may impact the timely availability of the information.

**What data sources will you use?**
You must be as specific as possible so that people will know exactly what data is required.

Consider the following questions when identifying your data sources:
- Is the data from an internal or external source?
  - Internal data is data obtained from sources such as the Data Warehouse or our community perceptions surveys. External data is data obtained from sources outside of the Tax Office, such as the ABS.
- Are there any issues around ownership of the data that will affect your ability to acquire it?
  - For example, data from state-owned property databases may require special permission before you can access it. You must allow sufficient lead time for that permission to be obtained.
- Are there any issues with regard to the availability of the data that would affect its usability?
  - For example, there is a significant lag time between the collection and subsequent availability of ABS census data that restrict its usefulness as a benchmark.
- When and how often do you need to collect the data?

You need to obtain data at regular intervals so that progress can be tracked over time.

- Can the data be replicated?
  - The data may need to be replicated for various reasons including:
    - the need to provide the same information at specific intervals for time-based analysis
    - quality assurance processes to independently verify your results.

There is often a cost associated with the collection of data. Potential costs to the ATO can range from payments to an external organisation to the opportunity cost of diverting internal resources to the task of collecting the data. Potential costs to the community include an additional burden on the time and resources of respondents.

Before making a final decision on the data source consider the following:
- Are there any existing data sources that will suit your needs?
  - Collecting new data is generally more expensive and time consuming than collecting existing data. There may be existing data sources such as previous evaluation reports, other agency data or other internal and external data that will suit your needs.
- There is a range of existing surveys and studies that may also suit your purpose. But you must understand the purpose of the original study and any limitations it might have in relation to your particular indicator.
- You may be able to influence the questions that are included in a survey when it is being set up. This can reduce the costs involved with running a separate study.
- Our Corporate Research Centre can provide advice in relation to surveys and other research.
- Is the data in a form that is able to be analysed for your particular purpose?
  - Information which is collected at source (primary data) for the particular purpose of meeting the evaluation objective is more reliable than information derived from a pre-existing, data source. This is because the primary data has been collected with the particular purpose in mind.
  - Secondary data has usually been collected, summarised or interpreted for a different purpose and may not be entirely relevant when used as a substitute for primary data.
- This may weaken the reliability of your indicator.

Before you decide to use a secondary data source you must understand:
- the purpose for which the data was originally collected
- how the data was collected
- the currency of the data
- whether the data is representative of your population of interest.

For more information about the major methods and sources used for collecting data for evaluations, see Appendix 1.
Are there any data limitations?
You need to understand whether there are any limitations or issues that will impact on the reliability of the underlying data. These limitations include:
- changes over time in the method of data collection or compilation
- unavailability of data in some years
- possible time lags in the availability of internal and external data.

What comparisons will you use?
You need to be clear about which comparisons you intend to use to make sense of the changes for each indicator. In making a decision consider whether you:
- know how you’re going to measure your baseline
- will use a benchmark and whether it will be based on internal or external information
- need to select a control and target group
- need to identify a similar group for comparison purposes.

Once you have decided which comparison technique to use, you should identify what needs to be put in place to collect the data.

Documenting your intentions
You must record information relating to the underlying data sources, data limitations and comparisons associated with each indicator. This provides a sound basis for quality assurance processes and allows for independent verification or replication at a later date.
BEFORE PROGRESSING FURTHER, MAKE SURE YOU HAVE:

- validated your indicators
- formulated your data requirements and assigned responsibility for collection and analysis of your indicators
- a suite of indicators that can paint a defensible picture of effectiveness
- received endorsement of your suite of indicators from your RMC (or equivalent).
Data collection and analysis must stay focused on the original intent for which the indicators were developed. Otherwise you will end up “data rich and insight poor”.

GATHERING YOUR DATA
Once sufficient time has elapsed for your strategies to have an effect on voluntary compliance and/or community confidence you need to begin gathering the data to enable you to analyse each indicator.

Who’s involved?
The person or area responsible for gathering data and analysing each indicator should have been identified during the indicator identification and validation processes. It is highly likely that different people will be responsible for each indicator due to the specialised knowledge required to understand and interpret the data.

Data integrity
When collecting your data, ensure that there is a sound process in place to verify and validate it. You should consider:
■ relevance of the data – the data should actually measure the characteristic it claims to measure. This is particularly important if you’re using a data set that has been created for another purpose. If using secondary data, you need to rigorously review your data specifications to ensure that the data is appropriate for your purpose.
■ data availability and collection – ideally, data should be collected in a form that is tailored to your analysis requirements. This won’t always be possible particularly where secondary data is used. If using secondary data you need to explore the underlying data properties to understand any limitations they may pose on your analysis.
■ consistency of data – you must look for data that will be consistent over time. This allows for repeated observations to check whether the effects of your strategies are sustainable. If this isn’t possible, look for ways to mitigate the limitations of the data in the longer term.
■ timing of data – where data is sourced from external databases or reflects external market trends, you must make sure there are no issues with differing time frames when making comparisons. If there are any differences you must account for the time lag in your analysis.
■ expertise of data collectors – data quality can be impacted by the skills and expertise of the people involved in gathering the information. You may need to seek the advice of experts to ensure your data collection techniques are appropriate.

Quality assurance
Your data collection techniques must meet appropriate quality assurance standards. You need to address and document:
■ any limitations of the data source (for example, is the survey poorly designed, are there sampling errors and have these been articulated?)
■ any assumptions made
■ the approach taken and whether it is appropriate (for example, is the data coding approach correct?).

Data storage
Once collected, data needs to be appropriately stored. Storage requirements vary, depending on the type of data. The key elements to consider are:
■ security of the data, particularly where it is classified as in-confidence or above
■ ease of data retrieval.

You also need to store the detailed data specifications in a way that will allow the data to be reproduced at appropriate intervals for cyclical evaluations, ongoing trend analysis and quality assurance processes.

ANALYSING YOUR INDICATORS
Once the data for each indicator has been collected you then need to analyse it to understand where there is any evidence of a change in behaviour and/or community confidence.

The results for each indicator are brought together to determine whether, as a whole, they provide defensible evidence of the success of your strategies in meeting your success goals and ultimately your desired outcomes.

Who’s involved?
Indicator analysis should be done by business or data analysts who understand both the business and the underlying data. In many cases the analysts will need to work closely with other subject matter experts in order to deliver a comprehensive analysis. These experts may include risk and strategy experts, and economists and revenue analysts.

Making sense of the data
To evaluate your effectiveness you need to firstly explore whether there is a change in compliance behaviour and secondly, whether your strategies caused that change.
Is there a change?
An easy way to identify a change is to use a line graph to explore patterns and changes.

To do this you need to:
- plot your compliance data on a line graph, preferably using one that allows you to plot more than one data series. This will make it easier to compare your data with, for example, a benchmark
- identify the point in time when your strategies were implemented
- look at the historic observations before your strategies were implemented and identify any obvious trends
- draw a line that depicts that trend and extend this line into the future
- examine any observations taken after your strategies were implemented and compare them with the previous trend. If there is a difference you will be able to conclude that there is a change.

For more information on baselines and trends, see pages 12–13.

However, where the historical data shows considerable fluctuation it may be difficult in the short term to identify whether a change has occurred. You may need to continue monitoring over time to understand whether there is a difference.

Did the strategies cause the change?
Once you have identified a change you need to see if that change can be attributed to your strategies or whether it is a reflection of changes in the environment. Comparing your data to something else, such as a benchmark or control group, will help you to isolate the effects of your strategies from the general fluctuations caused by other changes in the environment.

To do this you need to:
- look at the comparison method identified earlier
- plot your comparison data on the graph
- compare the pattern of movements in both the comparison data and your compliance data to see if there are any differences.

If the data shows a difference in the pattern of movement after the strategies were implemented, you can say with some confidence that your strategies had an effect on compliance behaviour or community confidence. If, however, your compliance data displays the same pattern as the comparison statistic, you can conclude that the observed changes in behaviour were caused by environmental factors.

Comparison data can be generated through the use of benchmarks and inferential statistical techniques such as randomised control trials or matched groups.

Is the change in line with your expectations?
Sometimes your strategies can have a different effect than originally anticipated. Where you see a change that is unexpected you must explore it further.

If potential unintended consequences have been anticipated from the outset there may be indicators specifically designed to see whether these consequences have emerged.

Be careful not to view the emergence of an unintended consequence as necessarily negative. Some may, in fact, be positive and even if negative the insights they provide are valuable in re-shaping our compliance improvement strategies.

Identifying sustained change
A sustained change is one that lasts for the longer term. Identifying whether a behavioural change has been sustained needs to be done with repeated measures over time.

Figure 7 shows a change in the relationship between the compliance behaviour and the comparison statistic that corresponds with the treatment period. The 2008–09 observation indicates that the new relationship is holding.

FIGURE 7: Identifying a sustainable change in behaviour

![Diagram showing compliance behaviour and comparison statistic over time.](image)

If the behavioural change is sustained, you would expect to see the paths of the two groups continuing to track in parallel over an extended period of time as seen in figure 8.
There is no set period of time over which you should track your data to identify whether a change is sustainable. However, you will need to track behaviour for several years to really understand whether your strategies have had a lasting impact.

**DETERMINING YOUR EFFECTIVENESS**

While the analysis of each indicator provides some useful insights, it won’t provide you with a defensible picture of the extent to which your strategies have improved voluntary compliance and/or community confidence. You need to bring together the analysis results for all of the indicators and examine them as a whole to determine the extent of your effectiveness and draw defensible conclusions.

**Who should be involved?**

Evaluating compliance effectiveness should be done by people who understand both the risk and the environment affecting that risk. While the risk manager will generally take responsibility for making sense of the story, it’s good practice to involve other subject matter experts to help make sense of the information as a whole.

As for the indicator analysis, the people that should be involved in bringing together a defensible story should include the relevant risk and strategy experts as well as analysts who have an understanding of the environmental variables that can affect your risk area.

**Bringing the story together**

When the suite of indicators was originally chosen, a number of assumptions were made about what you’d expect to see if your strategies were effective. Now that you’re ready to bring the results together, you should revisit these early assumptions to see whether the indicator results are as expected.

If the behaviour is in line with your earlier expectations then determining the extent of your effectiveness should be relatively easy. If any indicators show a different result, however, you must objectively investigate the result to understand how it impacts on your conclusions about the extent of your effectiveness.

Where an indicator shows an unexpected result explore whether:

- the result is significant in relation to the overall suite of indicators – if the indicator is weak its impact on the overall outcome may not be significant. If it carries a lot of weight within the suite you need to understand how it affects the outcome.
- any other indicators show similar results – if the result is supported by a range of other indicators, then it would be reasonable to conclude that the strategies have not been effective.
Table 2 looks at evaluating the extent to which you have achieved a success goal. It shows the linkage between the indicators and the success goal. A good way to evaluate a success goal is to turn the goal into a question. In this case, the question would be: To what extent does the community demonstrate confidence in our ability to deal with non-compliance relating to [risk]?

In this example, the overall conclusion is supported by the evidence.

### Drawing your conclusions
Evaluating compliance effectiveness is based on understanding whether your compliance strategies have moved you towards achieving your success goals and ultimately your desired outcomes.

Look at the indicators aligned to each success goal and decide whether, collectively, they show the extent to which you have achieved each goal. If your evaluation shows that you have met your success goals then you can, with a reasonable degree of confidence, conclude that you have made progress towards achieving your desired outcomes.

### TABLE 2: Evaluating the extent to which a success goal has been achieved

<table>
<thead>
<tr>
<th>Success goals</th>
<th>Indicators</th>
<th>Evaluation outcome</th>
</tr>
</thead>
</table>
| The community is confident that the ATO is able to deal with non-compliance relating to participants in the risk | ■ Trend in the number of reports made by the community to the Tax Evasion Referral Centre (TERC).  
■ Responses to our surveys including community, business and tax practitioner perceptions surveys.  
■ Levels of engagement of representative bodies.  
■ Frequency and tone of media comment. | Conclusion  
While survey results have remained stable over the period 2006–08, the increased levels of engagement by the community indicate an improvement in the community’s confidence in our ability to deal with non-compliance relating to this risk.  
Summary of evidence  
■ There has been a 75% increase in the number of reports made by the community to the TERC in the 2007–08 income year compared to the 2005–06 income year.  
■ Perceptions survey results since July 2005 suggest that the community is confident in our ability to deal with participants in the risk.  
■ There has been an increased willingness by representative associations to work with us to design, develop and promote fact sheets to their members and others which is indicative of an improvement in the levels of engagement.  
■ Media commentary, specialist advice and letters to the editor in the past 12 months have been supportive of our approach. This contrasts with negative commentary in relation to several high profile cases in the previous period of review. |
Making recommendations
The lessons learned from evaluating effectiveness can help re-shape our compliance improvement strategies.

Consider the following questions when making your recommendations:
- Was the right mix of compliance strategies used to address the risk? If not, what mix would be better?
- How could your strategies be improved?
- Does more risk assessment work need to be done (for example, more research into a particular population’s compliance behaviour)?
- Did any unintended consequences emerge? How might any negative outcomes be mitigated?
- How can the application of the compliance effectiveness methodology be improved (for example, better articulation of success goals, data collection methods or analysis techniques)?

Presenting a quality outcome
A quality evaluation of whether, and to what extent, your strategies have achieved your success goals and ultimately your desired outcomes can provide you with a wealth of knowledge on which to base future risk treatment decisions.

To ensure that you can present a quality evaluation consider:
- the analysis methodology – choosing the right methodology is particularly important when looking at attribution issues. You need to select a methodology that can identify whether changes are caused by your strategies or by factors outside your control.
- any over-emphasis on one or more indicators – while it is acceptable for some indicators to carry more weight in the overall picture than others, you must take care not to rely too much on particular indicators simply because they tell the story that you want them to. Remember, the value you get from evaluating compliance effectiveness is not about success or failure but rather from the lessons learned and how they feed into the continuous improvement process.
- personal bias – you need to resist the temptation to manipulate your evaluation outcomes to paint only a positive picture. Evaluating effectiveness is not meant to be a statement of success or failure. It is meant to provide an understanding of whether you are on the right track with your strategies and to provide some guidance for future planning. Smoothing over any negative results can mask valuable lessons that could influence future risk treatment plans.
- defensible versus definitive – understanding what the indicators are really telling you can be difficult due to the subjectivity inherent in arriving at a conclusion.

Remember, evaluating your effectiveness is about presenting a defensible story. It is not about providing a definitive measurement. Knowing where to draw the line between definitive and defensible is quite difficult because your level of comfort in drawing conclusions about effectiveness is intrinsically linked to your personal risk stance.

When deciding on the level of information you need to support your story of effectiveness, think about how comfortable you would be if you were asked to justify your conclusions. Involving others with relevant business knowledge in the evaluation can go a long way to increasing your comfort levels.

Quality assurance
Involving the right people throughout the process will help you to increase the quality of your conclusions.

As a final step, however, you should have your claims independently verified to ensure that your conclusions are soundly based.
REPORTING ON EFFECTIVENESS
Evaluating compliance effectiveness presents an opportunity for continuous improvement based on an improved understanding of the risk. Presenting a report is the main means of communicating the outcome of your evaluation. It is usually prepared by the risk manager in consultation with other subject matter experts.

**When to report**
The regularity of the reporting cycle can be affected by a range of issues including:
- the planning cycle – you need to identify when the report will be needed for decision-making purposes. Understanding the planning and reporting cycles will help you to determine when the report is likely to be required.
- the evaluation cycle – the priority, pattern and severity of the risk may also require an evaluation to be undertaken on a more regular basis.
- Significant external factors such as volatility in the market place or the economic environment may also impact on the timing of an evaluation.
- other business reporting cycles – compliance effectiveness may need to be included in other reporting requirements at the business line and corporate levels.

You should consult your RMC (or equivalent) to identify the most appropriate time to produce and present your report.

**How to present your report**
You need to prepare a detailed report for reference purposes. You also need to prepare a summary document that presents the key points in a clear and concise form.

The actual format that you use to present your report may differ according to your business line’s reporting requirements and those of the forum to which it is being presented. The content, however, will be the same regardless of the intended audience or the required format.

**THE EXECUTIVE SUMMARY**
Your report should be presented in a way that makes the information easy to understand. The best way to do this is to present an executive summary supported by a more detailed report.

The executive summary should provide the reader with an understanding of your conclusions along with a brief outline of the evidence underlying those conclusions. It should also provide recommendations for the future. Your executive summary should usually be no longer than two pages.

**What should you include in the summary?**
You should include responses to the following questions:
- **Have you been effective?**
  Include a brief summary outlining your conclusions about the extent to which you achieved the desired outcomes.
  You should also discuss whether there have been any ripple effects and whether the results have been sustained.
- **How do you know?**
  Provide an outline of the evidence that supports your conclusions. Include brief details of any data limitations, assumptions and inferences that are relevant to your conclusions. You should also include any relevant contextual information, such as the effect that other factors may have had on the behaviour or the results, which will help the reader make sense of the story.
  Don’t at this point delve into the detail. Remember that the executive summary is meant to provide a high level view of the outcomes. A more detailed report can be made available for further reference.
- **What have you learned?**
  Include what you learned during the evaluation including:
  - improvements in your understanding of the risk
  - what treatments did or didn’t work
  - any unintended consequences.
- **Where to from here?**
  Include any recommendations for future risk treatment strategies. Consider:
  - the mix of strategies. For example, a greater emphasis on help and education rather than active compliance.
  - future risk assessment work. For example, there may be a need for more research to increase your understanding of the risk population.
  - the way the compliance effectiveness methodology is applied in the next evaluation cycle. For example, there may be a need to refine the desired outcomes, success goals or indicators if they were unrealistic, unsuitable or vague.
  - the evaluation process. For example, it may be necessary to set up systems to capture more useful data.
THE DETAILED REPORT
You need to prepare a detailed report of the evidence that supports your conclusions. It should include:

- the analysis results for each indicator focusing on the specific behavioural shifts that occurred and whether these changes were evident in the wider population or restricted to the specific target group
- a discussion of the evidence that supports your conclusion that the change in behaviour and/or community confidence can be attributed to your compliance strategies
- whether the suite of indicators demonstrates the extent to which you’ve achieved each of your success goals and ultimately your desired outcomes.

Where you’ve identified a positive shift in compliance behaviour or community confidence you should discuss whether:

- it is possible, at this point, to identify a sustained change. It may be too soon to determine whether a behavioural change can be sustained over the longer term.
- any unintended consequences have emerged. You should include suggestions for mitigating any negative outcomes in the future.

In your recommendations include any findings about:

- which compliance strategies worked and which didn’t and any suggestions for improvements in the future
- whether the right mix of strategies was used to address the risk
- whether the risk is still manifesting in the same way or whether it has changed
- whether more risk assessment work is necessary to understand the risk.

Presenting your report
Before submitting your report to your RMC you should submit both your detailed report and executive summary to your risk owner for endorsement.

Once that has been done you should submit your report through the normal channels.
APPENDIXES
## DATA COLLECTION METHODS AND SOURCES

<table>
<thead>
<tr>
<th>Method or source</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ATO data warehouse</strong></td>
<td>Data from tax returns, activity statements and other forms lodged by taxpayers and intermediaries.</td>
<td>Readily accessible</td>
<td>Misunderstandings sometimes occur between analysts and data extractors. Data specifications must be discussed in detail.</td>
</tr>
<tr>
<td><strong>Corporate surveys</strong></td>
<td>Data collected from a sample of the population. Market research surveys are generally made up of a range of questions aimed at gauging public opinion. ATO corporate surveys typically cover a range of issues, including awareness and understanding of tax obligations, perceptions of tax fairness and attitudes towards the ATO and tax compliance. These surveys include our Business Perceptions Survey, Community Perceptions Survey, Professionalism Survey and a range of tax practitioner surveys. Surveys are also commissioned, or run in-house, on specific topics such as influences on taxpayer claiming behaviour, awareness and satisfaction with specific products and understanding of specific tax obligations.</td>
<td>Offers anonymity, Relatively inexpensive to administer, Easy to compare over time provided questions remain the same or similar, Can cover many people, Able to collect large amounts of data, Many sample questionnaires already exist, Can be tailored to specific issues at the risk, product and market levels as well as the corporate level.</td>
<td>Might not get carefully considered feedback, Wording can bias clients’ responses, Not the full story – breadth rather than depth, Danger of ‘over-surveying’ certain segments (e.g. tax practitioners), Need a certain level of expertise for reliable sampling, questionnaire design, data collection and analysis. This limits the opportunity for in-house surveys and puts the onus on commissioning experts, Sampling and non-sampling errors, The sampling error is associated with inferring the whole population characteristics from a small sample, The non-sampling error includes non-response bias and biases introduced by the framing or order of questions.</td>
</tr>
<tr>
<td><strong>Interviews</strong></td>
<td>Used when you want to fully understand someone’s impressions or experiences, or to learn more about their answers to questionnaires.</td>
<td>Obtains the full range and depth of information, Develops a relationship with the client, Flexibility with the client, Data is collected with a specific purpose in mind which makes the data more consistent with the objective.</td>
<td>Often time-consuming, Difficult to analyse and compare, Potentially costly, Interviewer may bias clients’ responses which may make it difficult to legitimately compare results from different interviewers, Generally can’t provide statistically reliable samples.</td>
</tr>
</tbody>
</table>
# DATA COLLECTION METHODS AND SOURCES

<table>
<thead>
<tr>
<th>Method or source</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| **Focus groups**  | Used to explore a topic (for example, reactions to an experience or suggestion, understanding common complaints) in depth through group discussion. | - An efficient way to obtain good range and depth of information in a short time  
- Data is collected with a specific purpose in mind which makes the data more consistent with the objective  
- Valuable for exploring new areas and developing ideas and hypotheses. Often used as a preliminary data gathering exercise to help inform the design of a more comprehensive survey. | - Problematic to analyse responses  
- Need a good facilitator for safety and closure  
- Difficult to schedule a group of 6–8 people together at the same time  
- Unrepresentative of wider population  
- Tendency of participants to treat the experiment as a game rather than a real-life situation  
- Generally can’t provide statistically reliable samples |

| **Case studies**   | Provide an in-depth examination of a single instance or event. | - Provides a deeper understanding of the behaviour of a particular kind of taxpayer or segment which may lead to the generation of hypotheses and more extensive research  
- Can be undertaken, to some extent, without the need for taxpayer co-operation because existing data can be used. | - Usually time consuming to collect, organise and describe  
- Represents depth of information rather than breadth and may not be necessarily representative of the wider population  
- Costly to conduct when time and resource intensive  
- Generally can’t provide statistically reliable samples. |

| **Academic research** | Academic researchers publish tax-related research on a range of topics. | - Researchers are highly-qualified and specialised  
- You can commission academic researchers to do specific work  
- The work may already have been done. | - Information might be too broad and general for our purposes  
- Usually a time lag exists between research and publication which means the results might not be applicable to the period of interest. |

---

**Notes:**
- **Focus groups** provide qualitative data.
- **Case studies** provide qualitative data.
- **Academic research** provides both qualitative and quantitative data.
# Data Collection Methods and Sources

<table>
<thead>
<tr>
<th>Method or Source</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| **Public opinion research**  | Provides qualitative data Library & Information Services subscribe to several databases which may be useful for this kind of research. Can include market research surveys and scans of topical issues receiving media attention. | - Gives a broad and current view of public opinion  
- A partial analysis can provide some useful context to other findings. | - Can be difficult to analyse for various reasons including:  
- Media scans need to be fairly comprehensive over the period of review and range of media  
- Need to use consistent and standard definitions for terms such as positive and negative media perceptions  
- Resource intensive and likely to mean that most business areas wouldn’t conduct their own ongoing media analysis. |
| **Other government agencies** | Material can be found on the relevant government websites. A good place to start is the Australian, State, Territory and Local governments website at www.gov.au, which has links to federal and state government sites in Australia. Most government websites allow access to a range of published information which provides information on external factors or external benchmarks relating to tax compliance indicators. These include:  
- Australian Bureau of Statistics (ABS) – provides a wide range of statistics  
- Reserve Bank of Australia (RBA) – has a range of statistics on financial indicators  
- Productivity Commission – undertakes research on economic, industry, business and demographic issues including a number of business benchmarking studies. | Easy to access | - Might use units of measurement (for example, different definitions for particular population segments) that aren’t specific to our needs. |
| **Professional associations** | Professional associations and industry bodies represent the interests of their members and regulate their practice for the benefit of the community. Some associations publish regular updates and reports relating to their industry. May provide information on external factors or external benchmarks relating to tax compliance indicators | - Might use units of measurement (for example, different definitions for particular population segments) that aren’t specific to our needs.  
- Some reports are only available by paid subscription. |
APPENDIX 2

REFERENCES

Evaluation references
Australian National Audit Office 1998, Better practice principles for performance information, Australian National Audit Office, Canberra
Babbie, ER, 2006, The Practice of Social Research, 11th edn, Thomson/Wadsworth, California
Bamberger, M, Rugh, J, Mabry, L, 2006, RealWorld Evaluation: working under budget, time, data, and political constraints, Sage Publications, California
Department of Finance 1994, Doing Evaluation: A practical guide, Department of Finance, Canberra
Organisation of Economic Cooperation and Development 2006, Tax Administration in OECD and selected non OECD countries: Comparative information series, OECD, Paris
Owen, JM, 1993, Program Evaluation: Forms and Approaches, Allen & Unwin, Australia
Patton, MQ, 2002, Qualitative Research & Evaluation Methods, 3rd edn, Sage Publications, California
Posavac, EJ, 2007, Program evaluation: methods and case studies, Pearson Prentice Hall

Other useful tools and resources
Information is also available on these websites: Australasian Evaluation Society http://www.aes.asn.au/
American Evaluation Association http://www.eval.org/
## GLOSSARY

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison statistics</td>
<td>Statistics or data that are used as a comparison. For example, a benchmark or control group can be used as a comparison statistic.</td>
</tr>
<tr>
<td>Component</td>
<td>A part of a larger whole.</td>
</tr>
<tr>
<td>Data point</td>
<td>The individual point where a value is graphed, as a point on a line, bar or pie slice. Each data point maps to an individual cell in a spreadsheet's data range. For example, the percentage of activity statements lodged on-time for the 2005–06 income year.</td>
</tr>
<tr>
<td>Data series</td>
<td>A complete series of data, corresponding to the same type of data points. For example, the percentage of activity statements lodged on-time over a period of seven income years.</td>
</tr>
<tr>
<td>Inferential statistics</td>
<td>A group of methodologies that allow you to draw inferences about a population based on the behaviour of a sample.</td>
</tr>
<tr>
<td>Primary data</td>
<td>Data which is collected at source for the particular purpose of meeting the evaluation objective.</td>
</tr>
<tr>
<td>Risk management committee</td>
<td>Oversees all compliance risk work in a business line and is responsible for prioritising compliance risk work for that business line.</td>
</tr>
<tr>
<td>Sample</td>
<td>A small part or quantity intended to show what the whole is like. For example, a subset of the population of interest.</td>
</tr>
<tr>
<td>Secondary data</td>
<td>Data that has been collected, summarised or interpreted for a different purpose.</td>
</tr>
<tr>
<td>Trend</td>
<td>A general direction and tendency.</td>
</tr>
<tr>
<td>Trend line</td>
<td>A line indicating the general course or tendency of something. For example, a set of points on a graph.</td>
</tr>
<tr>
<td>Trend analysis</td>
<td>Involves the collection and presentation of information in order to identify patterns.</td>
</tr>
</tbody>
</table>
MORE INFORMATION

For more information or to provide feedback on this publication, email our Compliance Effectiveness Centre of Expertise at ComplianceEffectivenessCoE@ato.gov.au

More information is also available in our publications:
- Literature review – Measuring compliance effectiveness (NAT 71078)
- Measuring compliance effectiveness – Our methodology (NAT 72341)
- Measuring compliance effectiveness – Applying our methodology (NAT 72342)
- Developing effective compliance strategies (NAT 72788).

To view these publications visit www.ato.gov.au/complianceeffectiveness

To order printed copies visit www.ato.gov.au/onlineordering