UNDERSTANDING EVALUATION FOR EQUITY PROGRAMS

A guide to effective program evaluation

Dr Ryan Naylor, The University of Melbourne

Make tomorrow better.

ACKNOWLEDGEMENTS

CONTENTS

EXECUTIVE SUMMARY

The most important aspect of evaluating a program is that it has clear, unambiguous and measurable aims and objectives.

This guide is intended to help equity practitioners — particularly those who lack experience or confidence in program evaluation — to develop effective evaluation strategies for their programs. It is intended to complement existing institutional guidelines on evaluation, and acts as an initial introduction rather than a definitive guide to the subject.

Measuring the impact of equity programs is essential. Every program can benefit from having an evaluation plan, and evaluation should be a part of equity practitioners' normal business. Evaluation is an opportunity to highlight the strengths and innovations of a program, as well as to identify and address weaknesses. Its primary purpose is to help you maximise your program's effectiveness, not as an imposed performance management structure.

Evaluation involves gathering, reviewing and reflecting on information about your program. It follows a four step process, of establishing criteria; constructing standards; measuring performance and comparing with standards; and synthesising and integrating evidence into a judgement of worth. This same process applies whether you are evaluating a project plan, a pilot, or a fully operational project.
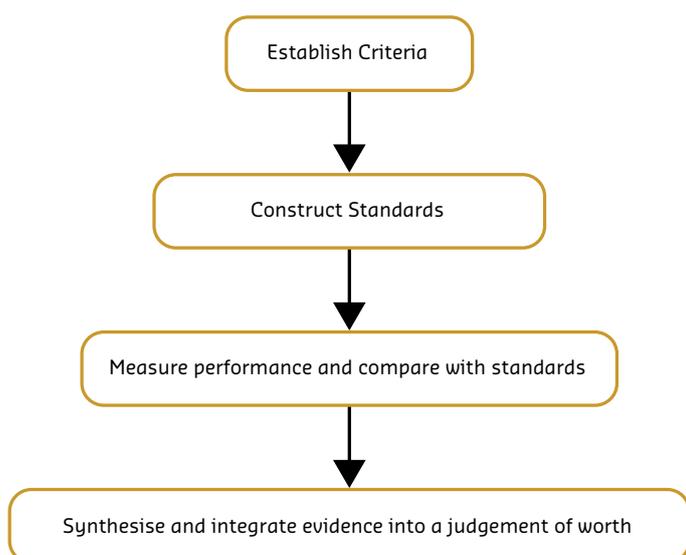
```
┌─────────────────────────────────┐
│       Establish Criteria        │
└─────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────┐
│       Construct Standards       │
└─────────────────────────────────┘
                │
                ▼
┌───────────────────────────────────────────┐
│ Measure performance and compare with standards │
└───────────────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────────────────────┐
│ Synthesise and integrate evidence into a judgement of worth │
└─────────────────────────────────────────────────┘
```

The most important aspect of evaluating a program is that it has clear, unambiguous and measurable aims and objectives. The program aims determine what evaluation questions are asked at each phase of the program's life; which evaluation methods are appropriate to gather evidence to answer those questions; and which indicators and measures should be captured as evidence.

Appropriate alignment between program aims, evaluation criteria, methods and evidence gathered is essential. Having a detailed evaluation plan can be really useful in any program, to ensure this alignment is achieved. There are no easy answers to evaluating equity programs with rigor and detail. Typically, using a mixed methods approach is the one most likely to provide a full understanding of an equity program.

Evaluation should not be a burden. A clear evaluation plan will ensure data is gathered methodically, at time-appropriate intervals, and without wasting time gathering unnecessary data. The purpose of this guide is not to insist that practitioners collect more information.

Equity programs will benefit from formal or informal evaluation throughout their lives. As programs mature, the sorts of evaluation questions being asked should change. Evaluation during the planning phase may strengthen a project by clarifying its aims and improving the alignment between its aims and design, or with institutional targets. It may also establish criteria and standards for evaluation during later phases. Monitoring the program during its operation ensures the program is achieving its aims, and is being delivered effectively and consistently across multiple sites or practitioners. Finally, evaluation during the summative phase seeks to establish which aspects of the program worked and why, and what its impact has been. To fully capture the impact of your project, examining both unintended and intended outcomes can be very useful.

This guide is divided into 3 major parts.

1. **Introduction**
   How do you evaluate an equity program?
   This section of the guide provides a brief description of the major concepts in evaluating equity programs, and how to design an evaluation plan for a program.

2. **When Do You Evaluate An Equity Program?**
   The next 3 sections provide resources to develop effective evaluation strategies at each of the 3 main phases of an equity program:
   1. The planning phase;
   2. Monitoring;
   3. The summative phase.

3. **Further Resources**
   This section provides references for further reading, and some questions and prompts to assist you in planning specific programs.

Dr Ryan Naylor
Lecturer in Higher Education
Melbourne Centre for the
Study of Higher Education
(MCSHE)
The University of Melbourne

INTRODUCTION

This guide is intended to help equity practitioners to develop effective evaluation strategies for their programs.

Measuring the impact of equity programs is essential. Evaluation should be a part of your normal business.

A clear evaluation plan can be really helpful in any equity program.

Evaluation is an opportunity to highlight the strengths and innovations of a program, as well as to identify and address weaknesses. Its primary purpose is to maximise your program's effectiveness, not as an imposed performance management structure.

Why is evaluation of equity programs important?

Equity in higher education is based on the idea that those with the capability and desire to attend university should be able to, and that staff and students should work in an inclusive and supportive environment, regardless of their backgrounds. Equity programs are strategies to achieve this. The most important aspect of an equity program is therefore its impact. Evaluation measures impact and helps you understand how and why you're achieving those outcomes. Its primary purpose — and value — is to help you most effectively achieve your program's goals.

Equity programs that are not having an impact are not helping anyone. In fact, because they divert funds from other potential programs that might be more successful, they actually undermine those goals. Evaluation is simply how this evidence of impact is gathered, from the initial stages of a program, through monitoring during its operation, to a detailed summation at the end of an iteration of the program. This is simply a matter of good practice. Intuition is not enough to establish what works or how successful an equity program is. Good practice demands a strong evidence basis. Evaluation — even informal evaluation — must be part of the design of all equity programs, and, where possible, should be built in from the very beginning (rather than an ad hoc report at the end, as is sometimes the case).

What is evaluation?

Evaluation is a part of life for equity practitioners. An evaluation plan is a requirement of most grants, whether funded by national bodies or through an institution's internal funding. Some practitioners are very comfortable with the idea and have a lot of experience in the area. Others are less comfortable.

Evaluation and reporting are not the same thing. Reporting is primarily an operational process to ensure that, for example, external funding has been spent appropriately or performance management targets are being met. The primary audience is typically people outside the equity program itself.

In contrast, the evidence collected during program evaluation is as much for your own purposes as for external audiences. Evaluation highlights the strengths of the program and pinpoints innovative practice, as well as identifying and addressing areas of weakness for future practice. It is as much about establishing what is good about a program as addressing deficits, and is an opportunity to reflect and learn from your work and that of your colleagues.

Evaluation of equity programs seeks to answer three basic questions:

1. Are our programs consistent with our strategic objectives (institutional or government, as well as program goals) in relation to equity and social inclusion?
2. Are our programs achieving the desired outcomes?
3. Do any aspects of our programs need improvement?

Evaluation is therefore not a difficult task (theoretically, at least). Evaluating an equity program just involves collecting rigorous, detailed evidence to answer these questions.

In practice, however, evaluation can be a complex task. Establishing causal relations between your program and a particular outcome is often difficult. People may be acted on by multiple programs within a broader policy and program environment, or too many variables exist to control for in any rigorously methodological way. Definition and measurement issues may also complicate things.

What is the purpose of this guide?

This guide is intended to help those involved in equity programs — particularly those who lack experience or confidence in program evaluation — to build effective evaluation strategies into their daily work. It is intended to be a practical resource to assist at each stage of the evaluation process. "Equity program" will be used as a general term for any projects associated with social inclusion or widening participation in higher education.

Evaluation should not be burdensome. Everyone collects evidence of effectiveness. This guide is intended to help you ensure you're collecting the right types of data, and to help you analyse and reflect on it, rather than insist that you collect more.

THE FUNDAMENTALS OF EVALUATION

Evaluation involves gathering, reviewing and reflecting on information about your program.

The logic of evaluation is a four step process, of establishing criteria; constructing standards; measuring performance and comparing with standards; and synthesising and integrating evidence into a judgement of worth.

It is essential that the aims of the program and the evaluation questions being asked are completely clear and unambiguous.

The criteria of evaluation (or evaluation question) determine what types of evidence and data collection methods should be used.

Mixed methods approaches are often helpful to gain a full understanding of an equity program.

A key aspect of evidence-based decision-making is having access to the information relevant to making that decision. Evaluation is simply the process of gathering, reviewing and reflecting on that information. For new programs, it should be built into the design of equity programs from the program's initial proposal to ensure that the right information can be collected through the life of the project. For pre-existing programs, explicit aims and a clear evaluation strategy should be developed as soon as possible (where they don't already exist). Evaluation helps you effectively and efficiently collect the right information. If it's part of your routine activities, you and your program will benefit.

The Logic of Evaluation

The process of evaluation, which Fournier calls "the logic of evaluation", consists of 4 steps:

```
┌─────────────────────────────────────────────────────────┐
│                   Establish Criteria                      │
└─────────────────────────────────────────────────────────┘
                            │
                            ▼
┌─────────────────────────────────────────────────────────┐
│                  Construct Standards                      │
└─────────────────────────────────────────────────────────┘
                            │
                            ▼
┌─────────────────────────────────────────────────────────┐
│        Measure performance and compare with standards     │
└─────────────────────────────────────────────────────────┘
                            │
                            ▼
┌─────────────────────────────────────────────────────────┐
│   Synthesise and integrate evidence into a judgement of worth │
└─────────────────────────────────────────────────────────┘
```
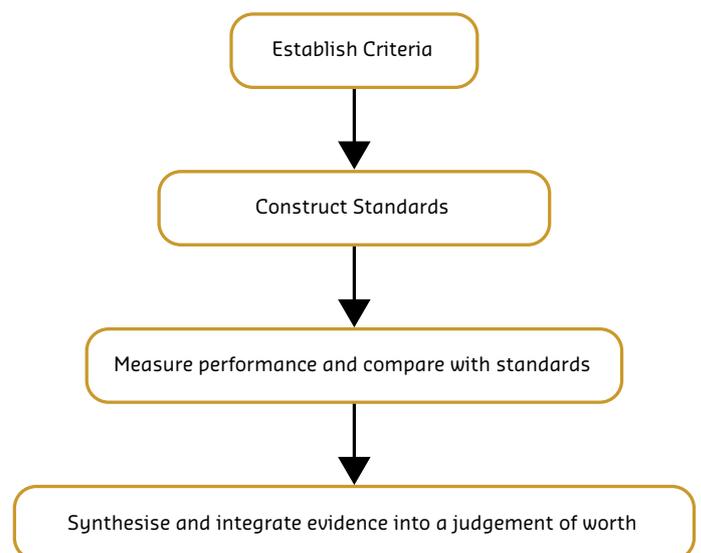
Establishing criteria – asking the right questions

The first step in performing an evaluation is establishing which aspects of an equity program are to be examined — that is, deciding which questions are being asked of it. Programs at different stages of maturity may have different criteria for evaluation; a list of typical questions that may be asked at each stage is provided in the latter half of this guide. You might sometimes evaluate specific aspects of a program, rather than the whole. However, whether the program is achieving its desired outcomes, and contributing to the strategic objectives of the institution, remains a vital, higher level question at each stage.

The "ultimate goal" of equity in higher education is not a simple or uncontested concept. Generally speaking though, equity programs are based on at least one, and sometimes several, of the following broad principles:

- Informing aspirations and developing expectations towards higher education;
- Developing academic capacity and providing academic support;
- Establishing inclusive processes;
- Supporting students in dealing with broader issues outside of their academic lives.

Outreach programs in schools, for example, may seek to inform the aspirations of students regarding university, increase preparedness (by informing subject choices and increasing school retention rates, increasing familiarity with the university environment, or boosting academic engagement or achievement), and improve support for students (by boosting parental and teacher involvement and support, and clarifying pathways into university). In contrast, financial aid and discounted childcare for students are intended to minimise obstacles from "real life" that might otherwise interfere with their study, whereas providing training for academics or peer mentors on cultural inclusiveness is about establishing inclusive processes.

These four aims form the basic criteria against which the impact of the program should be evaluated, and the types of evidence that should be collected. For this reason, it is essential that the aims of the program are completely clear and unambiguous. This will be discussed in more detail below. Some aims may be very concrete (such as increasing the number of enrolments from a particular school), whereas others may be less so (increased community engagement with the university, or increased awareness about structural inequalities in promotion panels, for example). Intangible outcomes can be very important in equity programs, but are often harder to measure. Being clear about your aims helps you to make sure you are getting at the right outcomes, even when those outcomes are difficult to evaluate.

Other evaluation criteria, such as cost, may also be considered, although what is considered cost-effective depends on the institution, the specifics of an individual program, and its impact. Again, these could be concrete (e.g. cost), or not (e.g. reputational benefits for universities).

Constructing standards
In order to establish whether a program is successful, it is obviously important for you to have some idea of what "counts" as success, or how well the program should perform. This can be a very complicated and program-dependent process.

Typical standards include key performance indicators (KPIs) and institutional targets, in which case the standards are often determined by those outside the program; or comparing the program to other, similar programs.

Comparing one program with another (benchmarking) is hampered by the lack of rigorous, detailed studies published. Even between superficially similar programs, there can be great variation in cost, intensity and approach, which may confound comparison. However, it does provide some guidance on what results are realistic with a particular type of program. Ideally, KPIs should be based on benchmarking via the literature and experience, but this may not always be possible and you might not be able to do it routinely.

Constructing appropriate standards frequently requires experience and flexibility. There are no general solutions, although the chosen standards should reflect an institution's context and objectives, as well as the aims of the equity program itself, and the evaluation questions being asked. Discussion with your colleagues and managers may be useful.

Measuring performance
There are multiple types of evidence that can be used to measure a program's impact: attendance numbers, test results, surveys of user perceptions, interview data, and so on. The most important aspect of choosing what methods to use and what types of data to collect is the aims of the program and the criteria used to evaluate it. The criteria of evaluation determines what evidence is appropriate.

For example, using a survey to evaluate an outreach program in terms of its success in informing aspirations is probably appropriate. However, a survey of participant satisfaction probably won't effectively measure an increase in academic preparedness (unlike pre- and post-program testing). Simply reporting the number of participants in the program or how enjoyable it was does not provide evidence of impact for either aim.

Many equity programs define impact in terms of changes in attitudes, performance or behaviour. It is therefore often important to measure these factors before and after the program, which is why having an evaluation plan from the earliest stages of program design is useful. It may be very difficult to reconstruct some types of evidence at the end of a program, and not having this evidence may make it impossible to determine whether the program was successful or not.

Mixed methods research is often seen as an appropriate tool to gain a full understanding of an equity program. Evaluation is a specialised form of research, and many projects may need a range of evaluation methods to measure all of their outcomes.

It is beyond the scope of this guide to provide a full description of mixed methods research. Creswell and Plano Clark (2007) is an authoritative textbook in the area, and Creswell et al. (2008) describe some important methodological issues and designs to consider when designing a mixed methods study. Further reading is also provided in the last section.

Quantitative methods tend to be useful to assess broad trends, group responses, and measures of efficiency. Examples of quantitative methods include surveys, econometrics and mathematical modelling.

Advantages of quantitative methods include:

- Relatively quick to conduct and analyse data (and many data analysis software packages are available to do so);
- Findings such as effect sizes and statistical significance are relatively independent of researcher subjectivity;
- The apparent precision of findings may lend them higher credibility with some stakeholders.

However, researchers may miss important phenomena if the instruments used are not well designed, or the findings may not accurately reflect participant's understandings.

Qualitative research is able to give more detailed views of individuals and their experiences, but is often more time- and labour intensive to use. Examples include individual or focus group interviews, case studies or action research.

Advantages of qualitative methods include:

- More complex phenomena can be studied according to participants' own categories of meaning;
- Rich detail can be generated;
- Contextual factors may be identified;
- More responsive approaches may be used to gather data.

However, qualitative approaches may create findings that may not generalise to other settings (making it more difficult to make predictions), may be more easily influenced by the researcher, and may have less credibility with some audiences.

Central to mixed methods research (and this guide) is that the evaluation question is fundamental, and the evaluation methods used should follow logically from the questions being asked . The methods used may therefore be combined in a number of different ways: a qualitative component may precede the quantitative, or vice versa, or they may be conducted concurrently, depending on the criteria of evaluation.

It is important that the methods selected are feasible within the cost and time constraints of the program. Qualitative methods, for example, are often much more time consuming than many quantitative methods, and so are more suitable for evaluating pilot groups, small programs, or test groups within larger programs. It may not be possible — or desirable — to qualitatively assess the impact on every student in a large program. On a related note, it is very difficult to determine impact for some types of programs (outreach to primary or early secondary school students, for example, or to pin down the effects of one program when students are involved in several different programs) in completely definitive ways. External conditions constrain the types of evaluation you can perform.

This is not to say that you shouldn't design your programs to incorporate the best possible evaluation strategies, but what is possible has limits. Effective planning can maximise the research potential of your study while still operating within your budget and available time. If you don't know what is possible, it is important to collaborate with others. Other equity practitioners will be able to advise whether a particular strategy will work "in the field," or whether there are confounding factors with that type of program that have been seen before. You can get advice to avoid having to "reinvent the wheel" regarding your evaluation program and have a colleague peer review your plan. Some academics may also be eager to collaborate as part of their service to the university or for research possibilities, although making sure both parties are clear on what the collaboration will involve and what its purpose is may avoid any confusion or conflict later.

Collaboration with other professional and academic staff is a win-win situation, and looking for opportunities to do so could be useful for you.

A common method for determining program outcomes is to use an experimental design approach. Here, participants' responses are compared with a control group (a group of similar participants who didn't undergo the program, or at least, the aspect of the program being evaluated) to try to confirm that any differences in outcomes are caused by the program and not other influences.

Technically, participants should be randomly assigned to the control or the experimental group for the design to be considered a formal experiment. As this is often not possible in equity programs (because participants volunteer to participate, or because it may be unethical to withhold support from students for experimental reasons), these evaluations are often "experiment-like" rather than experimental.

It can sometimes be very difficult to find an appropriate control (particularly in some widening participation programs), so in well-designed programs, several different control groups may be used for comparison. An alternative is to compare the participants' responses before and after the program, although there are confounding factors here too, such as other factors influencing participants during the course of the program. Being upfront about these limitations is both desirable and to be expected; no study is perfect. Sources for further reading on experimental design is provided in the Further Resources section of this guide.

Making judgements

The ultimate aim of evaluation is to measure the impact of an equity program, and to decide whether its impact has justified its costs. This judgement may not necessarily be made by the evaluator; it may be decided by managers or senior staff. However, this judgement should be based on the criteria and standards collected as part of the logic of evaluation, bearing in mind the constraints or limitations of the methodology used.

PLANNING FOR EVALUATION

It is never too early to develop an evaluation plan for a project.

All equity programs can benefit from having an evaluation plan.

The aims and objectives of the program determine what the evaluation questions should be.

"Smart" objectives are specific; measurable; achievable; realistic; and timed.

A clear evaluation plan will help you gather data methodically, at the right time, and without wasting time gathering unnecessary data.

Indicators help you keep track of your project aims and the likely impact of your program. They have no value in and of themselves.

The types of data you gather and the evaluation questions asked will change over the life of the program.

It is never too early to develop an evaluation plan for a project. All equity projects can benefit from having an evaluation plan from their very early stages. Ideally, this would be developed alongside the original proposal, but this may not always be possible due to management decisions and practitioners being moved from other programs, for example. A good evaluation plan will make collecting and reviewing data much easier for you, as well as increasing the quality and rigor of your findings. Many institutions will have a template for practitioners to use, but a series of prompts have also been provided at the end of this guide to help you. Having a detailed evaluation plan is simply good practice.

A vital part of planning a program is budgeting the time it will take to evaluate it. Although evaluation shouldn't be onerous, it does take time to plan, collect and analyse data, and reflect on the findings. In some cases, up to 25% of a practitioner's time on a project may be spent in evaluation. A good evaluation plan can make this as efficient as possible, but you must be realistic about allocating time in your program proposal.

As noted above, having clear, unambiguous program aims is essential. An aim, in this context, is a broad statement of desired outcomes from the program.

For example, the broad aim of an undergraduate support program might be to increase students' academic capacity. A more focused aim might be increasing students' academic literacy in first year.

Objectives are subsidiary to aims, and describe the steps needed to accomplish those aims. They must be highly focused and feasible, and emphasise how the program aims will be accomplished. An objective does more than describe what you will do or deliver (which is an activity or output of the program instead). A frequently used mnemonic is to create "smart" objectives:

Specific
Measurable
Achievable
Realistic, and
Timed.

These aims and objectives will be used, throughout the life of the program, to determine what the evaluation questions should be; what methods are used to gather data to measure the program's impact; methodological considerations such as recruiting control groups, potential confounding variables and instrument design; and so on.

Evaluation questions, objects of analysis and methods change based on the maturity of the program, so a good evaluation plan anticipates and accommodates these changes. A list of typical questions for evaluating programs in the planning, monitoring and summative phases is provided in the next major section.

Objectives are often tied to measures and indicators. These are typically developed based on standards (for example, to develop academic literacy to a particular level), and can be used to monitor changes over time, to compare with standards, or to compare program outcomes at different sites. Although the two terms are often used interchangeably, indicators are used to monitor a program's progress, while measures are used to sum up its outcomes.

Establish Criteria

Aims

↓

Objectives

↓

Evaluation questions

↓

Methodology

Construct Standards   Methods → Controls

Measure performance   Indicators/measures   Standards

↓

Analysis

↓

Make judgement   Judgement
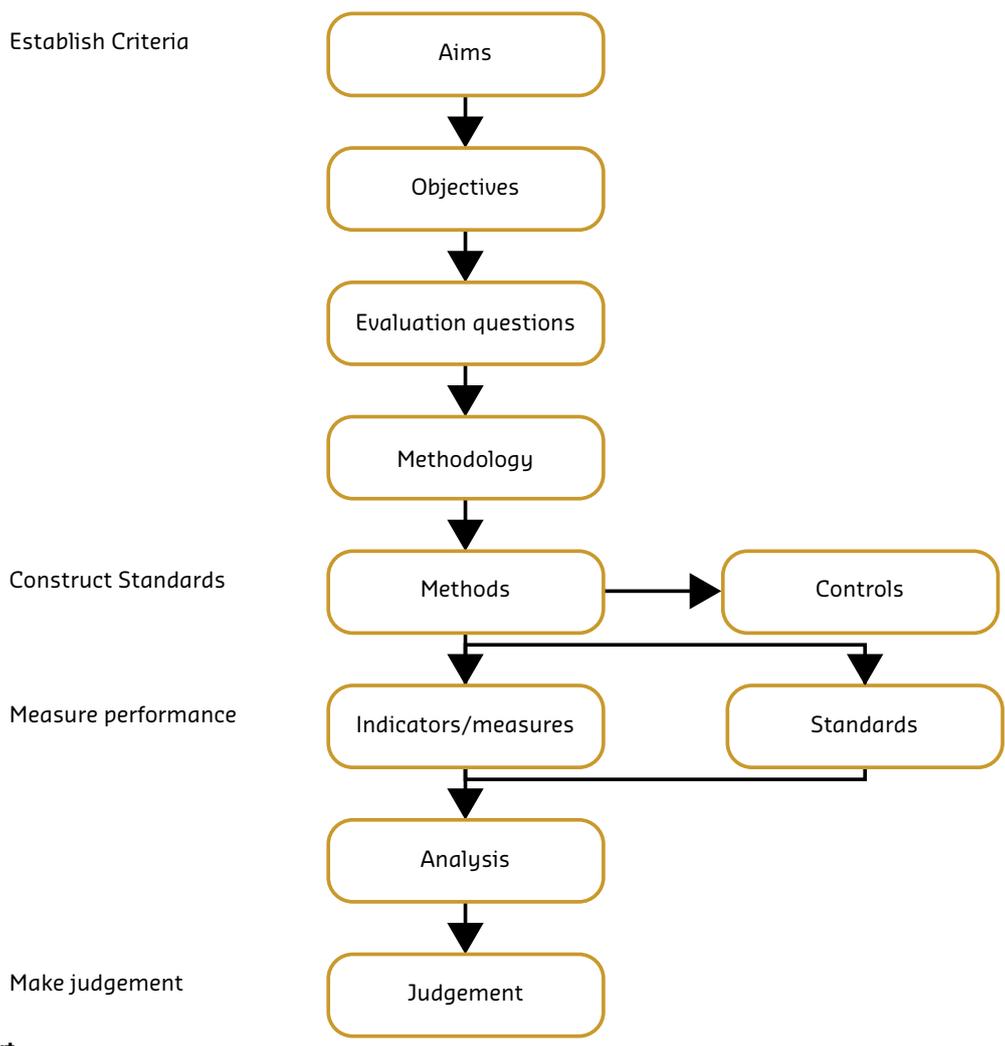
An evaluation flowchart.

It is important to remember that indicators have no value in and of themselves; they are used to inform judgements about the effectiveness of the program. Too much focus on indicators, or the wrong indicators, can lead to a misaligned program, in which the true aims of the project are ignored in favour of the pursuit of particular indicators. In developing indicators, it is therefore important to ensure that the indicators used are the best available, and accurately measure progress towards the project aims. These considerations are detailed in the sections on monitoring and evaluating during the summative phase, but establishing the right indicators is an important part of the evaluation plan.

It should also be noted that students are often involved in several programs at once. Several outreach programs may operate in the one school, for example, or a student already at university may participate in a library skills seminar, visit the academic writing unit, and have weekly tutorials with an undergraduate mentor. In these cases, it can be particularly important to have a clear evaluation plan to be able to articulate the effects of each program. As noted above, if the programs have different aims or objectives,

your evaluation questions, or the questions you ask of participants, should focus on the unique features of your program. Finding the right control group may be effective (although it may be challenging, depending on the type of program). You may choose to evaluate all of the programs together, as a set, rather than evaluating each separately (if funding arrangements allow), which allows you to pool your resources with your collaborators.

Again, planning for best practice within your operational constraints, and consulting with your colleagues and manager can be useful to ensure you're not aiming too high, or too low.

Finally, being able to accommodate obvious contingencies might be a useful addition to your plan. Will it still be possible to do a summative evaluation if your funding is cut part way through the program? What will you do if the parameters have to change due to staffing issues? How will that affect your operation and evaluation? How can you make your plans flexible enough to deal with these outside influences? Planning for contingencies like these will make your project more robust, and make dealing with these risks less stressful.

Typical questions that are useful to consider during the development of an evaluation plan include:

Who are the participants in this program?

Who are the stakeholders for this program? Are there other stakeholders involved in the evaluation? (Eg: Colleagues interested in your findings)

How will these stakeholders and participants be engaged during the implementation of the evaluation plan?

What ethical requirements are there in evaluating each stage of the program? How long will obtaining the necessary permissions (from the institution, partner organisation, government, participants, etc.) take?

What methodological requirements are there in evaluating each stage of the program?

What methods will be used to evaluate the program? Why were these methods chosen?

How will the evaluation findings be used? What are the reporting requirements?

What are the critical time points for the program? How will these affect program monitoring and evaluation?

EVALUATING A PROJECT PLAN

Evaluation during the planning phase will help you strengthen a project by clarifying its aims and improving the alignment between its aims and design, or with institutional targets.

It may also establish criteria and standards for evaluation during later phases, which will make later evaluation easier for you.

Evaluation during planning may involve a needs analysis (whether the program is needed) and/or a plausibility analysis (whether the program will plausibly achieve its aims).

Sources of evidence include literature reviews, benchmarking similar programs, statistical analysis, document analysis, concept mapping, stakeholder opinion, and program logic models.

When do you evaluate an equity program?

All programs undergo some form of evaluation during the approval process, but formal, detailed evaluation can considerably strengthen a program. This section details a framework for evaluating programs during planning: both during development and before a major revision of existing programs.

Unfortunately, equity practitioners may have to "hit the ground running" if they're put in charge of a pre-existing project, or in response to new grants or decisions from their managers. This may prevent a full, detailed examination of the project. Sometime the previous program's evaluation plan might have been inadequate, or non-existent. The most important thing is to be clear about the aims and objectives of the program and how it operates. From that, you can start to build an evaluation plan into the pre-existing program as you have time, and create a detailed plan for evaluation during its later stages.

Evaluation during these early stages fundamentally functions as a needs analysis or a plausibility analysis of the program, depending on when in the planning process the evaluation occurs.

A needs analysis, as the name suggests, seeks to determine whether the program is needed. This is typically an evaluation of the program's aims, and often involves articulating whether it is aligned with and will advance institutional targets. A plausibility analysis asks whether the design of the program will plausibly achieve its aims. It is therefore an evaluation of the alignment between the program's aims and its method of operation. Clearly, the two analyses are complementary rather than mutually exclusive, and are frequently performed together, but they are evaluating different things. At the strategic level, needs analyses are more common, while both forms of evaluation can apply at the project level.

As with any evaluation, the criteria for evaluation are determined by the intended aims of the program; standards are established and evidence collected according to the logic of evaluation (described above).

Resources used for evaluation during this phase include reviews of the literature (particularly descriptions of previous studies) and benchmarking to establish standards, identify best practice and support the needs and plausibility analyses.

Other sources of evidence used during the planning phase might include:

- Statistical analysis; e.g.
  - ◇ Demographic information about regions of interest;
  - ◇ Enrolment information;
  - ◇ Academic results (completion, success, and retention rates).

- Document analysis; e.g.
  - ◇ Government policy documents;
  - ◇ University policy or strategy documents;
  - ◇ Grant conditions or agreements.

- Concept mapping.

Typical questions asked during the planning phase include:

Which equity group will be targeted by the program?

How do the aims and objectives of this program align with broader strategic goals (institutional, government, etc.)?

What underlying causes contribute to the under-representation from these groups?

How can these causes be addressed?

Will this program plausibly address those discrepancies?

- Stakeholder opinion (e.g. students, parents, service providers, equity practitioners or managers, academics or other stakeholders); e.g.
  ◊ Individual interviews/one-on-one discussion;
  ◊ Focus groups;
  ◊ Surveys;
  ◊ Delphi technique (experts individually propose a range of needs or designs, and then collectively discuss or rank the proposals; typically, several iterations are used to reach consensus).

- Program logic models (project design is reviewed by articulating causal links between means or activities, and ends or outcomes; how and why the links exist is often also explained).

While evaluation during planning needn't use all of these forms of evidence, using multiple strands of evidence is most likely to ensure your program is likely to succeed.

MONITORING

Evaluation during the operation of the program (monitoring) will help you make sure the program is achieving its aims, and is being delivered effectively and consistently across multiple sites or practitioners.

Effectiveness, efficiency and delivery processes can all be monitored.

Monitoring at regular intervals through operation can help you keep track of the program and identify problems or new opportunities early.

Monitoring program outcomes typically involves tracking qualitative and quantitative indicators via a mixed methods analysis.

Monitoring program processes typically involves action research methods, stakeholder reviews and examining program logic models.

This section provides advice on monitoring, or evaluating programs during operation. This advice applies whether you are managing a pilot, or a full program. Monitoring is particularly important during pilot programs and the early stages of full programs, to ensure that the program is getting off to a good start, to detect problems early, or (for pilot programs) to provide a solid foundation for a future program.

During their operation, three aspects of an equity program can be evaluated:

1. Program effectiveness (impact);
2. Efficiency (cost-effectiveness or cost/benefit analyses); and
3. Program processes (delivery).

Program effectiveness and efficiency together make up the program's outcomes.

Both processes and outcomes need to meet appropriate standards for a program to have the desired impact. For that reason, it is important that all three aspects of the program are monitored. You don't want to deliver a less effective program at a lower cost per head, or carefully develop a program that runs over several sites and then not ensure that all practitioners deliver it consistently.

Monitoring is used to ensure that delivery processes and outcomes are on track, and to develop systems to enable that monitoring or to fine-tune the program over its life. Monitoring is an on-going aspect of most programs, except perhaps the very short ones. Some programs lend themselves to small, continuous evaluations, whereas large evaluations at particular intervals may be more appropriate for other programs. You could also monitor specific aspects of the program at different intervals, or one aspect of a program may demand in-depth review if it is performing much better or worse than expected.

Gathering and assessing evidence is the primary activity during the monitoring phase. Lots of evidence is created as part of the normal business of the program, ranging from informal observations of participant engagement to formal outputs such as survey data, academic results, and costs. The important step is to collate and interpret it at appropriate intervals.

Appropriate methods for collecting evidence for evaluating processes include:

- Critical friends and stakeholder reviews, where an independent observer provides a critique based on their expertise;
- Action research.

Reviews by critical friends or stakeholders rely on an independent observer to provide a critique based on their expertise. In the case of a critical friend, it is typically a colleague with experience in monitoring similar programs, although anyone you trust to provide an independent, expert critique could be appropriate. In a stakeholder review, it is another stakeholder in the program who provides the critique, rather than an independent colleague, but the process is fundamentally the same. In both cases, the goal is to receive formative feedback to ensure the delivery of the program is progressing appropriately and consistently.

Action research is a methodology where practitioners conduct systematic inquiries to help them improve their own practices. It is commonly used in healthcare, but it's also particularly good for equity practitioners because it's focused on creating solutions to practical problems. It is also led by practitioners themselves, which empowers them within their programs.

Typical questions asked during the monitoring phase include:

Is the program reaching its target population?

Is it meeting its objectives?

Is it meeting its intended standards?

What is the trend over time? Is it changing?

Is it being implemented in the ways specified in the original plan?

Is the delivery of the program working as intended?

In action research, practitioners undergo cycles of reflecting on aspects of the program, proposing and instituting plans of action, and then observing and reflecting further. In some ways, it's a more formalised approach to the sorts of activities equity practitioners do as part of the everyday management and improvement of their projects.

More information for action research is provided in the Further Resources section.

Both methods typically involve close interaction with those involved in program delivery. Thus, the methods of collecting evidence described in the previous section, such as interviews, concept mapping, and reviewing program logic diagrams, are important aspects of both techniques.

In contrast, program outcomes are typically monitored via mixed methods approaches. Developing indicators is discussed in more detail in the Planning for Evaluation section of this guide.

.

However, it must be emphasised that most projects require a combination of qualitative and quantitative indicators for their outcomes to be monitored appropriately. These indicators must also be appropriately aligned with the aims of the program; it is important that the indicators used measure outcomes, not outputs (activities). Once again, impact is the primary goal of equity programs. Indicators that measure the number of workshops given or the total participants do not measure impact, and are of little rigorous value in evaluating impact. They may, however, be valuable in measuring efficiency.

Finally, during monitoring, you should try to be aware of unexpected outcomes (positive or negative), which may start to become apparent during the operation of the program. You may need to develop new indicators to effectively capture these new outcomes.
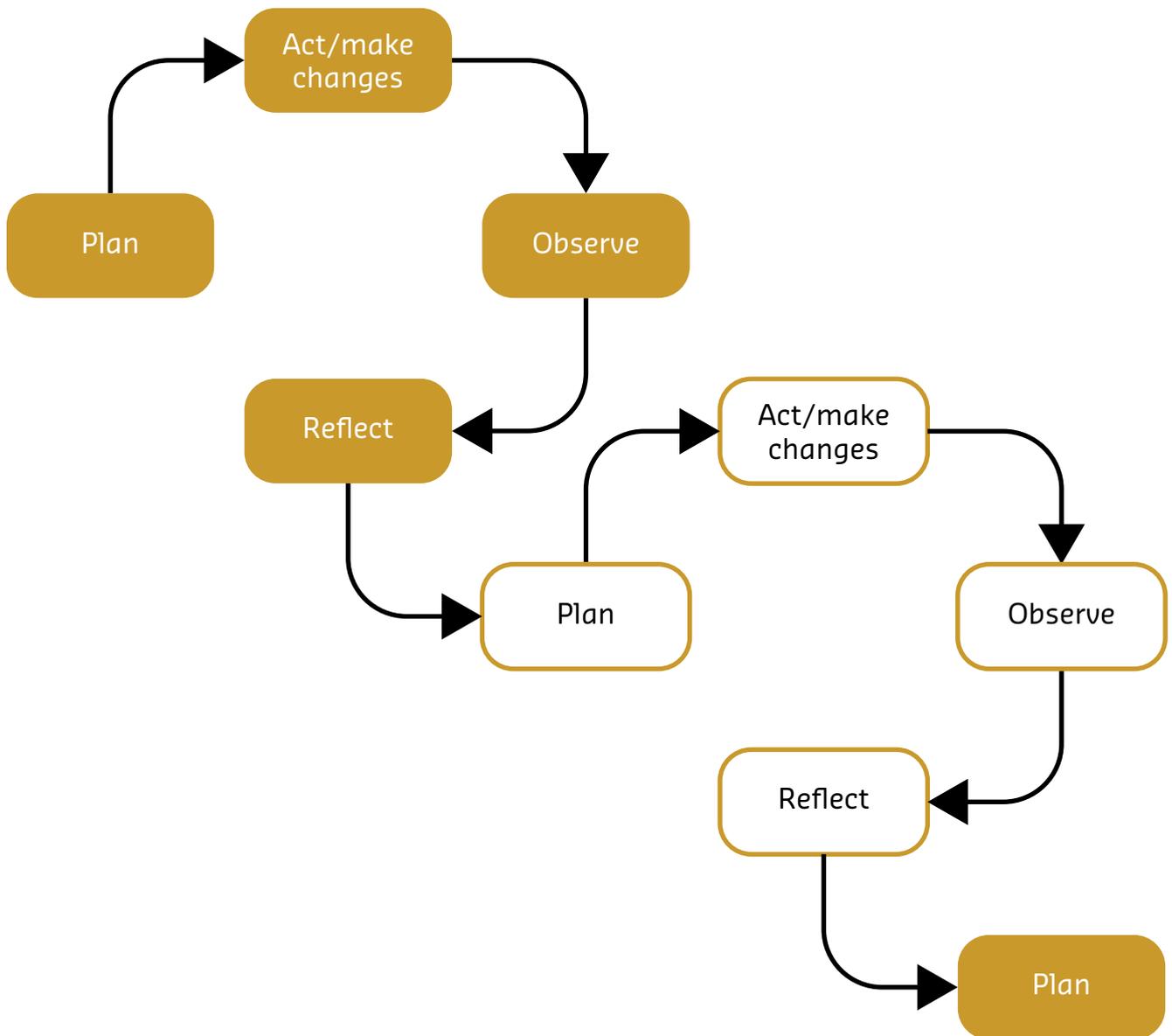
The action research process.

EVALUATION DURING THE SUMMATIVE PHASE

Evaluation during the summative phase of an equity program seeks to establish which aspects of the program have worked and why.

Evaluation during the summative phase will help you work out which aspects of the program worked and why, and what its impact has been.

Summative evaluation is usually based on indicators developed during the planning and monitoring phases.

Unintended and intended outcomes are both useful to examine.

An experimental or experiment-like design is often used to establish the impact of the program. Several different control groups, or pre- and post-program responses from participants, may be used for comparison. Causal links may still be very difficult to establish, however.

Evaluation during the summative phase of an equity program seeks to establish which aspects of the program have worked and why. Many practitioners will be familiar with this type of evaluation; it is what is commonly thought of by the term "evaluation." This is the type of evaluation most commonly published in equity literature, and most funding bodies require this type of evaluation.

Evaluation during the summative phase is typically retrospective. It is performed when programs have reached maturity — not necessarily at the end of the program's life, but when all the bugs have been ironed out of the program's operations, and at least one full cycle of its activities have taken place.

Clear, measurable aims for the program should have been established long before this phase, and are essential to enable summative evaluation to occur. If the program has a robust and valid evaluation plan from the planning and monitoring phases, this should be much easier, as evaluation and collection of evidence should be built into the everyday practices of equity practitioners. However, the indicators used to monitor a program may not be the same as the measures used to sum up its final outcomes (which is why the terms "indicators" and "measures" are used). In many programs, the two will be identical, but in others, there may be logistical or methodological reasons why one type of evidence is collected to monitor the progress of the program, and another used to evaluate its final outcomes.

For example, an academic development program may monitor student progress via attendance, class performance and engagement, but measure overall impact in terms of their final marks, ATAR scores, or ability to enter the major, degree or employment of their choosing.

As with monitoring, it can be equally useful to look at the unintended consequences of the program as well. To measure these effects, mixed methods approaches that provide an opportunity for open-ended responses (for example, Further Comments sections on participant surveys) are typical.

Typical questions asked during the summative phase include:

What are the range and extent of the program's outcomes?

Have the program's aims been achieved?

Was the program implemented as planned?

How did the implementation affect the outcomes of the program?

How did differences in implementation affect outcomes at different stages or sites of the program?

What were the unintended outcomes of the program?

Was the program cost effective? What were its benefits relative to its costs?

The answers to these questions inform decisions concerning the replication or extension of the program.

Translation of a project's aims into valid measures of outcomes is a major methodological issue. Given the complexity of determining causal links in these types of program, it is still often an inexact science. At the very least, your evidence of impact should have face validity — that is, a colleague or critical friend should be able to understand the link between the data collected and the program aims.

FURTHER RESOURCES

Mixed methods research
Major textbooks in the field include:

Bergman, M. M (Ed.) (2008). Advances in Mixed Methods Research. London: SAGE Publications Ltd.

Creswell, J. W., & Plano Clark, V. L. (2007). Designing and conducting mixed methods research. London: SAGE Publications Ltd.

Teddlie, C., & Tashakkori, A. (2009). Foundations of mixed methods research. London: SAGE Publishing Ltd.

The following annotated bibliographies provide detailed references for quantitative and qualitative research methods:

McMillan, J. H., Mohn, R. S., & Hammack, M. V. (2013). Quantitative research designs in educational research. Oxford: Oxford University Press. http://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0113.xml

Onwuegbuzie, A. J., & Denham, M. (2014). Qualitative data analysis techniques. Oxford: Oxford University Press. http://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0078.xml

Action research
Major textbooks in the field include:

Greenwood, D. J., & Levin, M. (2006). Introduction to action research: Social research for social change. London: SAGE publications Ltd.

Koshy, V. (2009). Action research for improving educational practice: A step-by-step guide. London: Sage Publications Ltd.

Stringer, E. T. (2004). Action research in education. Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.

The following annotated bibliographies provide detailed references for action research methods. Further information can also be found in Onwugbuzie and Denham (2014):

Maruyama, G., & Van Boekel, M. (2014). Action research. Oxford: Oxford University Press. http://www.oxfordbibliographies.com/view/document/obo-9780199828340/obo-9780199828340-0149.xml

Program evaluation
Major textbooks in the field include:

Owen, J. M., & Rogers, P. J. (1999). Evaluation Fundamentals. London: SAGE Publications Ltd.

Royse, D., Thyer, B., & Padgett, D. (2009). Program evaluation: An introduction. Boston: Cengage Learning.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). Evaluation theory, models, and applications. New York: Jossey-Bass.

Other useful references may include:

Cooksy, L. J., Gill, P., & Kelly, P. A. (2001). The program logic model as an integrative framework for a multimethod evaluation. Evaluation and program planning, 24 (2), 119-128.

Frechtling, J. A. (2007). Logic modeling methods in program evaluation. New York: Jossey-Bass.

Gale, T., Sellar, S., Parker, S., Hattam, R., Comber, B., Tranter, D., & Bills, D. (2010). Interventions early in school as a means to improve higher education outcomes for disadvantaged (particularly low SES) students. Australia: DEEWR.

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. Practical Assessment, Research & Evaluation, 12 (10), 1-8.
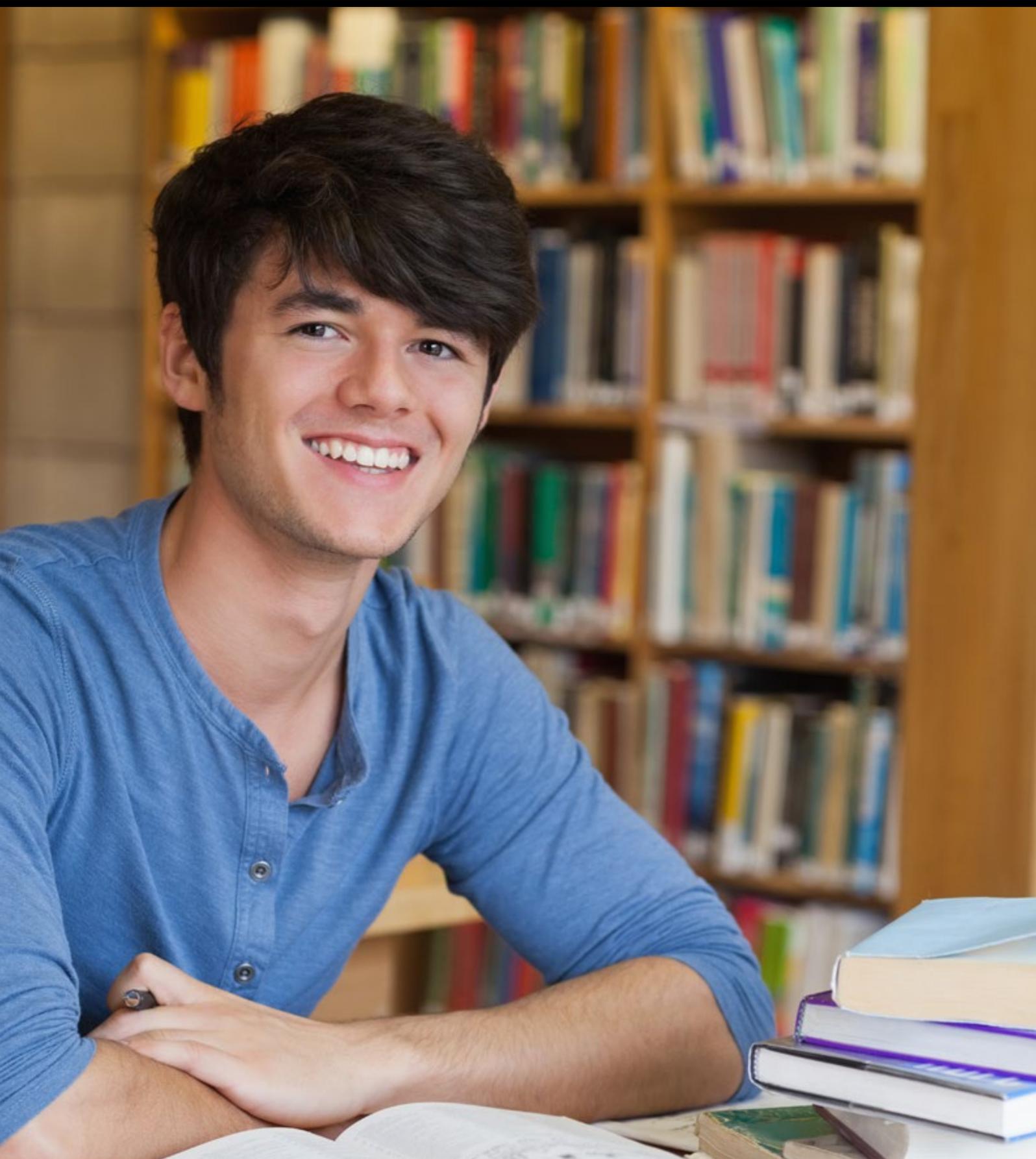
REFERENCES

Cave, M. (1997). The use of performance indicators in higher education: The challenge of the quality movement: Jessica Kingsley Publishers.

Cooksy, L. J., Gill, P., & Kelly, P. A. (2001). The program logic model as an integrative framework for a multimethod evaluation. Evaluation and program planning, 24 (2), 119-128.

Creswell, J., Plano Clark, V. L., & Garrett, A. L. (2008). Methodological Issues in Conducting Mixed Methods Research Designs. In M. M. Bergman (Ed.), Advances in Mixed Methods Research. (pp. 66-84). London: SAGE Publications Ltd.

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals an objectives. Management Review, 70 (11), 35-36.

Fetterman, D. M. (1994). Empowerment evaluation. Evaluation practice, 15, 1-15.

Fournier, D. M. (1995). Establishing evaluative conclusions: A distinction between general and working logic. New directions for program evaluation, 68, 15-32.

Frechtling, J. A. (2007). Logic modeling methods in program evaluation (Vol. 5). New York: Jossey-Bass.

Gale, T., Sellar, S., Parker, S., Hattam, R., Comber, B., Tranter, D., & Bills, D. (2010). Interventions early in school as a means to improve higher education outcomes for disadvantaged (particularly low SES) students. South Australia.

Gale, T., Sellar, S., Parker, S., Hattam, R., Comber, B., Tranter, D., & Bills, D. (2010). Interventions early in school as a means to improve higher education outcomes for disadvantaged (particularly low SES) students. Australia: DEEWR.

Greenwood, D. J., & Levin, M. (2006). Introduction to action research: Social research for social change: SAGE publications.

Hatt, S. (2007). Measuring success: A guide to evaluation for Aimhigher. UK: Aimhigher evidence good practice group.

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. Practical Assessment, Research & Evaluation, 12 (10), 1-8.

Koshy, V. (2009). Action research for improving educational practice: A step-by-step guide: Sage.

Maruyama, G., & Van Boekel, M. (2014). Action research. from http://www.oxfordbibliographies.com/view/document/obo-9780199828340/obo-9780199828340-0149.xml

McMillan, J. H., Mohn, R. S., & Hammack, M. V. (2013). Quantitative research designs in educational research. Retrieved 15 January, 2015, from http://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0113.xml

Meyer, J. (2000). Qualitative research in health care: Using qualitative methods in health related action research. BMJ: British Medical Journal, 320 (7228), 178.

National Centre for Student Equity in Higher Education (NCSEHE). (2011). Progress Report: 2008-2011. Adelaide: NCSEHE.

Neuman, W. L., & Neuman, W. L. (2006). Social research methods: Qualitative and quantitative approaches.

Onwuegbuzie, A. J., & Denham, M. (2014). Qualitative data analysis techniques. Retrieved 15 January, 2015, from http://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0078.xml

Owen, J. M., & Rogers, P. J. (1999). Evaluation Fundamentals. London: SAGE Publications.

Roth, J. (1990). Needs and the needs assessment process. Evaluation practice, 11, 39-44.

Royse, D., Thyer, B., & Padgett, D. (2009). Program evaluation: An introduction. Boston: Cengage Learning.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). Evaluation theory, models, and applications. New York: Jossey-Bass.

Teddlie, C., & Tashakkori, A. (2009). Foundations of mixed
  methods research. London: SAGE Publishing Inc.

Witkin, B. R., & Altschuld, J. W. (1995). Planning and
  conducting needs assessments. California: Sage.

PROMPTS AND QUESTIONS FOR PLANNING

The following pages contain a program assessment framework for you to review and reflect upon the role of evaluation when designing new equity programs or considering existing ones.

What are the general aims of your program?

Informing aspirations towards higher education?

Developing academic capacity and providing academic support?

Establishing inclusive processes?

Supporting students in dealing with broader issues outside their academic lives?

Other?

What are the specific objectives of your program? (What specific areas of change do you intend?)

Which of these aims and objectives are of highest priority?

How do these aims and objectives fit within institutional and government guidelines/goals?

What is the current situation which this program will try to improve? How is the situation measured at the moment?

What specific targets do you intend to achieve?

Why is this situation the way it is? Why do (e.g.) these problems/discrepancies exist?

What examples exist for benchmarking to ensure these targets are realistic?

What specific activities will be part of this project?

What methodology will you use?

How will those activities address the discrepancies/ reach those targets? (Are they plausible solutions)?

What is your strategy for collecting and analysing data?

What sources of evidence (indicators) can be used to monitor progress?

How will that data be collected?

What data collection techniques will you use?

What data analysis techniques will you use?

Who are the participants? (Are the right people being identified to ensure the program aims are met?)

Do you need ethics or other permissions to collect data?

What control groups are being used?

How often will monitoring occur to ensure progress? (Will different aspects of the program be monitored at different intervals?)

How can the program be made more effective or efficient?

What are the costs associated?

What are your reporting requirements? (How often? What audience will be reading it?)

Has the program met the original targets? Why?

How could the program be improved?

CONTACT

ncsehe.edu.au