



**MINISTRY OF SOCIAL
DEVELOPMENT**

TE MANATŪ WHAKAHIATO ORA

25 MAY 2020

Dear [REDACTED]

On 8 May 2020, you emailed the Ministry of Social Development (the Ministry) requesting, under the Official Information Act 1982, the following information:

- *A copy of "What have we learned and where to next? A review of evaluations of employment programmes and interventions, 1992-97" written by Diane Anderson.*

Please find a copy of the document attached.

The principles and purposes of the Official Information Act 1982 under which you made your request are:

- to create greater openness and transparency about the plans, work and activities of the Government,
- to increase the ability of the public to participate in the making and administration of our laws and policies and
- to lead to greater accountability in the conduct of public affairs.

This Ministry fully supports those principles and purposes. The Ministry therefore intends to make the information contained in this letter and any attached documents available to the wider public. The Ministry will do this by publishing this letter and attachments on the Ministry of Social Development's website. Your personal details will be deleted, and the Ministry will not publish any information that would identify you as the person who requested the information.

If you wish to discuss this response with us, please feel free to contact OIA_Requests@msd.govt.nz.

If you are not satisfied with this response, you have the right to seek an investigation and review by the Ombudsman. Information about how to make a complaint is available at www.ombudsman.parliament.nz or 0800 802 602.

Yours sincerely

Kate Satterthwaite
General Manager, Ministerial and Executive Services

Department of Labour
Evaluations

**What have we learned and where to next?
A review of evaluations of employment programmes
and interventions, 1992-97.**

Diane Anderson

November 1998

Labour Market Policy Group

DEPARTMENT OF
|L|A|B|O|R|
T E T A R I M A H I

TABLE OF CONTENTS

Table of Contents	2
Executive Summary	4
1. Introduction	8
1.1 Context and background information on organisations undertaking the evaluations	8
2. Key questions and Methodology	11
2.1 Key Questions	11
2.2 Methodology	11
2.2.1 Effectiveness of the programmes	11
2.2.2 Evaluation quality	13
2.2.3 Evaluation use	13
2.2.4 Limitations	14
3. Results	14
3.1 Types of programmes	14
3.2 Target groups	15
3.3 The focus of the evaluations	18
3.3.1 Cost and cost-effectiveness	18
3.3.2 A focus on individual programmes	19
3.4 Outcomes for participants	20
3.4.1 Do programmes place people into employment?	20
3.4.2 Training outcomes	21
3.4.3 Attitudinal outcomes	21
3.4.4 Difficulties in comparing outcomes	22
3.5 Operational strengths and weaknesses	24
3.5.1 Recruitment and selection of programme participants	24
3.5.2 Meeting the objectives or the intent of the programme	25
3.5.3 Programme monitoring	28
3.5.4 Case management and post-placement support	28
3.6 Evaluation Quality	30
3.6.1 Summary	30
3.6.2 The correspondence between the evaluation objectives and the rest of the report	30
3.6.3 The existence and quality of the key findings	32
3.6.4 Quality of conclusions	33
3.6.5 Evaluation report recommendations	34
4. Conclusion	36
5. Recommendations	38
References	40
Appendix 1: criteria	41
Appendix 2: evaluation utilisation Questions	46

<i>Appendix 3: Programme information</i>	47
<i>Appendix 4: The focus of the evaluations</i>	52
<i>Appendix 5: paper on evaluation methods</i>	54
<i>Appendix 6: Employment outcomes</i>	64
<i>Appendix 7: Training outcomes</i>	68
<i>Appendix 8: Evaluation quality</i>	70
<i>Appendix 9: Methods used in evaluations</i>	71
<i>Appendix 10: Utilisation of selected evaluations</i>	72

EXECUTIVE SUMMARY

Background

The Labour Market Policy Group (LMPG) of the Department of Labour has summarised evaluations of employment programmes from five agencies. The review focused on 35 evaluations conducted since 1991, and examined employment issues, operational trends, and evaluation quality. Results of the analysis are to inform development of a new department that combines income support and employment services. To date, this department and other government agencies have largely focused on evaluating individual employment and training programmes, and have tended not to evaluate broad employment strategies or to aggregate the results of individual evaluations. While evaluations of individual programmes will continue, there is a need to periodically examine common trends and differences across the evaluations. This evaluation presents the results of one such re-examination.

Key Results

As of 1 October 1998, Work and Income New Zealand (WINZ) will be focused on returning people to work using regional flexibility as a way to tailor programmes to the needs of the local labour markets. As a result, regional commissioners will need clear specifications on what each programme or intervention is trying to achieve and for which demographic or beneficiary groups. Regional commissioners will need to know how much programmes cost and what operational issues need to be resolved. The need for high quality evaluation to determine what works and what does not will be absolutely essential.

Twenty-two of the 35 evaluations identified employment outcomes for participants. The highest percentages gaining employment were Job Plus and Job Skills participants, with 66% and 65% respectively. These were both wage subsidy programmes. The Compass (1995) and the Hikoi ki Pae-rangi programmes had the highest percentage of participants entering further training. These were both life-skills or motivational programmes.

However, we could not tell, with any certainty, what programmes were most effective at improving the employment prospects of which groups and under what circumstances. This was partly because agencies did not use consistent outcome measures and the emphasis on particular outcomes differed across the agencies. For example, some programmes focused less on participants obtaining full-time employment than others. In addition, these evaluations were generally developed to meet the immediate information needs of internal and external stakeholders and not designed to address longer-term information needs. Evaluators were not required to evaluate suites or combinations of programmes. However, if Regional Commissioners are to reliably assess which programmes work, in what combinations and for which target groups the following will need to be developed:

- consistent outcome measures, including those for employment, training and attitudinal change as well as those for ethnic and demographic groups,
- predefined success criteria for programmes, and
- robust measures of the cost or cost-effectiveness of programmes.

Finally, assumptions about the benefits of particular types of programmes for particular groups will need to be tested. For example, the assumption that life skills or motivational programmes benefit Maori, Pacific Island peoples, youth and women is untested.

Several evaluations identified factors that could put the accomplishment of the programme aims and objectives at risk. These factors included the following:

- job seekers not being appropriately matched to programmes and interventions,
- the programmes failing to meet the programme objectives or the intent of the programme,
- inadequate monitoring of the programmes and job seeker progress, or
- inadequate case management or post placement support.

It was more difficult to identify the key strengths across the evaluations because the evaluations tended to focus on the operational weaknesses. However, evaluations also provide an opportunity to identify programme strengths. Such information can be used to develop 'best practices'.

The quality of the evaluations differed within and between agencies. To improve future evaluations, several issues surrounding the quality of the evaluations will need to be addressed. In particular, future evaluations will need to:

- clearly link evaluation objectives to the rest of the report,
- appropriately identify and support the key findings,
- provide conclusions with a reportable message,
- state in the Terms of Reference whether or not recommendations are to be included in the report and, if so, provide recommendations of sufficient quality.

Where to from here?

In looking across the agencies involved steps are being taken to increase the robustness of the evaluations and to assess them in a more strategic manner.

- NZES has taken steps to improve the robustness of its evaluations. For example, the Community Task Force (CTF) evaluation is employing the use of randomised control and treatment groups to measure outcomes for participants. While this has raised some ethical issues, it will provide more robust outcome analysis.
- NZES has also recognised the need to develop more robust and consistent measures of attitudinal change. For example the evaluation of Residential Motivational Training (currently underway) is attempting to measure attitudinal change of participants on a before and after intervention basis. In addition the evaluation is tracking cohorts of participants to measure the robustness of attitudinal change over time and across different providers.

- CEG has developed an evaluation strategy. The Mahi A Iwi and Pacific Island Peoples evaluations were part of a longer term strategy that started with exploratory case studies. CEG's next large evaluation project is an evaluation of the women's strategy. This evaluation will build on the understandings of CEG's work in the community gained from the previous evaluations and move towards measuring the outcomes of our involvement with groups. In another project CEG will attempt to get a clearer idea of its impact on communities, by tracking a sample of groups for three years to assess the short, medium and long term outcomes at the individual, group, and community levels. A third study will examine community access to CEG's services.

Nevertheless, further steps need to be taken to address issues raised in this meta-evaluation. The meta-evaluation recommends the development of an evaluation strategy which identifies priorities and timelines for selected strategic and operational evaluations along with the roles and project management responsibilities across the relevant agencies. While the strategy would cover a period of three years it could be updated annually. The strategy would also include ways to improve evaluation capacity. To this end, the meta-evaluation makes recommendations aimed at developing consistent definitions and measures of outcomes for participants in employment and training programmes. The development of appropriate impact measures for labour market and income support programmes, and criteria and guidelines for conducting evaluations are also recommended. The guidelines could build on existing guidelines such as those developed by LMPG and NZES. Finally, the meta-evaluation recommends that processes be put in place for tracking the recommendations or findings arising from evaluations. The aim is to encourage the creation of actionable recommendations where appropriate, and greater interaction between policy and evaluation units in developing recommendations and/or incorporating evaluation findings into future policy decisions.

Recommendations

It is recommended that:

1. DOL and SPA in consultation with WINZ and other relevant agencies work together to develop an evaluation strategy, dealing with employment and labour market issues over the next three years.
 - a) It is recommended that this strategy identify
 - a list of strategic and operational evaluations;
 - priorities and timelines;
 - the roles and project management responsibilities across the relevant agencies; and
 - ways to improve evaluation capacity.
2. The evaluation strategy includes an evaluation of which types or combinations of programmes and interventions are most effective for which demographic or beneficiary groups.
3. DOL and SPA in consultation with WINZ and other relevant agencies to work together on the following capacity building projects:

- a) develop consistent definitions and measures of outcomes for participants in employment and training programmes prior to finalising Work and Income New Zealand's systems. Consideration will need to be given to how outcomes are measured at the regional level;
- b) investigate and develop appropriate impact measures, including cost-effectiveness, for labour market and income support programmes;
- c) put in place processes for tracking the recommendations or findings arising from their evaluations; and
- d) develop and distribute criteria and guidelines for conducting evaluations.

RELEASED UNDER THE
OFFICIAL INFORMATION ACT

1. INTRODUCTION

The Labour Market Policy Group of the Department of Labour has summarised evaluations of employment programmes from five agencies. The review focused on 35 evaluations conducted since 1991, and examined employment issues, operational trends, and evaluation quality. The evaluations were completed by the New Zealand Employment Service, the Community Employment Group, the Department of Social Welfare, the Education and Training Support Agency and the Ministry of Youth Affairs. Results of the analysis are to inform development of a new department that combines income support and employment services. To date, this department and other government agencies have largely focused on evaluating individual employment and training programmes, and have tended not to evaluate broad employment strategies or to aggregate the results of individual evaluations. While evaluations of individual programmes will continue, there is a need to periodically examine common trends and differences across the evaluations. This evaluation presents the results of one such re-examination.

1.1 CONTEXT AND BACKGROUND INFORMATION ON ORGANISATIONS UNDERTAKING THE EVALUATIONS

In December 1997, Cabinet decided to integrate the New Zealand Employment Service, the Community Employment Group, New Zealand Income Support and the Local Employment co-ordination function into a one-stop shop. As of 1 October 1998, Work and Income New Zealand (WINZ) will be focused on returning people to work using regional flexibility as a way to tailor programmes to the needs of the local labour markets. As a result, Regional Commissioners will need to know what programmes work, in what combinations and for which target groups. They will need to know how much programmes cost and what operational issues need to be resolved. The need for high quality evaluation to determine what works and what does not will be absolutely essential.

This review of employment programme evaluations will assist officials to determine which programmes are most effective under what circumstances. Until recently, it would have been difficult to conduct such a review as there were only a small number of evaluated labour market programmes. However, following the 1993 Employment Task Force there was an increase in the number of labour market programmes, and a requirement that all new employment programmes be evaluated. Consequently, by the end of 1997, a sufficient number of employment programmes had been evaluated to examine common trends and issues across the evaluations.

A New Zealand Employment Service, Department of Labour

The New Zealand Employment Service (NZES) matches job seekers to vacancies; helps job seekers with training and job search skills; provides information and help to job seekers; and offers a variety of programmes to assist disadvantaged into employment. The aim of all NZES services and programmes is to assist people into suitable jobs. There is a particular focus on disadvantaged job seekers (i.e. NZES priority clients).

An evaluation unit within NZES carries out evaluations of NZES programmes or interventions. Most of the evaluations undertaken by NZES over the past five years have been in response to Government requirements, especially the Employment Taskforce. The Government has often determined what programmes will be evaluated and when they will be evaluated.

B Community Employment Group, Department of Labour

The Community Employment Group (CEG) works in partnership with communities and groups within communities to help them create their own opportunities for employment and positive activity leading to self sufficiency. CEG's delivery is at a pace and in a style that responds to a community's culture and development and is tailored to respond to the local circumstances of the community. The four groups targeted by CEG are:

- Maori,
- Pacific Island peoples,
- women, and
- disadvantaged rural and urban communities.

The CEG operational environment means that evaluation needs to be able to recognise the social, economic and employment benefits, risks and factors that relate to community development such as:

- the increase of community involvement through commitment and a sense of ownership that will help the project survive beyond government input,
- a small amount of government resource input (advice and funding) to lever a contribution from the community,
- the promotion of flexible responses to local needs and initiatives,
- fostering a wide variety of responses rather than a replicated response,
- testing innovative (and sometimes higher risk) local solutions to complex problems which often brings unexpected spin-offs,
- fostering the community to be an active partner in the identification and development of their own locally-appropriate employment solutions,
- allowing for results of longer-term developmental projects,
- allowing assistance to be directed to groups who are generally more difficult to reach through the traditional bureaucratic approach, and
- the contribution to social cohesion at the community level.

C Social Policy Agency, Department of Social Welfare

The Social Policy Agency (SPA) provides the government with policy advice on a range of major social and welfare issues, and advises on welfare policy change.

While the primary role of the SPA evaluation team is managing evaluation contracts, it also undertakes some small scale evaluations. Clients for evaluations include Social Policy Agency policy groups, Children Young Persons and their Families Service, New Zealand Income Support Service and Community Funding Agency. Most evaluations are for the purpose of providing information for policy development rather than operations.

D Education and Training Support Agency

The Education and Training Support Agency (ETSA) is an Education Crown entity with a Board appointed by the Minister of Education. The Agency works with employers, Industry Training Organisations, and training providers to raise the skill level of the New Zealand workforce. The Agency is responsible for the following major training initiatives:

- Skill New Zealand,
- Training Opportunities Programme¹, and
- Skill Enhancement.

E Ministry of Youth Affairs

The Ministry of Youth Affairs provides policy advice to the government on issues affecting youth; administers grants to sponsor organisations to operate the Conservation Corps and Youth Services Corps; and takes action to help young people help themselves (i.e. provision of information to young people and those who work with them and the promotion of youth development opportunities).

F Labour Market Policy Group, Department of Labour

The Labour Market Policy Group (LMPG) provides a range of strategic policy advice. This includes:

- developing strategic policy advice and analysis of macro economic and general labour market issues,
- providing strategic advice and analysis of macro economic and general labour market issues,
- monitoring and investigating changes in the labour market.
- facilitating the development of an evaluation strategy for the Department.
- advisory and research services to the National Advisory Council on the Employment of Women (NACEW).

The evaluation unit within LMPG is involved in major evaluation projects² carried out by the Services in the Department of Labour, such as NZES and CEG. LMPG's role is to assist in ensuring the evaluations are of sufficient quality. For example, LMPG provides technical advice on evaluation methods and in 1995 produced a set of evaluation guidelines for the department. LMPG is also represented on the steering groups managing the major evaluations. Steering group members are required to jointly sign off on key milestones (e.g. completion of the scoping, planning, implementation and review phases).

¹ TOP is the Government's largest labour market training intervention, providing around 15,000 places per year.

² Major projects are those that a) have an impact on the KRAs of other services, b) are relevant to major strategic developments in the Department, or c) pertain to the performance contracts between the Chief Executive and other General Managers.

2. KEY QUESTIONS AND METHODOLOGY

2.1 KEY QUESTIONS

The key questions to be addressed in this study are as follows:

1. What are the strengths and weaknesses of different employment programmes or interventions?
2. What employment programmes or interventions appear to be more effective in terms of participant employment and training outcomes, for whom, and why?
3. What are strengths and weaknesses of the evaluations?

This question looks at the quality of the evaluation and answers whether they are robust, supportable, accurate and useful.

2.2 METHODOLOGY

Common trends and issues related to participants' outcomes (i.e. employment, training and attitude), operational strengths and weaknesses, the quality of the evaluations, and the use made of the evaluations, were identified based on the data sources listed below. This process is referred to as a meta-evaluation. A meta-evaluation, sometimes referred to as an evaluation synthesis, is a procedure for systematically comparing the results of multiple evaluations in order to summarise what is already known about a group of policies or programmes. A meta-evaluation may be used on its own or to form the basis for further research.

This report summarises the contents of 35 evaluations of employment programmes undertaken between January 1992 and December 1997 by the following organisations:

- | | |
|---|-----------------|
| • New Zealand Employment Service | 22 evaluations, |
| • Community Employment Group | 3 evaluations, |
| • Department of Social Welfare | 5 evaluations, |
| • Education and Training Support Agency | 3 evaluations, |
| • Ministry of Youth Affairs | 2 evaluations. |

The evaluation reports were the primary source of data for this review. We relied on the information contained in the text of the report and did not replicate the work undertaken in original evaluation.

Other data sources used in the review included follow-up policy documents, and interviews with key staff in NZES, DSW and CEG. This information was primarily used to address issues such as the utilisation of the evaluation. A sample of six evaluations was taken to determine the utilisation of the evaluations (refer to section 2.2.3).

2.2.1 Effectiveness of the programmes

To determine the effectiveness of the programmes information was collected on a) range and cost of programmes, b) the percentage of participants achieving employment and/or training outcomes, and c) the operation of the programmes.

The categorisation listed below recognises that different types of programmes have different expected outcomes. Programmes were classified as:

- wage subsidy
- work experience
- job-related training
- life skills or motivational training
- job search assistance
- community development
- assistance to people in work
- knowledge building
- strategic initiatives.

Refer to Appendix 1 for a definition of each category. In some cases, the programmes that were categorised as assistance to people in work, knowledge building or strategic initiatives have been categorised as “other”.

The following target groups were identified:

- long-term job seekers,
- youth,
- women,
- Maori,
- Pacific Island people,
- sole parents,
- disabled,
- other.

Employment and training outcomes within each programme type were compared in order to determine the effectiveness of the programmes³. Employment outcomes were stated as the percentages of participants achieving the outcome after the intervention, and classified as follows:

- full-time employment (i.e. over 30 hours per week),
- part-time employment (i.e. between 15 and 30 hours per week),
- temporary employment (i.e. a temporary placement lasts between 11 days and 91 days and does not lapse the job seeker off the register),
- short duration employment (i.e. a short-duration placement by NZES is less than or equal to 10 working days and doesn't lapse the job seeker off the register),
- self-employment, and
- unspecified employment (i.e. where the content of the employment outcome is unclear or made up of a combination of the above outcomes).

Within each of the above categories a distinction was also made between subsidised and unsubsidised employment.

³ Not all of the evaluations were required to examine employment and training outcomes. For example, two of the CEG evaluations examined the process of programme delivery and long-term economic development rather than short-term employment outcomes.

Training outcomes were stated as the percentages of participants entering further training after the programme or intervention. No attempt was made to differentiate between training types as in most cases it was not clear as to what constituted further training.

The operational strengths and weaknesses of the programmes were reviewed and trends identified. Trends were identified by counting the number of times an issue was mentioned across different evaluations. Key trends were those that were mentioned in the greatest number of evaluations.

An analysis of trends in cost or cost-effectiveness could not be included because the information contained in the evaluations was insufficient. None of the evaluations adequately addressed cost or cost-effectiveness issues.

2.2.2 Evaluation quality

The quality of the programmes were assessed against criteria drawn from a report by the Australian National Audit Office, entitled *Program Evaluation in the Australian Public Service* (Canberra ACT, September 1997). The criteria were modified somewhat to fit the New Zealand context, with additional detail being added in some areas. Broadly the evaluation reports were assessed against the criteria in the following areas:

- completeness of the reports
 - a) table of contents
 - b) objectives
 - c) a description of background to the evaluation
 - d) a description of the programme being evaluated
 - e) a description of the scope of the evaluation
 - f) a description of evaluation methodology
 - g) key findings
 - h) executive summary
 - i) recommendations
- correspondence between the objectives and the rest of the report
- appropriateness and quality of methods of data collection
- relationship between the conclusions and the data
- quality of the recommendations
- presentation of the report

The complete list of criteria used is included in Appendix 1.

2.2.3 Evaluation use

A sample of six evaluations was selected to determine the utilisation of evaluations. For the purposes of this evaluation, utilisation refers to *documented actions* taken based on the results or recommendations of the evaluation (i.e. documented changes to the programme). Agency staff were asked a series of questions on the utilisation of the selected evaluations (refer to Appendix 2).

Two evaluations from each organisation (NZES, CEG, DSW) were selected from the list provided in Appendix 3 according to the following criteria:

- the evaluation was completed within the last two years, and
- enough time has elapsed for any recommendations to have been implemented (i.e. late 1996, early 1997).

NZES, CEG and DSW were selected because they will form part of the new agency.

2.2.4 Limitations

There were several constraints on this meta-evaluation:

- It was difficult to compare the results of the evaluations because the methodologies and robustness and rigour of the evaluations included in the study varied considerably. Comparisons were also limited because of the different nature of some of the programmes.
- There may be some publication bias. While every effort was made to ensure that all the evaluations of employment programmes undertaken since 1992 were included some may have been missed. Some recent evaluations were not included because they were not completed at the time the analysis took place (i.e. the evaluation of Job Plus Maori Assets).
- The focus of this evaluation was on programmes evaluated. These programmes do not necessarily reflect the range of employment programmes available in each agency.
- The review relied primarily on the evaluation reports as a source of data. As a result, the underlying causes of issues identified are not always clear. Ideally interviews would have been undertaken with key agency staff about the issues identified in each evaluation but time and resource constraints meant this was not possible.

3. RESULTS

3.1 TYPES OF PROGRAMMES

Nine different programme types were identified amongst the 35 employment programmes⁴ (Table 1). Approximately two-thirds of evaluated programmes were undertaken by NZES. Fifteen of these programmes were life skills or motivational training programmes and a further eight programmes were wage subsidy or work experience programmes. NZES undertook the greatest range of programmes.

Table 1: Organisations undertaking evaluations by programme type

Programme Type	Organisations undertaking evaluations				
	NZES	CEG	DSW	ETSA	MOYA
Wage subsidy	3, 7, 9, 14, 20, 21, 31				
Work experience	2				
Training: life skills, motivation	6, 17, 25, 27, 29, 32		10, 11, 34, 35	1, 4, 13	16, 22
Training: job or business-related	12	33	11	1, 4, 13	
Job search assistance	8, 15, 19, 30				
Community development		23, 24			

⁴ Refer to Appendix 2 for a description of each of the programmes included in this meta-evaluation.

Assistance to people in work			26		
Knowledge building	5, 18				
Strategic initiatives	28				

Key:

1. Access Training, 1992
2. Community Taskforce, 1992
3. Taskforce Green, 1993
4. TOP, 1994
5. The Behaviour and Perceptions of Long Term Job Seekers, 1994
6. Limited Service Volunteer Scheme, 1994
7. Job Plus, 1994
8. Joint NZES/NZIS Work Focus Interviews, 1994
9. TFG, Tourism Green and Possum Control, 1995
10. Compass Pilot Programme, 1995
11. Training Incentive Allowance, 1995
12. Job Plus Training Pilot, 1995
13. Survey of TOP Trainees, 1995
14. Job Skills, 1995
15. Job Action, 1995
16. NZCC and YSC programmes, 1995
17. Tane Atawhai, 1995
18. Barriers to Employment Facing Long-Term Job Seekers, 1996
19. Youth Action Programme, 1996
20. Job Support Programme, 1996
21. Enterprise Allowance and the Capitalisation Option, 1996
22. NZCC Two Year Post-Course Evaluation, 1996
23. CEG Pacific Peoples Evaluation Report, 1997
24. Mahi A Iwi Evaluation Report, 1997
25. Maori Youth Pilot, 1997
26. OSCAR DAP Pilot, 1997
27. Wahine Pakari Programme, 1997
28. IEA Pilot, 1997
29. Hikoi Ki Pae-rangi/New Horizons, 1997
30. Job Action and Youth Action Programmes, 1997
31. Job Connection, 1997
32. Tama Tane O Le Pasefika, 1997
33. Be Your Own Boss, 1997
34. Compass After Two Years, 1997
35. BOOST, 1997

3.2 TARGET GROUPS

The greatest range of programmes was targeted at the long-term unemployed. The long-term unemployed are those who have been unemployed for at least 26 weeks and include sub-groups such as Maori, Pacific Island people and women (Table 2). The range of programmes specifically targeting Maori, Pacific Island people, women, sole parents and the disabled was limited compared to that specifically targeting the long-term unemployed and youth (Table 2).

The majority of programmes directly targeting Maori, women, and sole parents were life skills or motivational training programmes. In addition, one of the two programmes directly targeting Pacific Island peoples was a life-skills or motivational training programme (Table 2). There are no wage subsidy or work experience programmes directly targeting Maori⁵, Pacific Island people, women, or sole parents included in this report. There were 19 evaluated programmes targeting the long term unemployed, including six wage subsidy or work experience programmes, three job or business-related training programmes, and six life skills or motivational programmes.

Table 2: Programme type by target group

Programme Type	Target Group							
	Long-term job seekers	Youth	Maori	Pacific Island peoples	Women	Sole parents	Disabled	Other
Wage subsidy	3, 7, 9, 31	14					20	21
Work experience	2							
Training: life skills, motivation	4, 13	6, 16, 22, 35	17, 25, 27, 29	32	27, 29	10, 34		1, 11

⁵ An evaluation of the Job Plus Maori Assets programme was not completed in time to include in this evaluation. The Job Plus Maori Assets programme is a wage subsidy programme targeting Maori.

Training: job or business-related	4, 12, 13								1, 11, 33
Job search assistance	8, 15	19, 30							
Community development			24	23					
Assistance to people in work							26		26
Knowledge building	5, 18								
Strategic initiatives									28

Key:

1. Access Training, 1992
2. Community Taskforce, 1992
3. Taskforce Green, 1993
4. TOP, 1994
5. The Behaviour and Perceptions of Long Term Job Seekers, 1994
6. Limited Service Volunteer Scheme, 1994
7. Job Plus, 1994
8. Joint NZES/NZIS Work Focus Interviews, 1994
9. TFG, Tourism Green and Possum Control, 1995
10. Compass Pilot Programme, 1995
11. Training Incentive Allowance, 1995
12. Job Plus Training Pilot, 1995
13. Survey of TOP Trainees, 1995
14. Job Skills, 1995
15. Job Action, 1995
16. NZCC and YSC programmes, 1995
17. Tane Atawhai, 1995
18. Barriers to Employment Facing Long-Term Job Seekers, 1996
19. Youth Action Programme, 1996
20. Job Support Programme, 1996
21. Enterprise Allowance and the Capitalisation Option, 1996
22. NZCC Two Year Post-Course Evaluation, 1996
23. CEG Pacific Peoples Evaluation Report, 1997
24. Mahi A Iwi Evaluation Report, 1997
25. Maori Youth Pilot, 1997
26. OSCAR DAP Pilot, 1997
27. Wahine Pakari Programme, 1997
28. IEA Pilot, 1997
29. Hikoi Ki Pae-rangi/New Horizons, 1997
30. Job Action and Youth Action Programmes, 1997
31. Job Connection, 1997
32. Tama-Tane O Le Pasefika, 1997
33. Be Your Own Boss, 1997
34. Compass After Two Years, 1997
35. BOOST, 1997

Table 3 shows participation rates in the evaluated programmes. Figures show participation in the programme or the evaluation at the time the evaluations were carried out, and may not reflect current levels. These figures were taken from the evaluation reports. In some cases the figures refer to the results of sample surveys. In other cases the figures are derived from an administrative database.

Maori, Pacific Island peoples and women were under-represented in subsidy programmes which enabled participants to start a business or remain in a job (i.e. Enterprise Allowance and Job Support). For example, Maori make up 28% of the NZES register but only 11% of those receiving the Enterprise Allowance, and 9% of those receiving the Job Support subsidy were Maori (Table 3).

Maori participation was similar or greater than the proportion of Maori on the NZES register in wage subsidy or work experience programmes aimed improving chances of obtaining stable employment (i.e. Job Connection, Taskforce Green). This was not the case for women or Pacific Island peoples. For example, women make up 37% of the NZES register but only 18% of Job Connection, and 17.8% of Taskforce Green participants were women. Similarly, 8% of those on the NZES register are Pacific Island peoples but they only represent 3.3% of Taskforce Green participants and 6.6% of Job Connection participants (Table 3).

Table 3: Participation rates in the evaluated programmes

Programme Type	Pakeha %	Maori %	Pacific Island %	Other %	Women %
Wage subsidy and work experience					
Evaluation of Community Taskforce	-	-	-	-	-
Evaluation of Taskforce Green	55.9	31.8	3.3	9.0	17.8
Evaluating the Effectiveness Job Plus	-	-	-	-	-
Evaluation of TFG, Tourism Green and Possum Control					
• TFG	51.9	45.5	3.8	1.8	22.8
• Tourism Green	61.3	37.4	0.0	1.3	5.2
• Possum Control Programme	64.2	34.3	1.3	0.5	2.0
Evaluation of Job Skills	83.0	-	-	-	40.0
Evaluation of Job Support Programme*	85.0	9.0	3.0	3.0	39.6
Enterprise Allowance and the Capitalisation Option					
• Enterprise Allowance	80.0	11.0	2.0	8.0	31.0
• Capitalisation Option	71.0	16.0	4.0	9.0	22.0
Evaluation of Job Connection	49.0	42.0	6.0	2.0	18.0
Training					
The Effectiveness of Access Training□	57.5	33.0	6.0	4.0	52.5
TOP Review□	-	-	-	-	-
Evaluation of Limited Service Volunteer Scheme	63.0	27.5	4.8	4.6	-
Evaluation of Compass Pilot Programme*	60.0	20.0	6.0	14.0	96.0
Evaluation of Training Incentive Allowance*□	68.6	24.5	2.2	4.7	91.0
An Evaluation of Job Plus Training Pilot□	59.0	40.0	6.0	3.0	52.0
Assessment of TOP: Survey of TOP Trainees*□	-	-	-	-	-
Evaluation of NZCC and YSC programmes					
• NZCC	34.6	58.0	2.5	-	24.7
• YSC	36.3	42.5	15.9	-	31.9
Tane Atawhai Evaluation Report	1.0	94.0	1.0	4.0	-
NZCC Two Year Post-Course Evaluation	-	-	-	-	-
Evaluation of Meoni Youth Pilot	-	94.0	6.0	-	50.0
Evaluation of Wahine Pakari Programme*	-	-	-	-	100.0
Evaluation of Tama Tane O Le Pasefika	-	-	-	-	0.0
Be Your Own Boss: National Outcomes Report*□	66.0	21.0	2.0	11.0	40.0
A Quantitative Evaluation of Compass After Two Years*	-	-	-	-	-
Job search assistance					
Evaluation of Joint NZES/NZIS Work Focus Interviews	-	-	-	-	-
Evaluation of Job Action	53.7	32.8	9.4	4.1	21.0
An Operational Review of the Youth Action Programme	59.0	34.0	7.0	-	44.0
Enhanced Job Action Programmes	-	-	-	-	-
Enhanced Youth Action Programmes	-	-	-	-	45.7
Community development					
CEG Pacific Peoples Evaluation Report*	-	-	-	-	-
Mahi A Iwi Evaluation Report*	-	-	-	-	-
Assistance to people in work					
Evaluation of OSCAR DAP Pilot*	45.0	33.0	18.0	4.0	-
Knowledge building					
Behaviour and Perceptions of Long Term Job Seekers	NA	NA	NA	NA	NA
Barriers to Employment Facing Long-Term Job Seekers	NA	NA	NA	NA	NA

Strategic initiatives

Evaluation of IEA Pilot

Proportion on the NZES Register at April 1998	54.0	28.0	8.0	10.0	37.0
a) Programmes with an * include participants who may not be registered with NZES (i.e. DPB beneficiaries).					
b) Programmes with a □ are job or business-related training programmes. Some programmes such as TOP, also life skills training programmes.					
c) A dash (-) indicates the relevant data was not included in the evaluation report.					

- a) Programmes with an * include participants who may not be registered with NZES (i.e. DPB beneficiaries).
- b) Programmes with a □ are job or business-related training programmes. Some programmes such as TOP, also life skills training programmes.
- c) A dash (-) indicates the relevant data was not included in the evaluation report.

Maori had high rates of participation (e.g. over 28%) in other programmes such as Access, Job Plus Training, Job Action, Conservation Corps and Youth Services Corps, Youth Action, OSCAR DAP, and Hikoi ki Pae-rangi. Pacific Island peoples high rates of participation (i.e. over 8%) in Job Action, Youth Services Corps, OSCAR DAP, and Hikoi Ki Pae-rangi. Women had high participation rates (i.e. over 37%) in the following programmes which are not directly targeted at sole parents or women:

- Access
- Job Plus Training
- Job Skills
- Youth Services Corps
- Youth Action
- Job Support
- Maori Youth Pilot
- Be Your Own Boss.

Except for Job Support, these are all training or life-skills and motivational programmes. It is important to note that some programmes are open to women who are not registered as unemployed with NZES (Table 3).

3.3 THE FOCUS OF THE EVALUATIONS

The evaluations were categorised according to whether they focused on outcomes (summative), process, cost or basic knowledge building. The majority of programmes were focused on measuring the impact of a programme and how the programme was being implemented (Appendix 4). Refer to Appendix 5 for a fuller discussion of process and summative evaluations.

3.3.1 Cost and cost-effectiveness

Only seven of the 35 evaluations included in this study examined the cost of the programme in any form (Appendix 4). None of the evaluations adequately examined the programme cost or cost effectiveness. For example, none the evaluations included administrative costs such as staff time in their attempts to evaluate cost or cost-effectiveness. The lack of focus on cost results primarily from a lack of comprehensive cost information. Adequate information is currently not collected on programme costs and administrative costs (i.e. cost of staff time and support costs). NZES commented that until there is a reliable estimate of staff time then developing a robust analysis of cost effectiveness is difficult. However, information on programme costs and administrative costs is essential for determining the cost and/or cost-effectiveness of a programme or suite of programmes. All financial information should be collected in a way that allows aggregation by activity and by outputs. The

importance of evaluating the cost-effectiveness of programmes and interventions will only increase with introduction of the new agency and greater regional flexibility.

3.3.2 A focus on individual programmes

All of the evaluations included in this study focused on individual programmes (Appendix 4). There has not been a requirement for the agencies to evaluate suites or combinations of programmes or interventions, sequences of programmes or interventions, or outcomes for demographic or beneficiary groups across programmes or interventions. However, as of October 1 1998, WINZ will be focused on returning people to work using regional flexibility as a way to tailor programmes to the needs of the local labour markets. As a result, regional commissioners will need to know what programmes work, in what combinations and for which target groups. Moreover, the emphasis on case management and other forms of "individualised" assistance, means that in addition to evaluating discrete programmes there will need to be an additional focus on evaluating the impact of broader strategies or inter-linked series of interventions on individuals' labour market outcomes.

- a) *Types of programmes or interventions*: An evaluation of a type of programme or intervention would examine the effectiveness of the programme type in terms of job seekers achieving employment and training outcomes. Life skills and motivational programmes are a clear example of a type of programme that needs to be evaluated. The majority of programmes targeting the Maori, Pacific Island peoples, women, and sole parents are life-skills or motivational programmes. Some evaluations have examined outcomes for individual life skills and motivational programmes. However, there has not been an evaluation which examined the extent to which life-skills or motivational programmes led to employment and training outcomes, and how effective this programme type is compared to other programme types (e.g. wage subsidy programmes).
- b) *Sequence of programmes or interventions*: An evaluation focusing on a sequence of programmes would look at how particular programmes build on each other to ultimately lead job seekers into stable employment. Evaluating sequences of programmes has particular relevance for NZES as the concept of staircasing job seekers into work underpins much of NZES' work with job seekers. However, no work has been done to evaluate whether or not staircasing is effective in placing job seekers into work, which groups are most likely to benefit from staircasing and which combinations of programmes are most effective at placing people into work.
- c) *Demographic or beneficiary groups*: An evaluation could be undertaken looking across programmes available to particular demographic or beneficiary groups to determine which programmes or interventions are most beneficial and under what circumstances. For example, an evaluation could determine which of the programmes available to Maori are most effective at meeting their employment and training needs, and identify gaps in the provision of programmes or interventions.
- d) *Regional variation*: An evaluation could examine the extent to which the selection and implementation of programmes or interventions, and participant outcomes vary across the regions. With the move towards greater regional flexibility, policy makers will need to know which programmes and interventions are most effective

for which groups in which regions. Without such information it will be difficult for policy makers to develop new, or assess existing, broad coverage programmes.

3.4 OUTCOMES FOR PARTICIPANTS

3.4.1 Do programmes place people into employment?

Twenty-two evaluations identified employment outcomes for participants (Appendix 6). However, as Appendix 4 indicates, not all of the evaluations were required to focus on outcomes for participants. These results must be treated with caution owing to the constraints outlined in section 3.4.4.

A Wage subsidy and work experience programmes

Of the eight wage subsidy and work experience programmes included in this study, seven examined outcomes for participants⁶. The highest percentages gaining employment were Job Plus and Job Skills participants⁷. Forty-seven percent of Job Connection participants obtained full-time employment. These studies did not take into account displacement, substitution or dead weight effects. Job Support was also successful at placing participants into employment. However, it is important to point out that a significant number of the participants were already in employment prior to receiving assistance through Job Support. The Community Taskforce, and Taskforce Green programmes appear to get similar numbers of people into employment.

None of the wage subsidy and work experience programmes specified how much of an impact would be necessary to consider the programme “successful”. A study looking at wage subsidy programmes in OECD countries found that none of the schemes were particularly successful in terms of net jobs created. Most of schemes where target group additionality⁸ was measured came out with estimates of between 10% and 30% (NERA 1995: 13). However, Martin (1998: 12) points out that:

wage subsidy programmes do give an advantage to the target group compared with other job seekers, and the resulting redistribution of job opportunities may be justified on equity grounds.

Martin (1998) goes on to say that wage subsidy programmes may serve to maintain workers’ attachment to the labour force. However, he recommends that the subsidies be of short duration, targeted and closely monitored.

B Employment outcomes for other programmes types

Employment outcomes were measured for training programmes. However, with the possible exception of job-related training programmes, these programmes are not designed to lead participants directly to employment. Life skills and motivational programmes, for example, are often aimed at increasing the confidence of and skills of

⁶ The evaluation of Enterprise Allowance Programme recorded time spent off the register.

⁷ 65% of Job Skills and 66% of Job Plus participants obtained employment outcomes. This is consistent with employment outcomes achieved by JOS participants. JOS, an earlier version of Job Plus, was used as a comparison group in the evaluation of ACCESS.

⁸ Additionality, the number of extra jobs created by a programme, is not the only measure of a programmes success. For example, in the US the focus is on the degree to which programmes result in increased earnings for participants.

job seekers so as to improve their chances of finding work. Such programmes are not expected to achieve high employment outcomes.

Nevertheless, 13 of the 16 training programmes included in the study were required to examine outcomes. The highest percentages gaining employment were ACCESS⁹ and Wahine Pakari¹⁰ participants. The lowest percentages of participants were placed into employment following Tama Tane O Le Pasifika. None of evaluations focusing on training programmes specified how much of an impact would be necessary to consider a programme “successful”. International literature suggests that the outcomes for participants in public training programmes are mixed. The most consistently positive results have been recorded for adult women. Some programmes lead to positive results for men but very few training programmes lead to positive results for youth (Martin 1998). However, as Martin (1998) points out, most evaluations have very short time horizons, there is very little evidence on which types of and content of training programmes work best, and it is unclear why particular training programmes work well for particular groups (i.e. adult women).

It was not possible to draw firm conclusions from the data about employment outcomes for life skills and motivational programmes compared with wage subsidy and work experience programmes. Most of the wage subsidy and work experience programmes (i.e. Job Connection, Job Skills and Job Support) had better employment outcomes than most of the life skills and motivational programmes. However, some life skills and motivational programmes (i.e. NZCC and YSC) had better outcomes than the some wage subsidy programmes.

Programmes focusing on job search assistance appeared to be successful at placing people into employment, particularly Job Action and Youth Action. International literature indicates that job search assistance is usually the least costly active labour market programme and shows consistently positive outcomes (Martin 1998).

Insufficient data means it is not possible to comment on the employment outcomes for the programmes focusing on community development, IEA and OSCAR.

3.4.2 Training outcomes

Life skills and motivational programmes were more successful at placing people into further training than other types of programmes. However, placing people into further training is not the objective of other programme types such as wage subsidy and work experience programmes. The Compass (1995) and Hikoi Ki Pae-rangi (1997) programmes had the highest percentage of participants entering further training (Appendix 7). Both of these programmes target women, whom overseas literature suggests derive the greatest benefit from training programmes (Martin 1998).

3.4.3 Attitudinal outcomes

Approximately half of the evaluations measured the changes in attitude the course may or may not have had on participants. In these evaluations the majority of participants reported positive attitudes, which most frequently meant improved self-

⁹ 41% of participants achieved an employment outcome

¹⁰ 40% of participants achieved an employment outcome.

esteem and confidence, and/or improved motivation. In other cases, programmes were said to have a positive impact on the following:

- work habits
- presentation at work
- group interaction skills
- job search commitment
- broadening horizons in terms of options available
- sense of achievement
- attitude towards getting a job
- attitude towards agency (i.e. NZES, IS).

The approaches used for examining programme impacts on participant attitudes were not consistent or systematic. In most cases, attitude changes were measured by means of focus groups or surveys, in which programme participants were asked how they felt about their experiences. In some cases, participants were asked whether they believed their employment prospects had improved as a result of taking part in some programme. Some evaluations, such as the evaluation of Hikoi ki Pae-rangi, used a before and after study to examine changes in attitude. The evaluation reports do not generally show whether a) improvements in attitude continue over time or, b) more importantly, improved attitudes lead to improved employment outcomes. This last point is particularly relevant given the large number of life-skills and motivational programmes evaluated. One of the aims of life-skills and motivational courses is improve job seeker confidence so that they have improved chances of finding a job.

3.4.4 Difficulties in comparing outcomes

This evaluation was not able to state with any confidence which interventions or programmes worked best for which target groups owing to difficulties in comparing employment and training outcomes across programmes.

- a) *Employment outcome measures*: Most of the evaluations used several different means of measuring employment outcomes. For example, some evaluations measured employment outcomes in terms of full-time, part-time, short duration and temporary work while others only looked at those obtained full- and part-time work and in some cases the nature of the employment was unspecified. Within each of these categories some evaluations also distinguished between subsidised and unsubsidised work.

However, the difficulties associated with measuring outcomes across the evaluations need to be placed in context. There are several different government agencies represented in this meta-evaluation. Each of these agencies service different populations and have different goals. Under these circumstances it is unlikely that they would use the same outcomes measures. Within agencies, the focus has been on evaluating individual programmes rather than suites of programmes. As a consequence, evaluators have tended to develop outcome measures suited to the programme and the circumstances at the time. For example, it is only relatively recently that NZES developed a standard definition of stable employment.

- b) *Counterfactuals*¹¹: In several cases the evaluations did not specify how the programme would be judged a success or lacked reliable means for determining counterfactuals. There appear to be two reasons for this: a) the size of the impact necessary to consider a programme “successful” was not specified, and b) there were problems in designing comparison groups and counterfactuals. In the case of the latter, some evaluations relied on comparison groups of job seekers matched to those who participated in the programmes by demographic criteria or by location, in order to estimate what outcomes would have resulted in the absence of a given intervention. The development of counterfactuals was limited by the implementation approaches used for the programmes under study. In some situations, for example, a programme was offered to all eligible job seekers at the same time, so no control group could be established. No experimental designs were used, which would have relied on randomly assigned control and treatment groups.

There are indications that design of appropriate counterfactuals is changing. An evaluation of the expanded Community Taskforce, currently underway in NZES, is using randomly assigned control and treatment groups to measure outcomes for participants. However, such methods will not be appropriate for every evaluation (Appendix 5).

- c) *Administrative data*: The interpretation of administrative data was inconsistent. For example, in the case of NZES register data, the following measures were used by different evaluations:
- number of people still on the register after a specified period;
 - number of people who have left the register within a specified period,
 - total number of days spent on the register.
- d) *Definition of further training*: There was no consistent definition as to what constituted further training. In some cases “further training” was not defined while in other cases a detailed description was given of the courses participants went on after completing the programme. Some only looked at participation in TOP courses.
- e) *Outcome measurement period*: The duration over which outcomes were measured varied from zero months to two years. Time constraints placed on evaluations (e.g. Cabinet reporting requirements) led to inconsistent outcome measurement periods.
- f) *Definitions of ethnic or demographic groups*: The reporting of outcomes for particular ethnic or demographic groups varied across the evaluation reports. For example, outcomes for Pacific Island people were often not reported or were combined with other ethnic groups (e.g. “other” or “non-Maori”). Differing definitions of ethnic and demographic groups combined with other difficulties listed above, meant that it was not possible to compare outcomes for particular groups across different programmes or interventions.

¹¹ Counterfactuals are an estimate of what would have happened if the programme did not exist. The counterfactual could be outcomes for a group of non-participant, outcomes for a group before the intervention is introduced, or some other construct.

g) *Measuring attitudinal change*: Examining the impact of a course on participants' attitudes does not appear to have been examined in any systematic manner. None of the evaluations that examined participants' attitudes, attempted to measure the extent to which attitudinal change was sustained over time. NZES has recognised the need to develop more robust and consistent measures of attitudinal change. For example the evaluation of Residential Motivational Training (currently underway) is attempting to measure attitudinal change of participants on a before and after intervention basis. In addition the evaluation is tracking cohorts of participants to measure the robustness of attitudinal change over time and across different providers.

3.5 OPERATIONAL STRENGTHS AND WEAKNESSES

The review identified a number of key operational issues evident in several evaluations that will need to be resolved if the aims and objectives of programmes are not to be undermined. The issues were:

- job seekers were not appropriately matched to programmes and interventions,
- the programmes often failed to meet the programme objectives or the intent of the programme,
- monitoring of the programmes and job-seeker progress was frequently inadequate, or
- case management or post placement support was frequently inadequate.

It was more difficult to identify the key strengths across the evaluations because the evaluations tended to focus on the operational weaknesses. However, evaluations also provide an opportunity to identify programme strengths. Such information can be used to develop 'best practices'.

3.5.1 Recruitment and selection of programme participants

The review revealed that the recruitment and selection of participants for programmes and interventions varied. Some evaluations identified positive aspects about the programme recruitment and selection process. However, a number of evaluations indicated that, as a result of inadequate programme recruitment and selection processes,

- places are sometimes taken by job seekers who do not belong on the programme, and
- the most appropriate job seekers (i.e. those with the greatest need or those who will get the most out of the programme) are not selected to participate in the programme.

Recruitment was also identified as an issue in a recent review of the Employment Taskforce programmes¹². For example, problems with recruitment were evident across all of the seminar based programmes.

A Recruitment and selection procedures

¹² Eight ETF programmes were evaluated: Enhanced Job Action and Youth Action; Job Connection; Tama Tane o le Pasefika; Maori Youth Pilot; Wahine Pakari; Hikoi ki Pae-rangi/New Horizons; and the Maori Women's Development Fund. All of these programmes, except Job Connection and the Maori Women's Development Fund, were seminar based programmes.

A key factor determining whether or not job seekers receive the appropriate assistance is the programme recruitment and selection process. The review indicated that the recruitment and selection process varied across the programmes. For example, seven evaluations stated that poor recruitment and selection procedures and practices led to poor matching of job seekers to programmes and interventions. The evaluation of the Maori Youth Pilot (1995), for example, stated that:

In attempts to fill courses, inappropriate recruitment meant that places were not always filled by the most suitable participant.

In another example, recruitment to the enhanced Job Action workshops was also problematic. Employment Advisors sometimes over-recruited for workshops (i.e. they got up to 30 job seekers to say that they will attend a workshop so that 15 turn up). The problem with this 'over recruiting' is that at times, in excess of 20 job seekers turn up at a workshop (and some have been turned away), and at other times as few as eight participants.

On the other hand, five evaluations identified positive aspects about the recruitment and selection process. For example, the evaluation of the Wahine Pakari programme indicated that the recruitment process—using Maori networks to recruit Maori women—contributed to program success. In other examples, the Tane Atawhai evaluation noted that the programme relied on personal, face-to-face recruitment approaches to good effect.

B Recruitment guidelines and criteria

Another key factor determining whether or not job seekers receive the appropriate assistance are the recruitment guidelines and criteria. Agency staff are less likely to select the appropriate job seekers for a programme or intervention if the recruitment guidelines and criteria are inadequate. The review indicated that the quality of the recruitment guidelines and criteria varied across the programmes.

Seven evaluations indicated that inadequate recruitment guidelines and criteria were inadequate. For example, the evaluation of Job Skills (1995) indicated that Operational practice was not within the Cabinet guidelines and a greater number of under 26 week job seekers have entered Job Skills than agreed by Cabinet. In another example,

A high proportion of participants (44%) were already in work, indicating that Job Support funds have often been used to help clients to improve their work performance, rather than directly save jobs. This indicates Workbridge staff were not rigorously applying the eligibility criteria for applicants already in employment (Evaluation of Job Support, 1996).

Alternatively, there were some instances where the evaluations made positive statements were made about programme recruitment guidelines and criteria. For example, in the evaluation of Job Connection, NZES staff stated that found the programme easy to administer because the eligibility criteria were simple.

3.5.2 Meeting the objectives or the intent of the programme

The evaluation reports indicated that some programmes did not meet their objectives or the intent of the programme. Where the programme objectives or intent are not met job seekers are less likely to get the full benefit of the programme. The following were identified as key factors influencing whether or not the objectives or intent of the programme were met:

- programme guidelines and criteria,
- agency staff skills and knowledge of the target group and/or the programme,
- resources available to agency staff and providers to deliver the programme, or
- course presenters or providers.

A programme guidelines and criteria

There were some instances where the evaluations made positive comments about adherence programme guidelines or criteria. For example, the evaluation of the New Zealand Conservation Corps and Youth Services Corps revealed that projects met the criteria in the majority of cases.

On the other hand, eight evaluations indicated that staff or providers did not adhere to the existing programme guidelines or criteria. In the case of Job Skills (1995), operational practice was not within the Cabinet guidelines - a greater number of job seekers unemployed for less than 26 weeks had entered Job Skills than agreed by Cabinet, not all the employers were from the private sector, and over 80% of participants were not available on the fifth day for other activities. In another example, the Job Support (1996) evaluation revealed:

A high proportion of participants (44%) were already in work, indicating that Job Support funds have often been used to help clients to improve their work performance rather than directly save jobs. This indicates Workbridge staff were not rigorously applying the eligibility criteria for applicants already in employment

Eight programmes did not have adequate guidelines or criteria in place to ensure that the objectives or intent of the programme was met. The evaluation of the Youth Action programme (1995) revealed that inadequate guidelines and procedures to support case management was one of the reasons the number of job seekers receiving case management was low. In another example, the evaluation of the OSCAR programme revealed that:

There were insufficient policy and funding distinctions between the three activities which are core to the establishment and delivery of any viable programme: - feasibility analysis, service development, and service delivery.

B Agency staff skills and knowledge of the target group or the programme

The skills of agency staff and their knowledge of the target group or the programme is a contributing factor in whether or not the objectives or intent of the programme are met. The evaluation reports indicated that skills of agency staff and their knowledge of the target group or the programme varied.

Fifteen evaluations indicated that agency staff had insufficient a) knowledge of the target group, b) knowledge of the programme, or c) skills to implement the programme as required. For example, the evaluation of the Youth Action programme

(1995) indicated that NZES staff were not skilled enough to assess job seekers with a wide range of social and psychological problems. In another example, the evaluation of Boost (1997) revealed that staff had little or no awareness of how to provide a culturally safe service. In the case of the Pacific Island Peoples evaluation, some of the groups interviewed felt CEG lacked real knowledge of business practices.

Alternatively, positive comments were made in six evaluations about agency staff knowledge of and relationship with the target group. For example, the Mahi A Iwi and Pacific Island peoples evaluations both made positive comments were made about the ability of CEG staff to motivate and form close relationships with the target groups. The evaluations of the Job Skills and Youth Action programmes both made comments about benefits of one-to-one contact between agency staff and job seekers.

C Resources available to agency staff and providers to deliver the programme

A number of evaluations revealed that constraints on resources available to the agency or the provider agency was the reason the programme purpose or objectives were not being met. Eight evaluations indicated that constraints on resources available to the agency were a problem. For example, the evaluation of Job Action and Youth Action (1997):

some Employment Advisors stated that they do not have the resources to re-do Action Plans or show the provider what is required.

Many of the Employment Advisors interviewed stated that the time provided to implement Youth Action was inadequate.

The evaluation of Job Action (1995), Youth Action (1995), Job Connection and Tama Tane all indicated that Employment Advisor had insufficient time to undertake adequate case management. For example, the Tama Tane evaluation revealed:

This issue of case management is directly related to the workload of EAs. Unless the workload of EAs who are responsible for case managing of Tama Tane o le Pacifica participants is reviewed the problem will continue. Adequate resourcing is therefore required.

Four evaluations indicated that constraints on resources available to the provider was the reason the programme purpose or objectives were not being met. For example, in an evaluation of TOP (1994), the majority of providers mentioned increased costs as a growing constraint on the way they ran their courses (i.e. costs associated with complying with the National Qualifications Framework, OSH). In the case of the Job Support programme the lack of disability assessors sometimes led to delays in the provision of services to clients.

D Course presenters or providers

The evaluation reports indicated that the quality of the course presenters is a contributing factor in whether or not the objectives or intent of the programme are met. The quality of the course presenters or providers varied across the programmes, according to the evaluation reports.

Eight evaluations, for example, indicated that the programme objectives or the intent of the programme were not being met because course presenters or providers were

inadequate. For example, the Workbridge was contracted to run the Job Support programme but the evaluation revealed they did not rigorously apply the eligibility criteria for applicants already in employment, and nor did they undertake the post-placement follow-up contacts with programme participants as required in the contract.

Alternatively, six evaluations made positive comments were made about the course presenters or facilitators. Most of these comments were very general statements. However, evaluation of the Wahine Pakari programme indicated that the close attention paid by contractors to facilitator selection and training was the major factor ensuring seminar quality and consistency.

3.5.3 Programme monitoring

Ten evaluations indicated that there were problems associated with monitoring. These evaluations identified one or more of the following problems:

- a failure to carry out monitoring as required,
- a failure to carry out monitoring because it was not requested, or
- an inappropriate approach to monitoring (i.e. an ad hoc or inconsistent approach; use of monitoring tools that are not user-friendly).

Monitoring at various stages during and after the implementation of programme is essential to ensure the goals and objectives of the programme are met. If monitoring is inadequate there is a risk that the criteria for programme implementation will not be met (i.e. recruitment of inappropriate job seekers; providers failing to meet course implementation criteria; staff failing provide required support services such as case management; funding inappropriately allocated). Another risk is that new problems will remain unidentified and therefore unresolved. For example, three evaluations indicated that funding had been misallocated as result of inadequate monitoring. The following quote illustrates this point:

The distribution of the transport allowance needs to be monitored to ensure that job seekers, not employers, are receiving it. At present, 58% of job seekers are receiving only a proportion of the subsidy or no subsidy at all (Job Connection, 1997).

Other evidence suggests that insufficient or inadequate monitoring was a problem in more than just the ten evaluations mentioned above. For example, sixteen evaluations, indicated inadequate resourcing was a constraint on programme delivery but none of these evaluations adequately quantified the extent to which resources constrained delivery of the programme. In addition, few of the evaluations adequately defined "resources" and why they were constrained. If there were appropriate monitoring procedures to measure resource use, the agencies could determine whether or not resource levels (i.e. staff time) were inadequate. It was not clear from the evaluations why monitoring was not carried or not carried out effectively.

3.5.4 Case management and post-placement support

The evaluation reports revealed that the quality of case management and post-placement support differed across the evaluations. Positive comments about case management and post-placement support were made in five evaluations. For example,

in the evaluation of the IEA pilot, all participants regarded the opportunity to undertake regular follow-up for reporting against the work plan as a positive element of the new IEA process. In another example, the operational review of Youth Action indicated that Employment Advisors saw case management as a useful means of assisting participants to achieve goals set out in their Youth Action plans, and identifying further needs for assistance.

On the other hand, eight evaluations indicated that the benefits of the programme or intervention for job seekers were reduced because of inadequate case management or post-placement support. The evaluation of the Enhanced Job Action (pp23) for example, revealed that:

Follow-up is often late and sometimes non-existent ... While this does not influence the initial impact of the workshop, it significantly reduces the programme's overall effectiveness. This is for the following three reasons:

- *many job seekers need guidance in developing their Plans further.*
- *some job seekers need assistance in identifying training and employment opportunities.*
- *some job seekers need to be assisted in moving through their Plans including staying motivated, especially if they experience difficulties in implementing the Plan.*

Even when the follow-up does take place, the time available appears to be too little in many cases to really be useful.

Similar sentiments were expressed in the evaluation of the Hikoi ki Pae-rangi (pp20) seminar:

All participants interviewed reported that the post seminar follow-up needed development. The main impact of lack of follow-up is a loss of impetus for focusing on future direction. For many women to successfully move closer to employment or training, support and follow-up by NZES centre staff was crucial to achieving a successful outcome.

Case management and post placement support difficulties were also identified in the summary of the recently evaluated Employment Taskforce programmes. Case management was variously described as too late in commencing, inconsistent (intensive for some job seekers and non-existent for others) and or inadequate.

The reason why a significant number of job seekers did not receive any or sufficient case management, follow-up or post-placement support was unclear. However, the following suggestions were made in the evaluations:

- insufficient resources (three of the eight evaluations),
- these activities were not seen as a high priority for staff (two of the eight evaluations),
- inadequate guidelines and criteria (two of the eight evaluations), and
- low staff knowledge of, and commitment to the programme (two of the eight evaluations).

For example, the evaluation of Youth Action suggested that case management was inadequate because guidelines and procedures to support case management practices were inadequate, Employment Advisors lacked the time and an inefficient computer system compounded the problem (i.e. there were no Youth Action interview recording functions on the NZES computer system). The evaluation of Tama Tane o le Pacifica reported the following:

This issue of case management is directly related to the workload of EAs. Unless the workload of EAs who are responsible for case managing of Tama Tane o le Pacifica participants is reviewed the problem will continue. Adequate resourcing is therefore required.

The evaluations of Wahine Pakari and Hikoi ki Pae-rangi suggested that a lack of knowledge of, and commitment to the programme contributed to the low levels of follow-up. For example, evaluation of Hikoi ki Pae-rangi (pp20) reported the following:

The lack of knowledge of the knowledge of the seminar, commitment and ownership by IS and NZES contributed to poor follow-up. Contracting out the programme to external providers may have decreased the connection between IS, NZES and the programme. Most women attending the seminar were not NZES clients.

3.6 EVALUATION QUALITY

3.6.1 Summary

Evaluation quality was assessed according to the criteria listed in Appendix 1. The quality of the evaluations differed within and between agencies. To improve future evaluations, several issues surrounding the quality of the evaluations will need to be addressed. In particular, future evaluations will need to:

- clearly link evaluation objectives to the rest of the evaluation report;
- appropriately identify and support the key findings;
- provide conclusions with a reportable message; and
- state in the Terms of Reference how conclusions will be addressed.

3.6.2 The correspondence between the evaluation objectives and the rest of the report

The quality review identified 29 evaluations that met some of the criteria regarding links between the conclusions and the rest of the report (Appendix 8). For example, in most cases the objectives were explicitly addressed, and the evaluation was carried out a stage that was appropriate to the development of the programme (Table 4). However, review also identified a key weakness that will need to be addressed.

Twenty-six evaluations did not make clear links between the objectives, methods of data collection, conclusions and recommendations (Table 4). Examples of this weakness include:

- using methods that would not provide the information required to answer the objectives (e.g. the objective requires the evaluation to examine labour market

outcomes and the evaluation only provides an analysis of days spent by participants on and off the register),

- having conclusions that do not address the evaluation objectives, or
- not being able to link the objectives to the conclusion because the evaluation did not have a conclusion.

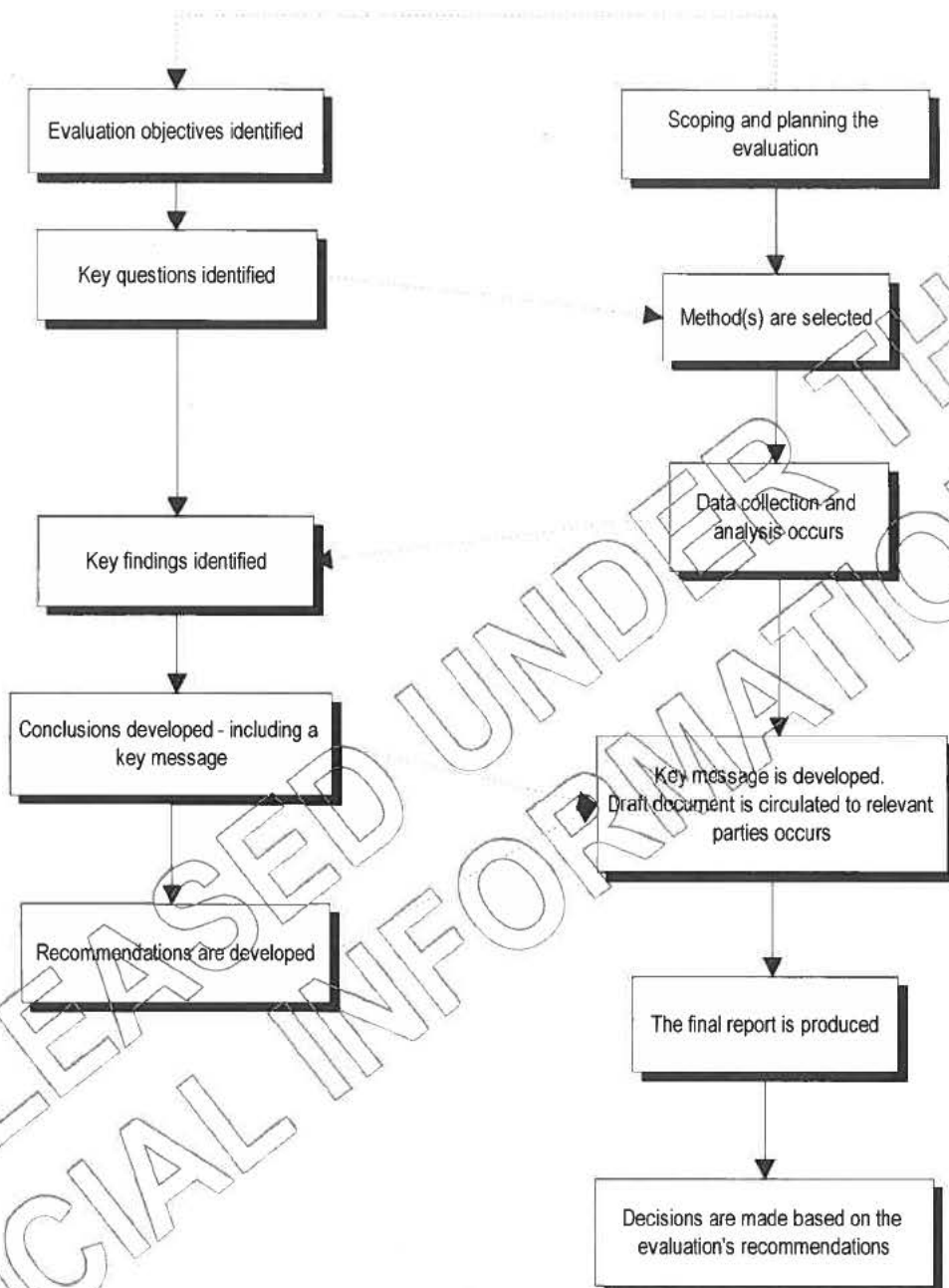
Table 4: Correspondence between the objectives and the rest of the report

Criteria	Objectives not included in the report	Evaluation did not meet criteria	Evaluation met criteria
a) all objectives should be explicitly addressed. For example, if the objectives related to efficiency, the information reported against it should be about efficiency not effectiveness	4	8	23
b) clear links between the objectives, methods of data collection, conclusions and recommendations	4	26	5
c) objectives are doable or any limitations are adequately described	4	13	18
d) objectives are appropriate to the stage of development of the programme	4	1	30

The evaluations used a range of methodologies (Appendix 9) to evaluate the implementation of programmes and/or outcomes for programme participants. However, further work is needed to determine which methods are most appropriate under which circumstances (e.g. when is it appropriate to use comparison groups such as that used in the 1995 evaluation of the Compass pilot?).

The evaluation objectives should specify the condition being dealt with along with one or more measurable criteria for success and should be clearly linked to other key aspects of the evaluation. As Figure 1 illustrates, the objectives should be developed during the scoping phase of the evaluation. The key questions, which flow from the objectives, should then inform the selection of the evaluation method(s). The findings, conclusions and recommendations should all be clearly related to the evaluation objectives. If the links between the objectives and rest of the report are weak, the implication is that the objectives were poorly developed (i.e. too vague) and/or the underlying problem or condition has not been addressed.

Figure 1: Links between the objectives and the rest of the report



3.6.3 The existence and quality of the key findings

Findings¹³ are the summation of the facts developed to meet the objective(s) of an assignment and form the basis for conclusions and recommendations. If the evaluation is to have any impact on decisions made about the programme it is critical that the findings are well supported. Once the evaluation is complete the evaluators often have little control over decisions made about the evaluated programme and decision-makers are likely to ignore poorly supported findings.

¹³ Findings may be negative and point up situations requiring correction or they may be positive and highlight programmes, policies, and procedures that work well and could be effectively applied to other areas. Findings are most useful when they point to the need for future improvements rather than placing undue emphasis on past deficiencies (GAO 1997: pp9.0-1).

Table 5: Key findings

Criteria	Key findings not included in the report	Evaluation did not meet criteria	Evaluation met criteria
The existence of key findings			
a) have been appropriately identified (i.e. address the underlying problem)	0	21	14
b) state the criteria employed	0	9	26
The quality of the key findings			
a) are supported by the evidence (i.e. effectively supported by appropriate levels of evidence ¹⁴ and source(s) ¹⁵)	0	17	18
b) contain no errors in logic or reasoning	0	10	25
c) answer the objectives	0	10	25
d) are appropriately qualified and cannot easily be misinterpreted or misquoted	0	11	24
e) identify the data sources on which the findings are based	0	2	33
f) are described adequately	0	10	25
g) identify the views of key stakeholders	0	2	33

Twenty-six evaluations stated the criteria employed in identifying the key findings. The criteria are the standard by which the programme is to be judged. The criteria may take the form of a comparison or a control group, a programme logic, the legislative intent of the programme, expert opinion, specified outcomes, or national or international experience. However, the key findings were not appropriately identified in 21 evaluations (Table 5). Examples of this include:

- providing a list of findings and leaving readers to determine which are the most important,
- providing key findings identified that did not answer the objectives, and
- providing key findings that were not supported by the data.

With regard to the quality of the key findings, twenty-six of the evaluations met some of the criteria. For example, almost all the evaluations identified the data sources on which the findings are based and the views of key stakeholders (Table 5). However, only nine evaluations met all the criteria (Appendix 9). Most frequently the findings were not supported by the evidence. For example, several evaluations indicated that the programme was a success but this was not supported by evidence from the evaluation.

3.6.4 Quality of conclusions

Most of the evaluations (25) included conclusions (Table 6). However, two key weaknesses were identified. Some evaluations (10) did not include conclusions (i.e. none of the evaluations undertaken by the SPA contained conclusions), and frequently, they did not contain a reportable message. This was concerning because the conclusion is the part of the evaluation that draws together the main findings into a reportable message to key stakeholders. If the report does not contain a conclusion

¹⁴ Evidence must be competent, relevant, and sufficient to support findings, conclusions, and recommendations.

¹⁵ If there is reason to doubt the evidence it should be corroborated by other evidence.

with a key message decision-makers may take no action or may not address the most important issues in the evaluation (Joint Committee on Standards for Educational Evaluation, 1994).

Table 6: Existence and quality of the conclusions

Criteria	Conclusions not included in the report	Evaluation did not meet criteria	Evaluation met criteria
Existence of the conclusions			
a) are clearly stated, not implied	10	11	14
b) have a reportable message	10	13	12
c) highlight evidence of significance (to get management to take corrective action)	10	10	15
Relationship between the conclusions and the data			
a) relevant data or argument should be provided to support conclusions	10	9	16
b) the conclusions should follow from the data and links should be clearly stated (direction and intensity of conclusions should be based on the findings)	10	9	16

The conclusions should also be defended and defensible. If stakeholders receive insufficient information on the strength of the conclusions, they may disregard them. Conclusions that are inadequate (i.e. not based on the findings) can lead decision-makers into taking inappropriate action. Only 13 evaluations met all the criteria regarding links between the conclusions and the data.

3.6.5 Evaluation report recommendations

A The existence and quality of the recommendations

This review of evaluation quality found that the recommendations were either non-existent or of poor quality (Table 7). Recommendations may not be required in the evaluation report. Generally the requirement that an evaluation have recommendations is outlined in the Terms of Reference but these were included in only one of the evaluations (Australian National Audit Office, 1997). Without the Terms of Reference it was unclear whether or not the recommendations were expected in the evaluations. However, if recommendations are present they should flow logically from the findings and be practical, actionable and able to solve the problems uncovered. If recommendations are not made or are of poor quality, decision-makers may take no action when action was warranted or may not address the most important issues arising from the evaluation.

Table 7: Existence and quality of the recommendations

Criteria	Recommendations not included in the report	Evaluation did not meet criteria	Evaluation met criteria
Existence of the recommendations			
a) clearly identifiable as recommendations rather than embedded in an undifferentiated way in the findings	25	2	8
Quality of the recommendations			
a) flow logically from findings	25	1	9

b) are practical in implementation	25	4	6
c) identify responsibilities for implementation and time frames	25	10	0
d) identify potential benefits and cost implications	25	10	0
e) identify the root cause of the problem	25	10	0

B Tracking the recommendations or findings from the evaluations

Discussions with officials in NZES, CEG, IS, ETSA and Youth Affairs revealed that none of these organisations have a formal system in place to track the recommendations resulting from the evaluations. Tracking the recommendations allows evaluators to monitor the extent to which the recommendations were implemented. This is a form of quality control as recommendations are less likely to be implemented unless they are actionable and relevant.

There were several reasons given as to why the agencies did not formally track recommendations.

- a) *No one is assigned responsibility for tracking the findings or recommendations.* For example, the NZES evaluation team generally develops recommendations in conjunction with (NZES) policy. The recommendations are then handed over to the Operations Groups (NZES) in a handover document. Once the handover has occurred, the implementation of the recommendations is the responsibility of the Operations Groups (NZES). There is no formal mechanism or system within NZES that checks that recommendations are acted upon. The manager of the NZES evaluation team indicated that it does not have the resources to check that recommendations are acted upon.
- b) *The number of evaluations undertaken by the agency was very small and the agencies do not see the need to implement a formal tracking process.* For example, CEG¹⁶ and Ministry of Youth Affairs had undertaken three and two evaluations respectively.
- c) *It is generally not the responsibility of the evaluation unit to develop and track recommendations or findings.* For example, within the Department of Social Welfare, most evaluations are for the purpose of providing information for policy development rather than operational purposes. Consequently most SPA evaluations do not include a list of recommendations. Recommendations are more likely to be included in a SPA evaluation report when the focus is on operations. Policy evaluations will provide conclusions that are just one source of information used to form the basis of policy recommendations. It is the clients responsibility to develop, implement and monitor any recommendations resulting from the evaluations. Any major policy initiatives or strategies arising from any information

¹⁶ Given there have only been two reports and they had an exploratory focus that will lead to further research, CEG has not seen a need to implement a formal system to track the recommendations. However, because of CEG's small size and field driven policy process there are many informal processes used to make use of research. The research is disseminated to all staff who then make use of it in their daily business as well as being used at a regional and national level to assess service gaps. The research recommendations were incorporated into the evaluation work programme.

source including evaluations are included in the SPA work plan and are closely monitored.

- d) *The findings or recommendations were incorporated into the policy process in an informal manner (i.e. CEG, SPA, ETSA, Ministry of Youth Affairs, NZES).* In some cases tracking the recommendations was made more difficult because of poor administrative or organisational structures. For example, it is possible to obtain copies of the NZES handover documents outlining the recommendations but it is not a straightforward process. The filing system is very fragmented and no-one has responsibility for the system. The NZES evaluation team commented that NZES was "not good at keeping paper and that more information on what happened regarding the recommendations was likely to be obtained by talking to operational and policy staff". It was unclear where the handover reports were stored - they could be on the evaluation files, policy files or operational files. There is no formal system for tracking the recommendations within Operations. Different people were involved in different evaluations.

C Utilisation of the evaluations

Monitoring the extent to which actions are taken based on the information contained in evaluations is another means of assessing the quality of the evaluations, as well as their usefulness. While DSW, CEG and NZES did not have formal systems for tracking the recommendations or actions arising from their evaluations, they were, to varying degrees, able to provide evidence on the extent to which their evaluations were utilised.

CEG provided clearly documented evidence of the actions taken as a result of the Mahi A Iwi and Pacific Island Peoples evaluations (Appendix 10). The results from the examination of actions arising from the two DSW and two NZES evaluations were mixed. Both agencies were able to provide the required information on only one of their evaluations.

4. CONCLUSION

As of 1 October 1998, the Work and Income New Zealand (WINZ) will be focused on returning people to work using regional flexibility as a way to tailor programmes to the needs of the local labour markets. As a result, regional commissioners will need to know what programmes work, in what combinations and for which target groups. They will need to know how much programmes cost and what operational issues need to be resolved. The need for high quality evaluation to determine what works and what does not will be absolutely essential.

Twenty-two of the 35 evaluations identified employment outcomes for participants. The highest percentages gaining employment were Job Plus and Job Skills participants, with 66% and 65% respectively. These were both wage subsidy programmes. The Compass (1995) and the Hikoi ki Pae-rangi programmes had the highest percentage of participants entering further training. These were both life-skills or motivational programmes.

However, we could not tell, with any certainty, what programmes were most effective at improving the employment prospects of which groups and under what circumstances. This was partly because agencies did not use consistent outcome measures and the emphasis on particular outcomes differed across the agencies. For example, some programmes were less focused on participants obtaining full-time employment than others. In addition, these evaluations were generally developed to meet the immediate information needs of internal and external stakeholders. Evaluators were not required to evaluate suites or combinations of programmes. However, if Regional Commissioners are to reliably assess which programmes work, in what combinations and for which target groups the following will need to be developed:

- consistent outcome measures, including those for training and attitudinal change as well as employment,
- predefined success criteria for programmes, and
- robust measures of the cost or cost-effectiveness of programmes.

Finally, assumptions about benefits of particular types of programmes for particular groups will need to be tested. For example, the assumption that life skills or motivational programmes benefit Maori, Pacific Island peoples, youth and women is untested.

The review identified a number of key operational issues evident in several evaluations that will need to be resolved if the aims and objectives of programmes are not to be undermined. The issues were:

- job seekers were not appropriately matched to programmes and interventions,
- the programmes often failed to meet the programme objectives or the intent of the programme,
- monitoring of the programmes and job seeker progress was frequently inadequate, or
- case management or post placement support was frequently inadequate.

It was more difficult to identify the key strengths across the evaluations because the evaluations tended to focus on the operational weaknesses. However, evaluations also provide an opportunity to identify programme strengths. Such information can be used to develop 'best practices'.

The quality of the evaluations differed within and between agencies. To improve future evaluations, several issues surrounding the quality of the evaluations will need to be addressed. In particular, future evaluations will need to:

- clearly link evaluation objectives to the rest of the report,
- appropriately identify and support the key findings,
- provide conclusions with a reportable message,
- state in the Terms of Reference whether or not recommendations are to be included in the report and, if so, provide recommendations of sufficient quality.

In looking across the agencies involved steps are being taken to increase the robustness of the evaluations and to assess them in a more strategic manner (e.g. the

use of control and treatment groups in the evaluation of the expanded Community Taskforce). Nevertheless, further steps need to be taken to address issues raised in this meta-evaluation.

The review recommends the development of an evaluation strategy which identifies priorities and timelines for selected strategic and operational evaluations along with the roles and project management responsibilities across the relevant agencies. The review of employment evaluations will be only one of several inputs to the development of an evaluation strategy. While the strategy would cover a period of three years it could be updated annually.

The strategy would also include ways to improve evaluation capacity. To this end, this review makes recommendations aimed at developing consistent definitions and measures of outcomes for participants in employment and training programmes. The development of appropriate impact measures for labour market and income support programmes, and criteria and guidelines for conducting evaluations are also recommended. The guidelines would build on existing guidelines such as those developed by LMPG and NZES. Finally, the review recommends that processes be put in place for tracking the recommendations or findings arising from evaluations. One option is to put in place a requirement that the evaluation units report annually to their agency CE on actions that have arisen as a result of their evaluations. This will encourage the creation of actionable recommendations where appropriate, and greater interaction between policy and evaluation units in developing recommendations and/or incorporating evaluation findings into future policy decisions.

5. RECOMMENDATIONS

It is recommended that:

1. DOL and SPA in consultation with WINZ and other relevant agencies work together to develop an evaluation strategy, dealing with employment and labour market issues over the next three years.
 - a) It is recommended that this strategy identify
 - a list of strategic and operational evaluations;
 - priorities and timelines;
 - the roles and project management responsibilities across the relevant agencies; and
 - ways to improve evaluation capacity.
2. The evaluation strategy include an evaluation of which types or combinations of programmes and interventions are most effective for which demographic or beneficiary groups.
3. DOL and SPA in consultation with WINZ and other relevant agencies to work together on the following capacity building projects:

- a) develop consistent definitions and measures of outcomes for participants in employment and training programmes prior to finalising Work and Income New Zealand's systems. Consideration will need to be given to how outcomes are measured at the regional level;
- b) investigate and develop appropriate impact measures, including cost-effectiveness, for labour market and income support programmes;
- c) put in place processes for tracking the recommendations or findings arising from their evaluations; and
- d) develop and distribute criteria and guidelines for conducting evaluations.

RELEASED UNDER THE
OFFICIAL INFORMATION ACT

REFERENCES

Australian National Audit Office (1997) *Programme Evaluation in the Australian Public Service*. Australian National Audit Office, Canberra.

Joint Committee on Standards for Educational Evaluation, 1994 *The Program Evaluation Standards: How to assess evaluations of educational programmes. Second Edition*. Sage Publications Inc, California.

Martin, John (1998) 'What works among active labour market policies: evidence from OECD countries' experience'. Paper presented at the conference on *Unemployment and the Australian Labour Market*, HC Coombs Centre for Financial Studies, Kiribilli, 9-10 June.

National Economic Research Associates (1995) *OECD Wage Subsidy Evaluations: Lessons for Workstart. Volume 1: Main Report*. Final Report for the Department for Education and Employment by NERA, London.

Funnell, Sue (1997) 'Making evaluation work for your organisation'. *Australian Evaluation Society Workshop*.

US GAO (March 1997) *General Policies/procedures and Communications Manual*

US GAO (April 1995) *Welfare to Work: Measuring outcomes for JOBS participants*. Report to the Ranking Monitoring Member, Committee on Finance, US Senate.

APPENDIX 1: CRITERIA

Programme type

1. *wage subsidy*: programmes where the job seekers wage is either fully or partially subsidised,
2. *work experience*: programmes that provide work experience but do not attract a subsidy (i.e. Community Taskforce),
3. *job-related training*: training programmes tied to a particular job or which provide vocational training,
4. *life skills/ motivational training*: programmes which provide job seekers with the necessary confidence, skills and knowledge look for work.
5. *job search assistance*: this refers to interventions which are aimed at improving job seekers chances of finding a job (i.e. work focus interviews; developing action plans; case management).
6. *community development*: programmes aimed at encouraging local economic development.
7. *support programmes*: this refers to programmes that assist low paid workers to maintain their jobs or provides support that enables job seekers to obtain employment (i.e. OSCAR which provides out of school care for the children of low income families).
8. *knowledge building*: research carried out to gain a better understanding of particular target groups or aspects of unemployment.
9. *strategic*: this refers to evaluations examining core business procedures (i.e. IEA).

Operational strengths and weaknesses

1. Recruitment, selection and preparation of participants in programmes
 - a) targeting (i.e. people, groups or areas selected)
 - b) recruitment and selection
 - c) pre-course preparation (i.e. provision of information about the course to participants; identification of participants needs prior to going on the course)
 - d) Marketing of the course/programme
2. The appropriate people skills and resources were resources were used:
 - a) Agency staff skills and resources
 - i) staff understanding of, and relationships with job seekers
 - ii) staff knowledge of the programme
 - iii) staff skills
 - iv) communications and relationship with providers
 - v) communication and relationship with other agency staff
 - vi) HO communications and relationship with field staff
 - b) Providers had the skills and resources to deliver the programme adequately
 - i) quality of facilitators, trainers, co-ordinators, providers etc

- c) Resources available for delivering the programme
 - i) resources available to the agency for delivering the programme (i.e. number of staff, time)
 - ii) resources available to providers for delivering the programme
- d) Timeliness (i.e. elements of the programme occurred on time; the programme occurred at the right time)
- e) The environment in which the course/ programme was delivered

3. Financial costs and benefits of the programme

4. Structure of the programme/course

- a) Goals and focus of the programme
- b) Flexibility within the programme to respond to differing needs of clients
- c) Extent to which cultural issues are addressed
- d) Duration
- e) Content of seminars/courses/workshop
- f) Number and range of places available
- g) Cost to job seekers of participating
- h) Subsidies or payments to employers or job seekers
- i) The extent to which the programme is linked with other programmes and supporting activities
 - i) Ease with which aspects of the programme could be used (i.e. plans)

5. Guidelines, criteria

- a) Quality or appropriateness of existing guidelines/ criteria
- b) Adherence to existing guidelines/ criteria
- c) Best practice guidelines or success factors

6. Monitoring

- a) Case management, follow-up, and post-placement support
 - i) existence and quality of guidelines
 - ii) frequency and type of contact with clients
 - iii) follow-up
 - iv) resources available for case-management (i.e. staff time)
 - v) perceived usefulness of case management and follow-up in assisting clients
 - vi) post-placement support

7. Tendering and contracting

8. Relevancy of the programme

- a) demand for the programme
- b) level of dead-weight, substitution and displacement
- c) participation rates of target groups or sections of the target group
- d) range of industries or providers using the programme
- e) uniqueness of the programme
- f) extent to which the programme is recognised by employers, education institutions and the community

- g) extent to the programme or aspects of the programme met the employment and training needs of the target group
 - i) training programmes
 - ii) life-skills programmes
 - iii) work-focus interviews
 - iv) work experience or wage subsidy programmes
 - v) impact on the prospects of clients who are particularly difficult to assist

9. General views of the programme

10. Organisational and management structures and processes

11. Operational aspects were not addressed in the evaluation.

Criteria for evaluating the quality of the evaluation

Completeness of reports

1. a table of contents that:
 - a) identifies the location of all important parts of the report
 - b) has page numbers or other clear identifiers
 - c) page numbers and other identifiers are correct and complete - they follow the structure of the report
2. objectives that are:
 - a) easily identifiable;
 - b) stated as neutrally as possible;
 - c) describe and explain any differences, if any, between the original objectives and those used in the evaluation.
3. a description of background to the evaluation that includes information on
 - a) what gave rise to the evaluation
 - b) how the evaluation will be used
 - c) who requested the evaluation and/or for whose use the evaluation is primarily needed
 - d) who is the primary audience
4. a description of the programme being evaluated that
 - a) is located near the beginning of the report
 - b) is located in one place
 - c) includes information on the background and context¹⁷ of the programme
5. a description of the scope of the evaluation that includes information on
 - a) the key objectives (refer above)
 - b) constraints on the scope of the evaluation

¹⁷ The context should include information on the following: geographic location of the programme, its timing, the political and social climate surrounding it, competing activities in progress, the staff, and the pertinent economic conditions.

- c) composition of samples, sample sizes, response rates etc,
 - d) lists of interviewees, their affiliations, reason for interviewing them
 - e) the time period covered by the data used in the analysis is stated
 - f) geographic locations and organisations included in the evaluation are identified, unless a valid reasons exists not to
6. a description of evaluation methodology that
- a) provides sufficient information to understand the methods being used including the type of instrument(s) being used and the method of data analysis
 - b) discusses any limitations of methods and implications for interpretation
7. key findings that
- a) have been appropriately identified (i.e. address the underlying problem)
 - b) state the criteria employed
8. conclusions that
- a) are clearly stated, not implied
 - b) have a reportable message
 - c) highlight evidence of significance (to get management to take corrective action)
9. executive summary that
- a) summarise the findings of full the report
 - b) is brief and strategic
 - c) provides sufficient but brief background information before the findings for the reader to understand what was being evaluated and why
 - d) includes recommendations of substance
 - e) includes major findings at minimum and recommendations where appropriate: clearly distinguishes between findings and recommendations
 - f) has a reportable message
10. inclusion of recommendations
- a) included
 - b) clearly identifiable as recommendations rather than embedded in an undifferentiated way in the findings

Correspondence between objectives and the rest of the report

- 1. all objectives should be addressed;
- 2. clear links between the objectives, methods of data collection, conclusions and recommendations; and
- 3. objectives should be explicitly addressed. For example, if the objectives related to efficiency, the information reported against it should be about efficiency not effectiveness
- 4. doable or any limitations are adequately described
- 5. are appropriate to the stage of development of the programme

Appropriateness and quality of methods of data collection

1. all methods used actually contribute to the evaluation and the way they do so is made clear;
2. design of the particular method is appropriate for its intended use (i.e. the sample size and composition are appropriate);
3. appropriate comparisons are used to make judgements;
4. the degree to which one can be confident in the data should be identified
5. a balanced view based on data collection and analysis is provided

Quality of the key findings

1. are supported by the evidence (i.e. effectively supported by appropriate levels of evidence¹⁸ and source(s)¹⁹)
2. contain no errors in logic or reasoning
3. answer the objectives
4. are appropriately qualified and cannot easily be misinterpreted or misquoted
5. identify the data sources on which the findings are based
6. are described adequately
7. identify the views of key stakeholders

Relationship of conclusions to data

1. relevant data or argument should be provided to support conclusions
2. the conclusions should follow from the data and links should be clearly stated (direction and intensity of conclusions should be based on the findings)

Quality of recommendations

1. flow logically from findings;
2. are practical in implementation;
3. identify responsibilities for implementation and time frames;
4. identify potential benefits and cost implications;
5. identify the root cause of the problem.

Presentation of report

1. is user friendly, well organised, accessible, good use of headings, sub headings etc
2. non-value adding repetition kept to a minimum
3. avoids excessive use of technical terms; includes glossary where appropriate
4. complementary use of narrative, graph, tables etc
5. is clear, concise and readable

¹⁸ Evidence must be competent, relevant, and sufficient to support findings, conclusions, and recommendations.

¹⁹ If there is reason to doubt the evidence it should be corroborated by other evidence.

APPENDIX 2: EVALUATION UTILISATION QUESTIONS

Questions to be asked of each service regarding the utilisation of the selected evaluations

- What *documented* actions arose from the evaluations?
- What were these actions based on:
 - recommendations in the evaluation?
 - recommendations developed by policy after the evaluation was completed?
 - other, specify?
- Is this the usual process?
- Is there documented evidence of the actions taken? (Demonstrate)
- If no actions were taken based on the evaluation, why was this the case?
- Were there undocumented benefits that arose from the evaluation?

Agencies selected

New Zealand Employment Service

Evaluations selected:

- An Evaluation of the Enterprise Allowance and the Capitalisation Option
- An Operational Review of the Youth Action Programme

Community Employment Group

Evaluations selected:

- Mahi A Iwi Evaluation, 1997
- Pacific Island People's Evaluation, 1997

Department of Social Welfare

Evaluations selected:

- Compass, 1995
- Training Incentive Allowance, 1995

APPENDIX 3: PROGRAMME INFORMATION

Table 8: Programme information

Evaluation	Programme Description	Target Group	Evaluation Completed	Agency
1. The Effectiveness of Access Training	ACCESS training aims to improve the job prospects of unemployed people, especially those who are disadvantaged in the labour market and for whom traditional training methods were unsuitable or unavailable. A broad spectrum of training was provided including life skills courses, vocational training and work based training.	Unemployed people, especially those who are disadvantaged in the labour market and for whom traditional training methods were unsuitable or unavailable	1992	ETSA
2. Evaluation of Community Taskforce (CTF)	CTF is a pilot programme that provides work experience to job seekers and enables projects of benefit to the community to be undertaken. Job seekers participate in CTF for 3 days per week. Community projects are between three and six months in duration. The Government's objectives were to increase the job readiness, personal self-esteem and motivation of participants while not displacing those in paid employment.	Job seekers registered 26+ weeks unemployed, aged 15 - 60 years.	1992	NZES
3. Evaluation of Taskforce Green (TFG)	TFG is a work experience programme, administered by NZES which provides full-time temporary jobs which give long-term unemployed job seekers skills and work habits to improve their chances of finding unsubsidised work, enable projects of environmental benefit to be undertaken, and enable projects of community benefit to be undertaken.	Those registered with NZES for 26+ weeks	1993	NZES
4. Training Opportunities Programme (TOP) Review	TOP is aimed at assisting the most seriously disadvantaged individuals in the labour market. The purpose of the programme is to provide second chance training for the unemployed with no or low qualifications and limited skills in order to assist them to achieve independence. TOP is the Government's largest labour market training intervention, providing around 15,000 places per year.	Young people and long-term registered unemployed with low qualifications.	1994	ETSA
5. The Behaviour and Perceptions of Long Term Job Seekers	NA. This was knowledge building survey.	-	1994	NZES
6. Evaluation of Limited Service Volunteer Scheme	Limited Service Volunteers is a six week programme of residential military training designed to build self-confidence and motivate unemployed job seekers.	18 -25 years olds who are unemployed	1994	NZES

7. Evaluating the Effectiveness of Job Plus	Job Plus is a wage subsidy programme which assists employers who hire people disadvantaged in the labour market for full-time permanent positions, via a wage subsidy. The duration of the subsidy is generally six months and the maximum is 12 months. Job Plus aims to facilitate entry or re-entry to the full-time permanent workforce.	Those registered with NZES for 26+ weeks	1994	NZES
8. Evaluation of Joint NZES/NZIS Work Focus Interviews	Joint work focus interviews were conducted by staff from both NZES and NZISS. The purpose of the trial was to investigate how interviews conducted jointly by staff from both services compared to normal work focus interviews conducted by NZES staff alone.		1994	NZES
9. Evaluation of TFG, Tourism Green and Possum Control	(TFG: see # 3). The Possum Control Project is a programme to use the long term unemployed to carry out labour intensive possum control and goat control where this is feasible, for TB and conservation related purposes, that would otherwise not occur. Tourism Green is a programme to establish fully funded work projects for DOC additional to its business plan and of benefit to the environment, employment and tourism.	Those registered with NZES for 26+ weeks	1995	NZES
10. Evaluation of Compass Pilot Programme	Compass assists sole parent beneficiaries to take steps towards employability and workforce participation by providing them with individual career counselling and help to access child care, education, training and employment opportunities.	Sole parents in receipt of the DPB or the Widows Benefit, particularly those who have been on the benefit for 1+ years and whose oldest child is 5+ years old.	1995	DSW
11. Evaluation of Training Incentive Allowance	The Training Incentive Allowance provides financial assistance in meeting study costs.	Those receiving the emergency maintenance allowance, the domestic purposes benefit, the widows benefit, and the invalids benefit.	1995	DSW
12. An Evaluation of Job Plus Training Pilot	Job Plus Training places disadvantaged job seekers into full-time work by overcoming their actual or perceived lack of skills or experience through providing targeted training that is usually tied to a specific vacancy.	Those registered with NZES for 26+ weeks or those otherwise disadvantaged in the labour market.	1995	NZES
13. Assessment of TOP: Survey of TOP Trainees	(Training Opportunities Programme: see # 4)	(see # 4)	1995	ETSA
14. Evaluation of Job Skills	The CHCH Youth Employment Initiative, called Job Skills, is a modification of the Job Plus wage subsidy programme which allows the CHCH city council's Youth Employment Service to provide fully subsidised workers to the private sector of 4 days per week.	Unemployed youth aged 16 - 24 years old	1995	NZES

15. Evaluation of Job Action	Job Action is an employment assistance strategy for long-term unemployed job seekers consisting of an interview, a workshop and follow-up case management.	The very long term unemployed (i.e. 104+ weeks unemployed).	1995	NZES
16. Evaluation of NZ Conservation Corps (NZCC) and Youth Services Corps (YSC) programmes	The NZCC aims to provide young people with opportunities for work skill acquisition and personal development through participation in conservation orientated projects which are (with one exception) 20 weeks in duration. The YSC applies the NZCC model to social services, such as restoration of war memorials. All projects include elements of community work, education, work experience and recreation, along with Te Ao Maori.	Young people aged 16-25 years old and registered unemployed for 13+ weeks.	1995	MOYA
17. Tane Atawhai Evaluation Report	Tane Atawhai involves participants attending a seminar designed to assist Maori men to recognise the barriers that they may have in obtaining work, and then to put these barriers into perspective.	Unemployed Maori men enrolled with NZES	1995	NZES
18. Barriers to Employment Facing Long-Term Job Seekers	NA. This was knowledge building survey.			
19. An Operational Review of the Youth Action Programme	Youth Action Programme is aimed at improving education, training and employment outcomes of eligible young people. Participants complete an action plan outlining what they will do to overcome their barriers to employment. EA's assist clients to find work and access employment programmes and training opportunities.	All 16-20 year old job seekers who had been registered unemployed for 13 or more weeks.	1996	NZES
20. Evaluation of Job Support Programme	The Job Support makes funding available to purchase support services and/or wage subsidies needed by a person with a disability to work in open employment, and receive the appropriate rate of pay for a job. Workbridge was contracted by NZES to run the Job Support.	Those who meet the following criteria as defined by Workbridge <ul style="list-style-type: none"> • have an identified disability • be taking up or in a job which is available to anyone in the community • be in self-employment where they are the main decision-maker and they receive the profits of the business. 	1996	NZES
21. Enterprise Allowance and the Capitalisation Option	The Enterprise Allowance is a programme that assists people into self-employment by learning the appropriate business skills. Eligible job seekers have access to a grant of up to \$600, and continue to receive income support while they are training. NZES assists the job seeker to develop a business plan, ensures the plan is vetted independently, and monitors the state of the business if the plan is approved and put into action. The capitalisation option operates in a similar manner, except the client receives an initial grant and a smaller cash flow during the subsidy period. As there are greater risks associated with this there are additional requirements that the job seeker	Those who meet the following criteria <ul style="list-style-type: none"> • work-tested beneficiaries; • registered unemployed for at least 15 weeks; • interested in self-employment; and • have a realistic chance of self-employment. 	1996	NZES

	invests \$1000 in the business and that the capitalisation is a last-resort funding, and these require additional checks at the beginning of the process.			
22. NZCC Two Year Post-Course Evaluation	(NZCC: see # 16)	(see # 16)	1996	MOYA
23. CEG Pacific Peoples Evaluation Report	The Pacific Island Employment Strategy aims to assist unemployed and disadvantaged Pacific Island people towards positive activity and opportunities for employment leading to self-sufficiency.	Pacific Island people within groups and communities that have high levels unemployment and disadvantage.	1997	CEG
24. Mahi A Iwi Evaluation Report	Mahi A Iwi: CEG uses a staircasing process to assist Maori towards increased labour force participation and self-sufficiency. CEG uses a holistic approach.	Maori within groups and communities that have high levels unemployment and disadvantage.	1997	CEG
25. Evaluation of Maori Youth Pilot	Maori Youth Pilot aimed at increasing self-esteem, instilling work habits and developing career goals through a one day seminar facilitated by NZES; a six week course delivered by Maori providers; and referral of the participants to other suitable training or work related activities	Maori job seekers registered with NZES and aged 16-24 years.	1997	NZES
26. Evaluation of OSCAR DAP Pilot	OSCAR programmes provide supervised out of school care and recreation services group and individual activities, and some provide opportunities for completion of homework. OSCAR is funded through the DAP.	Children aged 5-14 years, from low-income families.	1997	DSW
27. Evaluation of Wahine Pakari Programme	Wahine Pakari is a training programme designed to give Maori women skills to enter self-employment, businesses, further education and training, and employment.	Maori women	1997	NZES
28. Evaluation of IEA Pilot	IEA is a client focused model of service delivery		1997	NZES
29. Evaluation of Hikoi Ki Pae-rangi/New Horizons	The Hikoi Ki Pae-rangi seminar was developed as one of a range of pilot schemes designed to test the effectiveness of the facilitative approach in assisting sole parent beneficiaries into employment, education and training. The seminar was designed to target Income Support's Compass clients. The seminar addresses particular barriers faced by sole parents when considering entering or returning to the workforce. It aims to increase their knowledge, motivation and confidence. The seminar runs for two weeks, for four hours every day.	All women sole parents receiving an Income Support benefit for 12 months or longer.	1997	NZES
30. The Enhanced Job Action and Youth Action Programmes: An Evaluation	The Enhanced Job Action and Enhanced Youth Action programmes seek to assist registered unemployed job seekers to move into employment, training and education through a process of in-depth interviews and intensive case management between each participating job seeker and a NZES Employment Advisor.	Job Action: - all job seekers who reach 104 weeks registration of unemployment with the NZES and 5000 job seekers who reach 52 weeks. Youth Action: - all job seekers aged 16 to 20 who have been registered as unemployed for at least 13 weeks.	1997	NZES

31. Evaluation of Job Connection	Job Connection aims to assist very long-term job seekers by giving them an opportunity to work full-time for six months and gain skills and experience which will improve their chances of obtaining stable employment. Each person participating in Job Connection is placed on a fully funded work project for six months, receiving intensive supervision and assistance.	Very long-term job seekers (registered with NZES as unemployed for 208 weeks or more).	1997	NZES
32. Evaluation of Tama Tane O Le Pasifika	Tama Tane O Le Pasifika is a programme of five day seminars followed by case management. The aim is to help Pacific Island men into employment, education or training, by breaking down the barriers they face in seeking employment.	Long-term unemployed Pacific Island men, aged 35 and over.	1997	NZES
33. 1995-1996 Be Your Own Boss: National Outcomes Report	The Be Your Own Boss programme is an integrated package of business and enterprise skills training for unemployed and non-business wise people. The programme is provided by community based organisations on behalf of CEG.	Unemployed and non-business wise people	1997	CEG
34. A Quantitative Evaluation of Compass After Two Years	Compass (see # 10)		1997	DSW
35. BOOST	Boost is a pilot programme which has run in six Income Support districts. Income Support Customer Service Officer work with about 60 young people receiving Independent Youth Benefits to help them identify their goals and plan for their future.	Those in receipt of the IYB (i.e. 16-17 year olds).	1997	DSW

APPENDIX 4: THE FOCUS OF THE EVALUATIONS

Table 9: The focus of the evaluations

Evaluation	Outcomes: (summative) a	Process b	Cost	Basic knowledge building c
1. The Effectiveness of Access Training	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2. Evaluation of Community Taskforce	<input type="checkbox"/>	<input type="checkbox"/>		
3. Evaluation of Taskforce Green	<input type="checkbox"/>			
4. TOP Review		<input type="checkbox"/>	<input type="checkbox"/>	
5. The Behaviour and Perceptions of Long Term Job Seekers				<input type="checkbox"/>
6. Evaluation of Limited Service Volunteer Scheme	<input type="checkbox"/>	<input type="checkbox"/>		
7. Evaluating the Effectiveness Job Plus	<input type="checkbox"/>			
8. Evaluation of Joint NZES/NZIS Work Focus Interviews	<input type="checkbox"/>	<input type="checkbox"/>		
9. Evaluation of TFG, Tourism Green and Possum Control	<input type="checkbox"/>			
10. Evaluation of Compass Pilot Programme		<input type="checkbox"/>		
11. Evaluation of Training Incentive Allowance	<input type="checkbox"/>			
12. An Evaluation of Job Plus Training Pilot	<input type="checkbox"/>	<input type="checkbox"/>		
13. Assessment of TOP: Survey of TOP Trainees		<input type="checkbox"/>		
14. Evaluation of Job Skills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
15. Evaluation of Job Action	<input type="checkbox"/>	<input type="checkbox"/>		
16. Evaluation of NZCC and YSC programmes	<input type="checkbox"/>			
17. Tane Atawhai Evaluation Report		<input type="checkbox"/>		
18. Barriers to Employment Facing Long-Term Job Seekers				<input type="checkbox"/>
19. An Operational Review of the Youth Action Programme	<input type="checkbox"/>	<input type="checkbox"/>		
20. Evaluation of Job Support Programme		<input type="checkbox"/>		
21. Enterprise Allowance and the Capitalisation Option	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
22. NZCC Two Year Post-Course Evaluation	<input type="checkbox"/>	<input type="checkbox"/>		
23. CEG Pacific Peoples Evaluation Report		<input type="checkbox"/>		
24. Mahi A Iwi Evaluation Report		<input type="checkbox"/>		
25. Evaluation of Maori Youth Pilot	<input type="checkbox"/>	<input type="checkbox"/>		
26. Evaluation of OSCAR DAP Pilot	<input type="checkbox"/>	<input type="checkbox"/>		
27. Evaluation of Wahine Pakari Programme		<input type="checkbox"/>	<input type="checkbox"/>	
28. Evaluation of IEA Pilot		<input type="checkbox"/>		
29. Evaluation of Hikoi Ki Pae-rangi/New Horizons	<input type="checkbox"/>	<input type="checkbox"/>		
30. The Enhanced Job Action and Youth Action Programmes: An Evaluation	<input type="checkbox"/>	<input type="checkbox"/>		
31. Evaluation of Job Connection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
32. Evaluation of Tama Tane O Le Pasefika	<input type="checkbox"/>	<input type="checkbox"/>		
33. Be Your Own Boss: National Outcomes Report	<input type="checkbox"/>			
34. A Quantitative Evaluation of Compass After Two Years	<input type="checkbox"/>	<input type="checkbox"/>		
35. BOOST	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
TOTAL	24	27	7	2

- a) Summative evaluations are focused on measuring the impact of a programme, to make a judgement about whether or not the programme is meeting its objectives. Summative evaluations rely on quantitative methods such as surveys and analysis of administrative databases, in order to measure participants' outcomes relative to some counterfactual. The

counterfactual could be outcomes for a group of non-participant, outcomes for a group before the intervention is introduced, or some other construct.

- b) Process evaluations are usually focused on how a programme is being implemented. This type of evaluation can be used to test whether a programme is being implemented as intended, to improve its delivery, to understand better what elements or aspects of a programme are contributing to particular outcomes, or to determine how a programme works as a basis for future evaluations. Process evaluations may use a range of qualitative and quantitative methods, such as interviews, focus groups or surveys of administrators, programme participants and other stakeholders; analysis of administrative data or other records; direct observation; or in-depth case studies.
- c) This refers to two surveys focusing on the long-term unemployed.

RELEASED UNDER THE
OFFICIAL INFORMATION ACT

APPENDIX 5: PAPER ON EVALUATION METHODS

Evaluating Employment Interventions: Choosing the Appropriate Method

Julian Silver

Draft 3 June 1998

Policy Context

Forthcoming policy changes - including the structural integration of NZES, CEG and IS; the application of work capacity assessments, work testing and access to facilitative assistance to a wider group of beneficiaries; and the location of greater accountability for determining strategies and achieving results at a regional level - have implications for evaluation. In particular, these changes raise important issues about evaluation priorities, capacity and methodologies. For example:

- greater regional flexibility and discretion in determining the balance and mix of employment-related interventions needs to be backed up by good information on what works and what does not, and on the cost-effectiveness of alternative strategies. This sort of information should inform resource allocation decisions at a national, regional and site level; improve delivery and responsiveness; and form a part of the accountability framework;
- greater regional discretion offers the opportunity for piloting or trialing new strategies, programmes, and delivery approaches. For these pilots to be useful, they need to be properly evaluated;
- the emphasis on case management and other forms of "individualised" assistance, requires a shift away from evaluating discrete "programmes" and towards evaluating the impact of broader strategies or inter-linked series of interventions on individuals' labour market outcomes;
- the emphasis on holding the delivery department more accountable for results (as opposed to activities) requires a more rigorous understanding and evaluation of the assumed chain of causality between specific activities, short-term and intermediate outputs and outcomes, and ultimate outcomes.

This paper briefly summarises the debate on appropriate evaluation methodologies, particularly on the appropriateness of experimental methods, and concludes with some observations on how different methods can complement each other given the policy context outlined above. This section is not a technical guide, and further work will be needed to develop appropriate criteria and methodologies for evaluation in the new environment.

Types of Evaluation

Broadly speaking, evaluations can be grouped into two types:

- (a) *process* - these evaluations are usually focused on how a programme is being implemented. This type of evaluation can be used to test whether a programme is being implemented as intended, to improve its delivery, to understand better what elements or aspects of a programme are contributing to particular outcomes, or to determine how a programme works as a basis for future evaluations. Process evaluations may use a range of qualitative and quantitative methods, such as interviews, focus groups or surveys of administrators, programme participants and other stakeholders; analysis of administrative data or other records; direct observation, or in-depth case studies;
- (b) *summative* - these evaluations are focused on measuring the impact of a programme, to make a judgement about whether or not the programme is meeting its objectives. Summative evaluations rely on quantitative methods such as surveys and analysis of administrative databases, in order to measure participants' outcomes relative to some counterfactual. The counterfactual could be outcomes for a group of non-participant, outcomes for a group before the intervention is introduced, or some other construct. Ideally, a summative evaluation comprises a number of steps, set out below. The OECD notes that most summative evaluations answer only the first question; those answering the second do so in a partial manner; and very few address the third:²⁰
 - (i) what are the estimated impacts of the programme on the individual? This implies a counterfactual, i.e. what would have happened to the individual in the absence of the programme?
 - (ii) are the impacts of the programme large enough to yield net benefits? Costs and benefits can be measured in different ways, which need to be made clear in the evaluation objectives. For example, measuring the *fiscal impact for government* of an intervention can be distinguished from measuring the *economic impact for society*. Collecting sufficient information to measure economy-wide costs and benefits (eg: displacement effects, improved health or crime indicators) is usually problematic;
 - (iii) is this the best outcome that could have been achieved for the money spent? This implies a judgement about how significant the impact of the programme is, relative to alternative approaches or expectations.

There are two important methodological questions facing evaluators and their stakeholders:

- (a) what is the best method to measure impact?
- (b) what is the appropriate balance between summative and process evaluation?

²⁰ R. Fay, *Enhancing the Effectiveness of Active Labour Market Policies: Evidence from Programme Evaluations in OECD Countries*, OECD, May 1996.

Evaluation Stakeholders

The appropriate evaluation method will depend to a large part on the type of questions being asked, which will in turn depend on the audience and stakeholders. For Ministers, there will be occasions on which a decision is required on whether or not to continue a programme. This will usually be in the context of a pilot or when a specific appropriation expires. Often, Ministers and policy/ purchase advisors will want to use evaluations to make marginal decisions about the appropriate scale of a programme, the best mix of interventions, or to inform policy changes.

Evaluations can also be used by Ministers and policy/ purchase advisors to test the linkages between the results for which the department is being held accountable, and government's broader outcomes; and by policy/ purchase advisors to test the linkages between the different levels of results for which the department is being held accountable.

Under the new environment, Regional Commissioners and front-line staff will also be making decisions about the most appropriate mix of interventions, and on programme design (within broad parameters). Evaluations will provide some general information about "what works and what doesn't". The delivery agency at all levels should also be using evaluation findings to understand why a particular intervention is or is not working, and to improve delivery.

Impact Measurement

The evaluation literature focuses on two approaches to measuring impact for programme participants:

- (a) *experimental* - with this type of evaluation, the only factor differentiating programme participants and a control group (of non-participants) is random allocation between the two groups;
- (b) *quasi-experimental* - with this type of evaluation, random allocation does not occur, and attempts are made to control for the differences between the participant and the control (or, more accurately, comparison) group using statistical methods.

Quasi-experimental designs can take a number of different forms, ranging from simple "before and after" studies, or comparative outcomes for participant and comparison groups post the introduction of an intervention; to analysis of time series data or comparative "before and after" outcomes for participant and comparison groups; to econometric modelling (such as using regression analysis to estimate the differences between outcomes observed for participants, and those predicted on the basis of a model).²¹

There has been vigorous debate about the pros and cons of experimental methods. Reacting, in part, to studies which showed that the reliability of results from quasi-experimental evaluations are highly dependent on how the comparison group is chosen and on the types of statistical methods used, the US Government from 1986

²¹ For a summary of different approaches, see L. Mohr, *Impact Analysis for Program Evaluation*, Sage, 1995.

has only funded experimental evaluations.²² At the same time, however, questions have also been raised about whether the preponderance of the “experimental orthodoxy” in the US has been at the expense of other potentially useful methods that can provide insights into programme effectiveness.²³

The “experimental orthodoxy” in the US (typified by such large-scale studies as the Manpower Demonstration Research Company’s evaluations of welfare-to-work programmes) has not taken the same hold in other OECD countries, although there are a growing number of examples. The current evaluation of the expanded Community Taskforce programme is the first use of random assignment in the evaluation of a labour market programme in New Zealand, and will be a valuable test of the advantages and disadvantages of this approach.

The advantages and disadvantages of the two approaches are briefly summarised below.²⁴ In addition, there are a number of methodological and analytical issues common to both approaches.

Experimental Methods

The major advantage of the experimental approach is that it minimises selection biases and other matching problems. While it is possible, using quasi-experimental methods, to match participants and non-participants so that differences in observable characteristics (such as age, education, duration of unemployment, local labour conditions) are minimised, it is not possible to predict or match for all the relevant characteristics that are likely to affect participation and outcomes, or to control for all these variables. For example, while a highly motivated job seeker is more likely to seek to participate on a programme (for non-full coverage programmes), to be selected by an employer/ sponsor, and to have good outcomes, it is not possible, *a priori*, to match on degree of motivation (although various proxies, such as prior voluntary participation in programmes, or information gathered during assessment processes, could be used).

With an experimental approach, participation and non-participation is random, so any non-observable characteristics are randomly distributed as well. Given less reliance on potentially complex or discretionary matching procedures and model specification, an experimental approach is likely to produce more robust estimates of programme effects than a quasi-experimental approach.

On the other hand, a number of potential drawbacks with an experimental approach have been identified.

²² For example, T. Fraker and R. Maynard, *The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programmes*, The Journal of Human Resources, XXII, no. 2, 1987.

²³ For example, see L. Mead, *Optimising JOBS: Evaluation versus Administration*, Public Administration Review, vol 57, no. 2, 1997.

²⁴ This discussion draws on Fay (1996); Mead (1997); G. Burtless and L. Orr, *Are Classical Experiments Needed for Manpower Policy?*, The Journal of Human Resources, XXI, no. 4, 1986; T. Eardley and M. Thompson, *Does Case Management Help Unemployed Job Seekers? A Review of the International Evidence*, Social Policy Research Centre Reports and Proceedings, no. 132, 1997; and P. Rossi and H. Freeman, *Evaluation: A Systematic Approach*, Sage, 1985.

Firstly, experiments are very costly and can be difficult to implement. High implementation costs arise from the need to set up mechanisms to allocate and monitor participants and controls, to communicate to and train staff, and to establish specific data collection instruments for both participants and controls (rather than, for example, drawing the comparison group from existing databases). Implementation problems will arise if staff are not fully trained in the processes to be used, are not supportive, or are facing conflicting incentives (i.e. being required to meet high participation targets, at the same time as maintaining a control group who are not permitted to participate).

Implementation and design will be problematic if the programme being evaluated has several components or screening points. In these cases it is particularly important to be clear about what constitutes “the programme” (i.e. at what point random allocation occurs), who comprise the experimental and control groups, what groups outcomes are being measured for, and what hypotheses are being tested.

For example, Gueron and Pauly describe experimental methods for two different types of labour market programme, with different implications for the point at which random allocation occurs, and hence the groups for whom outcomes are measured.²⁵ “Broad coverage” programmes usually involve a number of components (i.e. job search followed by case management followed by training) which are mandatory for an identifiable group. For these programmes, random assignment occurs early in the process (amongst the mandatory group). The corollary is that average outcomes are measured across a broad group of people, only some of whom will have participated in particular sub-components.

On the other hand, “selective-voluntary” programmes involve individual programmes or components to which people are referred or self-select. For these programmes random assignment occurs once people have been referred or self-selected, and average outcomes are measured for this group. Some practical difficulties can be envisaged with the latter approach, particularly if a sub-group of people who are assessed as suitable for, or volunteer to go on, a programme are then told that they are in the control group. On the whole, random assignment is easier to manage and justify to clients and other stakeholders where the intervention being tested is mandatory.

Secondly, it is often claimed that ethical considerations preclude random assignment, particularly if potentially beneficial services are being withheld from the control group. However, in those (frequent) circumstances where eligible applicants exceed programme places, or where there are waiting lists, or where a new programme is being trialed, random assignment could be seen as a legitimate approach to rationing.²⁶ Furthermore, it cannot be assumed that every new programme being trialed (or even existing programmes, if they have not been evaluated before) will

²⁵ J. Gueron and E. Pauly, *From Welfare to Work*, Russell Sage Foundation, NY, 1991

²⁶ Burtless and Orr also discuss the more stringent ethical requirement that each subject, including each control, should consent to participate in the experiment. Such a requirement is likely to be most problematic where experimental methods are used to evaluate an existing programme where some people may be made worse off through losing access to services or an entitlement. In this case, Burtless and Orr propose a system of compensatory payments to people in the control group. It is unclear from evaluation reports whether this approach is used, and the reports themselves usually have little to say about ethical issues.

have a positive effect on participants. The possibility that a new programme may not be of benefit to participants is all the more reason for evaluating it in a rigorous way. In medical research, where the up-side and down-side risks are generally much higher than for labour market programmes, experiments are the standard research method, governed by strict protocols.

Thirdly, the tight controls and “artificial” conditions involved with running experiments may limit the extent to which any findings can be replicated in other sites or nation-wide, and may introduce other sorts of selection biases. The experiment may disrupt or change the patterns of demand for, or methods of delivery of, other services in a site. The nature of the experiment may deter job seekers, sponsors, providers or particular site administrators from becoming involved. The outcomes that are found for participants in an experiment (relative to the control group) may not be replicated when a programme is expanded to everyone in the eligible population, as there may simply be more eligible participants competing for a limited pool of outcome opportunities, which reduces average outcomes per participant.

Quasi-Experimental Methods

The major advantage of the quasi-experimental approach is that it is typically less costly and administratively demanding than an experimental approach. Depending on the sophistication of the econometric modelling techniques used, a quasi-experimental approach may also be able to produce a broader range of estimates than an experimental approach, such as estimates of the marginal effects of expanding or contracting a programme.

On the other hand, there are methodological problems with quasi-experimental approaches. As discussed above, a major problem is one of selection biases. Selection biases will be less severe if potential participants do not have the opportunity to select in or out of the programme, for example where the programme is mandatory for all of the eligible population in a site, or where measuring take-up amongst an eligible population is an objective of the evaluation.

A further problem is that the results of quasi-experimental evaluations are highly dependent on the way that comparison groups are selected and impacts estimated. There are many reasons why a comparison group may prove less than satisfactory. For example, “before and after” studies may be affected by external changes (such as changes in policy or in economic conditions). Comparison sites may be affected by unforeseen changes in labour market conditions or difficulties in matching on relevant variables. Information on important matching characteristics may not be available, or may not be known to be important at the time the comparison group is selected. The participant and comparison groups may be “contaminated” by people moving between them.

Many of these problems could be controlled for statistically, or through good design (eg: choosing a number of comparison sites in case external shocks affect one site), although it requires good information, high levels of expertise, and a recognition that estimates will be sensitive to the design. Few evaluations of labour market programmes have been done in New Zealand using of these more sophisticated quasi-

experimental evaluations, and those that have (such as Compass and Job Action) have been complex and at times difficult to interpret.

At the other extreme, there may be cases where a suitable comparison group cannot be found, due to likely high biases, administrative complexity, cost, or other factors. For example, finding a register-based comparison group for the Job Plus evaluation was deemed not feasible, as virtually all the long-term unemployed who were placed into work were done so through a Job Plus wage subsidy. While the absence of a comparison group severely limits the summative conclusions that can be drawn, there are a range of other standards against which the findings from these sorts of descriptive evaluations can be compared, eg: findings from similar programmes (i.e. were the employment outcomes from Job Plus within the range found from wage subsidy programmes in other OECD countries?), or expectations that are set when the programme is introduced, or outcomes derived from monitoring data.

Limits to Summative Evaluation

In addition to the issues discussed above, there are a number of limitations common, in varying degrees, to both experimental and quasi-experimental evaluations:

- summative evaluations typically rely on data gathering instruments such as surveys or administrative databases. Response biases or other data limitations will reduce the accuracy of results;
- summative evaluations require sufficient time to be able to measure outcome results and trends. This is often at odds with the need to fine-tune or more significantly modify programmes as they are being operated. It can also happen that subsequent policy decisions render a summative evaluation obsolete;
- because summative evaluations usually deal with programmes that are constrained in some way (i.e. by being run in a limited number of sites, for a set time period, for limited numbers of participants, or under tight administrative procedures), there are questions about the replicability and generalisability of results;
- summative evaluations treat the programme as a “black box”; ie: the evaluation measures the impact of a programme on participants’ outcomes, but does not attempt to explain why it produces these results, or how particular changes to the programme would affect outcomes. Summative evaluations are also unlikely to estimate the net impact of different stages or components of complex labour market programmes, or of policy and programme alternatives, unless a series of control or comparison groups are able to be constructed (which adds to expense and administrative complexity);
- summative evaluations (particularly experiments) usually measure impact for individuals. More sophisticated methods are required to estimate economy-wide effects, or to estimate the marginal effects of changes to the scale or mix of particular programmes. Summative evaluations, on their own, will provide little guidance about what sorts of policy changes might be appropriate.

The Role of Process Evaluation

A number of the problems listed above can be mitigated through the use of process evaluation as part of the overall evaluation design. Process evaluations are particularly useful for getting inside the “black box”, to assess whether it is being implemented as planned, and to determine why a particular intervention is producing particular outcomes, for particular groups of participants. For example, are poor outcomes predominantly due to implementation failures (such as inadequate training or systems, or lack of knowledge amongst the target group), or to flaws in policy design? Are particular elements of the programme affecting outcomes? Process evaluations will also be useful for fine-tuning new programmes, or for generating hypotheses for future evaluations.

The point is made in a number of studies that while summative evaluations work well for national programmes delivered within well-specified bounds, the greater the variability or discretion in policy and delivery, the greater the need for process evaluations to aid the interpretation of results.²⁷ Given the shift in New Zealand towards greater decentralisation, flexibility and individually-tailored assistance, there is a strong case for a process evaluation to form a part of any significant evaluation.

Conclusion

This paper has discussed how different evaluation methodologies can complement each other. The chosen methodology will depend on the sort of programme being evaluated, the phase of the programme during which the evaluation is occurring, and the types of questions being asked. While summative evaluations are intended to answer the question “does the programme work?”, process evaluations are required to answer the questions “why does the programme work or not work?” and “how can the programme be improved?” With regions having greater flexibility to modify programme criteria and delivery, and with programmes themselves becoming more multi-faceted, a good understanding of how processes affect results for particular clients, in particular regions or communities, is essential.

For summative evaluations, the international experience is that experimental methods produce more accurate results than quasi-experimental methods. However, there are high costs associated with experiments. The trade-off between accuracy and cost also needs to consider those factors which may reduce the accuracy, and hence the relative cost-effectiveness, of an experiment, such as response biases from surveys or poor generalisability.

Experiments have seldom been tried outside the US, but there seem few reasons why they should not be utilised, under appropriate circumstances, in New Zealand. Experiments require clear protocols (including informing people that they are participating in an experiment), close monitoring, and buy-in by Ministers, front-line staff and other stakeholders.

²⁷ See J. Stern et al, *OECD Wage Subsidy Evaluations: Lessons for Workstart*, NERA, 1995. This meta-evaluation quotes a Dutch evaluation which finds variations in implementation methods for the scheme across offices as leading to variations in placement rates of +/- 30%. Stern et al also note that very few wage subsidy evaluations attempt to cover these process issues.

Experiments seem best suited to pilots, and to relatively simple programmes, and are less suited to programmes that are already running or to those that consist of a number of stages or elements. On the other hand, pilots provide the necessary opportunity for fine-tuning the programme, which would undermine any summative evaluation. One solution to this dilemma is to break a pilot up into phases - a process evaluation (resulting in any necessary modifications) followed by an experimental evaluation.

While quasi-experiments are usually less costly and administratively demanding than experiments, their usefulness as a potential evaluation method will depend on factors such as judgements about the likely degree of selection bias, the information available for constructing comparison groups and other modelling purposes, and access to the skills needed to do the evaluation.

Whatever approaches are used, there are a number of more general principles which need to underpin good evaluation in the new policy environment:

- careful thought needs to be given to appropriate evaluation methods. There are opportunities for more rigorous use of both experimental and quasi-experimental methods, complemented by process evaluation. Where possible, a range of different methods should be used, to enable triangulation of different information;
- high quality data are essential, particularly data on outcomes and costs. Measures and definitions must be consistent if evaluations are to produce results that can be compared across different interventions, for participants with different characteristics, across regions or over time. These data need to be accessible to the people within the new department who are making decisions about resource allocation;
- opportunities should be taken to design and use monitoring data to support evaluations. For example, trends in register-based outcomes (such as exit rates) for different sorts of interventions could be used to measure how outcomes differ over the business-cycle; how outcomes are affected by expansions or contractions in participation or spending; or how participation in one programme affects outcomes for others (ie: substitution and dead-weight effects);
- it needs to be clear, when setting up the evaluation, what hypotheses are being tested. In particular, what is the counterfactual?
- there should be some criteria or benchmarks for setting hypotheses and assessing the significance of results. Ideally, these criteria should be established as part of the policy development process. This may simply involve a literature review or a meta-evaluation to determine what sort of results could be expected for different types of programmes in different contexts²⁸;
- evaluation requirements need to be built into policy development and, most importantly, into the way that programmes are implemented. The selection of trial sites or roll-out strategies, the design of information systems, and the development

²⁸ Mead (1997) argues that where an existing body of evidence is unambiguous on the impact of particular types of interventions (as he states it is for a "mandatory, work focused programme"), further experimental evaluations are unnecessary, and the focus should instead be on better use of administrative data to evaluate outcomes for different types of participants and different sub-programme components, using quasi-experimental techniques.

of guidelines and performance measures all provide valuable and unique windows of opportunity for evaluation.

Further Work

As noted, this paper is a discussion of issues, rather than a technical guide. Further work is needed to develop a more rigorous evaluation framework. Areas for further work include:

- developing criteria for selecting appropriate methodologies;
- developing a tool kit of analytical methods (including access to relevant expertise);
- defining “costs” and “benefits” for the purposes of impact assessment;
- specifying data requirements for monitoring and evaluation;
- defining outcomes and other key variables; and
- more clearly specifying what a programme or intervention is trying to achieve (e.g. enhance employability, increase access to work, increase people’s ability to stay in work).

APPENDIX 6: EMPLOYMENT OUTCOMES

Table 10: Employment outcomes

Programmes	N =	Full-time (%)	Part-time (%)	Temporary (%)	Short duration (%)	Self employed (%)	Unspecified employment outcome (%)	Duration over which outcomes were measured
Wage subsidy and work experience programmes								
□ ²⁹ Community Taskforce, 1992	439							0-8 Mths
A. CTF unsubsidised							A. 16.6	
B. subsidised							B. 9.6	
C. self-placed							C. 2.5	
D. Comparison unsubsidised							D. 5.4	
E. subsidised							E. 3.8	
F. self-placed							F. 3.2	
□TFG, 1993	828							3 Mths
A. TFG unsubsidised		A. 5.3	A. 4.1	A. 0.4	A. 4.2		A. ---	
B. TFG subsidised		B. ---	B. ---	B. ---	B. ---		B. 1.2	
C. TFG self-placed		C. ---	C. ---	C. ---	C. ---		C. 5.6	
D. Comparison group: No work unsubsidised		D. 2.1	D. 2.4	D. 0.4	D. 3.3		D. ---	
E. No work: subsidised		E. ---	E. ---	E. ---	E. ---		E. 1.4	
F. No work: self-placed		F. ---	F. ---	F. ---	F. ---		F. 3.0	
G. Comparison group: Some work unsubsidised		G. 2.5	G. 1.8	G. 0.3	G. 4.3		G. 1.1	
H. Some work: subsidised		H. ---	H. ---	H. ---	H. ---		H. 7.4	
I. Some work: self-placed		I. ---	I. ---	I. ---	I. ---		I. 7.1	
Job Plus, 1994	400							6 Mths
A. Job Plus							A. 65.0	

²⁹ □ = evaluations that relied on administrative data to determine outcomes.

□TFG, Tourism Green and the Possum Control,
1995

6 Mths

A. TFG 473 (TFG)
B. Tourism Green
C. Possum Control

A. 33.4
B. ---
C. ---

□Job Skills, 1995	395						66.0 ³⁰	6 Mths
-------------------	-----	--	--	--	--	--	--------------------	--------

Job Support Programme, 1996	406	-	-	-	-	-	82.0 ³¹	
-----------------------------	-----	---	---	---	---	---	--------------------	--

□EA and the Capitalisation Option, 1996	Not Reported							
---	--------------	--	--	--	--	--	--	--

□Job Connection, 1997	476							3 Mths
-----------------------	-----	--	--	--	--	--	--	--------

A. Job Connection A. 47.0³²
B. Comparison B. 7.0

Job-related training programmes								
---------------------------------	--	--	--	--	--	--	--	--

Access Training programme, 1992	1000							6 Mths
---------------------------------	------	--	--	--	--	--	--	--------

A. Access A. 41.0
B. JSR unsubsidised B. 35.0
C. JOS unsubsidised / subsidised C. 65.0

TOP, 1994 ³³	Not Reported							
-------------------------	--------------	--	--	--	--	--	--	--

□Job Plus Training Pilot, 1995	Not Reported							
--------------------------------	--------------	--	--	--	--	--	--	--

Be Your Own Boss, 1997	6,516						17.0	4-8 Mths
------------------------	-------	--	--	--	--	--	------	----------

Life-skills/motivational training programmes

□LSV Scheme, 1994								2.5 Mths
-------------------	--	--	--	--	--	--	--	----------

A. LSV A. 14.0
B. Control B. 7.0
C. NZCC C. 40.4

Compass Pilot Programme, 1995	Not Reported							
-------------------------------	--------------	--	--	--	--	--	--	--

³⁰ 66% remained with their Job Plus employer or got another job.

³¹ Of this number 44% were already in employment when they first received Job Support.

³² Of the 47% who were employed full-time 3 months after the completion of their Job Connection placement, 25% were employed without any subsidy from NZES. The remaining 22% were employed with the help of an NZES subsidy.

³³ This evaluation did not report on employment and further training outcomes for TOP participants. However, such outcomes are monitored continuously, reported to the Minister, and published annually in ETSA's annual accounts and other Agency publications.

Training Incentive Allowance, 1995	1203	2.40	13.8						
NZCC and YSC programmes, 1995									
A. NZCC immediate	A. 161	A. 25.2	A. 5.0				A. s 3.4	immediate	
B. NZCC 1 month	B. 159	B. 28.9	B. 0.6				B. s 1.6	1 Mth	
C. YSC immediate	C. 113	C. 28.3	C. 0.9				C. s 5.3	immediate	
D. YSC 1 month	D. 113	D. 28.3	D. 0.9				D. s 5.3	1mth	
□Tane Atawhai, 1995	737	14.0	2.0	5.0	13.0			6 Mths	
NZCC, 1996	45	46.0	6.0					2 Yrs	
Maori Youth Pilot, 1997	190								
A. After MYP							A. 27.0		
B. Before MYP							B. 6.0		
Wahine Pakari Programme, 1997	202							0-12 Mths	
A. After Wahine Pakari						A. 3.0	A. 40.0 ³⁴		
B. Before Wahine Pakari							B. 17.0		
□Hikoi Ki Pae-rangi/New Horizons, 1997	423								
A. Completed (3 Months)							A. 3.5		
B. Not Complete							B. 2.0	3mths	
C. Did not attend							C. 1.5		
D. Completed (6 Months)							D. 5.8		
E. Not Complete							E. 2.0	6mths	
F. Did not attend							F. 4.9		
Tama Tane O Le Pasefika, 1997	354							6 Mths	
A. Tama Tane		A. 4.0					A. 1.0		
B. Comparison		B. 2.0							
Compass 1997	1142								
Boost 1997									
A. Boost (former customers)	A. 98	A. 20	A. 6	A. 2					
B. Non-Boost	B. 105	B. 20	B. 13	B. 6					
Job search assistance									
□NZES/NZISS Work Focus Interview								3mths	

³⁴ 40% employment includes both full and part-time work.

A. Joint interview	A. 115	C. 11.3	C. 0	C. 0.9	A. 3.5	
B. Control	B. 115	D. 6.1	D. 2.6	D. 4.3	B. 2.6	
<input type="checkbox"/> Job Action, 1995	647					6 Mths
A. Job Action		A. 22.0	A. 21.0			
B. Comparison		B. 17.0	B. 23.0			
<input type="checkbox"/> Youth Action Programme, 1996	8425	24.5	4.5	4.4	6.3	- 0-14mths
<input type="checkbox"/> Job Action and Youth Action, 1997						5 Mths
A. JA pre-enhancement	A. and B. 16723	A. 23.1	A. 0.5	A. 3.2		
B. JA enhancement		B. 14.9	B. 0.3	B. 2.7		
C. YA pre-enhancement	C. Not Reported	C. 25.8	C. 0.6	C. 3.5		
D. YA enhancement	D. Not Reported	D. 21.0	D. 0.5	D. 3.4		
Community Development						
Mahi A Iwi, 1997	Not Reported				NA. This was a process evaluation	
CEG Pacific Peoples, 1997	Not Reported				NA. This was a process evaluation	
Other						
IEA Pilot, 1997	Not Reported				NA. This was a process evaluation	
OSCAR DAP Pilot, 1997	Not Reported					0-12mths

APPENDIX 7: TRAINING OUTCOMES

Table 11: Training outcomes

Programme	Training Outcomes		
	N =	% entering further training	Duration over which outcomes were measured
Wage Subsidy and Work Experience Programmes			
Community Taskforce, 1992	439		0-8 Mths
A. CTF		A. 6.4	
B. comparison		B. 4.8	
TFG, 1993	828		3 Mths
A. TFG		A. 4.7	
B. No work experience		B. 4.2	
C. Some work experience		C. 4.5	
Job Plus, 1994	400	-	6 Mths
TFG, Tourism Green and the Possum Control, 1995	TFG 473	Not Reported	6 Mths
Job Skills, 1995	395	4.0	6 Mths
Job Support Programme, 1996	375	Not Reported	
EA and the Capitalisation Option, 1996	Not Reported	Not Reported	
Job Connection, 1997	476	2.0	3 Mths
Job-related training			
Access Training programme, 1992	1000		0-7 Mths
A. Access		A. 11.0	
B. JSR unsubsidised		B. 9.0	
C. JCS unsubsidised / subsidised		C. 3.0	
TOR, 1994	Not Reported		
Job Pilts Training Pilot, 1995	Not Reported	100.0	
Be Your Own Boss, 1997	6516	7.0	4- Mths
Training: life-skills, motivational			
LSV Scheme, 1994	140		2.5 Mths
A. LSV		A. 6.0	
B. Control		B. 3.4	
C. NZCC		C. 18.1	
Compass Pilot Programme, 1995	Not Reported		
A. Compass volunteers		A. 64.9	
B. Target, Never on Compass		B. 39.2	
C. Non Target		C. 38.8	
Training Incentive Allowance, 1995	1203	37.6	
NZCC and YSC programmes, 1995			
A. NZCC immediate	A. 161	A. 20.2	immediate
B. NZCC 1 month	B. 159	B. 22.0	1 Mth
C. YSC immediate	C. 113	C. 36.3	immediate
D. YSC 1 month	D. 113	D. 36.3	1mth
Tane Atawhai, 1995	737	7.0	6 Mths
NZCC, 1996	45	18.0	2 Yrs
Maori Youth Pilot, 1997	190	21.0	
Wahine Pakari Programme, 1997	202	26.0	0-12 Mths

Hikoi Ki Pae-rangi/New Horizons, 1997	421	57.0	6mths
Tama Tane O Le Pasefika, 1997	354		6 Mths
A. TOP		A. 3.0	
B. other programmes		B. 9.0	
C. other training		C. 7.0	
Compass Programme			
BOOST, 1997			
A. Boost		A. 53.0	
B. Non- Boost		B. 37.0	
Job search assistance			
NZES/NZISS Work Focus Interview	115	Not Reported	3mths
Job Action, 1995	647		6 Mths
A. Job Action		A. 27.0	
B. Comparison		B. 23.0	
Youth Action Programme, 1996	8428		0-16mths
Job Action and Youth Action, 1997			5 Mths
A. JA pre-enhancement	A/B 16723	A. 11.2	
B. JA enhancement		B. 6.5	
C. YA pre-enhancement	C/D NR	C. 9.9	
D. YA enhancement		D. 10.6	
Community Development			
Mahi A Iwi, 1997	NA. This was a process evaluation		
CEG Pacific Peoples, 1997	NA. This was a process evaluation		
Other			
IEA Pilot, 1997	NA. This was a process evaluation		
OSCAR DAP Pilot, 1997	316	10.0	0-12mths

APPENDIX 8: EVALUATION QUALITY

Table 12: Number of evaluations meeting or not meeting quality criteria

Aspects of the report being evaluated	Not included in the report	No. evaluations not meeting any criteria	No. evaluations meeting some of the criteria	No. evaluations meeting all the criteria
1. Completeness of the report	6	0	1	28
1.1 A table of contents				
1.2 Evaluation objectives are stated	4	0	4	27
1.3 A description of the background to the evaluation	1	3	21	10
1.4 A description of the programme	2	1	26	6
1.5 A description of the scope of the evaluation	0	1	22	11
1.6 A description of the evaluation methodology	1	12	9	13
1.7 Key findings	0	6	19	10
1.8 Conclusions	10	4	16	5
1.9 An executive summary	7	0	25	3
1.10 Inclusion of recommendations	25	0	2	8
2. Correspondence between the objectives and the rest of the report	4	0	28	3
3. Appropriateness and quality of the methods of data collection	1	0	18	16
4. Quality of the key findings	0	0	27	8
5. Relationship between conclusions and data	10	6	6	12
6. Quality of the recommendations	25	1	9	0
7. Presentation of the report	0	0	5	30

APPENDIX 9: METHODS USED IN EVALUATIONS

Table 13: Methods used in evaluations

Evaluations	Method	Evaluations	Method
1. Access Training programme, 1992	15	19. Youth Action Programme, 1996	1, 5, 8, 11, 18
2. Community Taskforce, 1992	15, 18	20. Job Support Programme, 1996	1, 5, 8, 12
3. Taskforce Green, 1993	7, 13, 15, 18	21. Enterprise Allowance and the Capitalisation Option, 1996	12, 18
4. TOP, 1994	7	22. NZCC Two Year Post-Course Evaluation, 1996	6
5. Behaviour and Perceptions of Long Term Job Seekers, 1994	8	23. Mahi A Iwi, 1997	4
6. Limited Service Volunteer Scheme, 1994	9, 15	24. CEG Pacific Peoples, 1997	4
7. Job Plus, 1994	8	25. Maori Youth Pilot, 1997	1, 5, 9, 18, 19
8. Joint NZES/NZIS Work Focus Interviews, 1994	15, 18	26. OSCAR DAP Pilot, 1997	7, 8, 12, 18
9. TFG, Tourism Green and Possum Control, 1995	15, 18	27. Wahine Rakari Programme, 1997	2, 5, 8, 9
10. Compass Pilot Programme, 1995	5, 6, 10, 18	28. IEA Pilot, 1997	7, 11, 18, 19
11. Training Incentive Allowance, 1995	9	29. Hiko ki Paerangi/New Horizons, 1997	3, 5, 8, 9
12. Job Plus Training Pilot, 1995	1, 8, 17, 18	30. Job Action and Youth Action, 1997	1, 7, 18
13. Survey of TOP trainees, 1995	8	31. Job Connection, 1997	1, 5, 8, 15, 18
14. Job Skills, 1995	1, 18	32. Tama Tane O Le Pasifika, 1997	1, 7, 11
15. Job Action, 1995	1, 5, 8, 15, 16	33. Be Your Own Boss, 1997	8, 18
16. NZCC and YSC programmes, 1995	12, 18	34. Compass After Two Years, 1997	16, 18
17. Tane Atawhai, 1995	1, 11, 12, 18	35. BOOST, 1997	5, 8, 15, 18
18. Barriers to Employment Facing Long-Term Job Seekers	8		

Key

Method

1. focus groups
2. affinity interviews
3. hui
4. case studies
5. in-depth interviews
6. semi-structured interviews
7. structured interviews
8. sample survey
9. before and after questionnaire
10. longitudinal survey

No. of evaluations

- 10
- 1
- 1
- 2
- 9
- 2
- 5
- 14
- 5
- 1

Method

11. interviews with key groups or people
12. analysis of existing documents
13. expert review
14. control and treatment groups
15. comparison groups
16. modelling
17. construction of databases
18. analysis of existing data
19. development of case profiles

No. of evaluations

- 4
- 5
- 1
- 0
- 9
- 2
- 1
- 19
- 2

APPENDIX 10: UTILISATION OF SELECTED EVALUATIONS

Documented actions arising from the evaluations

New Zealand Employment Service (NZES)

The two evaluations selected were:

- An Evaluation of the Enterprise Allowance and the Capitalisation Option
- An Operational Review of the Youth Action Programme

A handover document provided information on proposed actions and actions already taken as a result of the Operational Review of Youth Action. These actions appear to have been based on the findings. There were no recommendations included in the evaluation report.

Interviews with staff revealed there were undocumented benefits arising from the Operational Review of Youth Action. The key undocumented benefit was that, along with other ETF evaluations, the evaluation contributed to the development of IEA and SOLO. NZES believes that IEA will address many of the problems identified in the Youth Action evaluation, such as providing the appropriate response for clients, and case management.

While NZES staff were able to list actions taken as a result of the evaluation of Enterprise Allowance, they were unable to provide any documented evidence of these actions, which included:

1. Targeting was tightened. Enterprise Allowance was restricted to priority status clients (e.g. 26+ weeks unemployed)
2. The operational guidelines were updated.
3. A senior contracts adviser was hired

The actions taken as a result of the evaluation of the Enterprise Allowance and the Capitalisation Option appear to have been based on the findings of the evaluation. There were no recommendations in the evaluation.

Community Employment Group (CEG)

The two evaluations selected were:

- Mahi A Iwi Evaluation, 1997,
- Pacific Island People's Evaluation, 1997.

CEG had clearly documented evidence of the actions taken as a result of the Mahi A Iwi and Pacific Island peoples evaluations (Table 14). Actions arising from both evaluations were based on the recommendations in the evaluations. The Mahi A Iwi recommendations were more research focused. The Pacific Island Peoples evaluation recommendations were more policy focused. In both cases this is the usual process within CEG. Undocumented benefits arising from the evaluations were also identified (e.g. Staff involved in doing the two CEG evaluations were up-skilled in evaluation and monitoring. Staff are now being encouraged to undertake evaluation and monitoring of their own projects).

Table 14: Documented actions arising from CEG evaluations

Evaluation	Documented actions arising from the evaluations
Pacific Island People's Evaluation	<ol style="list-style-type: none"> 1. The recommendations were incorporated into CEG's National Pacific Island Strategy. 2. CEG has taken, or is to take, a number of measures to improve the delivery of services to Pacific Island peoples: <ul style="list-style-type: none"> • an additional three Pacific Island field advisers have been appointed; • One of the aims of a national fonotaga (meetings) held for all key workers, was to provide the opportunity for information exchange at a national level. • One of the aims of the regional fonotaga (meetings) for key workers and representatives of the Pacific Island communities, currently underway, is to identify sources of PI projects, and potential key workers. • workshops will be undertaken in all regions to identify and develop regional strategies for effective delivery to PI people 3. CEG has taken, or is to take, a number of measures to increase their profile with Pacific island peoples: <ul style="list-style-type: none"> • A communications strategy is being developed. The strategy includes interviews on radio, displays at community festivals, and the publication of a booklet and poster. • One of the aims of the regional fonotaga (meetings) for key workers and representatives of the Pacific Island communities, currently underway, is to raise CEG's profile. 4. CEG is undertaking work to up-skill all CEG field staff on PI protocols and cultures: <ul style="list-style-type: none"> • CEG is conducting pilot training package for CEG staff which will be fine tuned for delivery to all staff. 5. CEG is undertaking work to up-skill PI groups in management and administration: <ul style="list-style-type: none"> • One of the aims of a national fonotaga (meetings) held for all key workers, was to provide training in management and administration. • One of the aims of the regional fonotaga (meetings) for key workers and representatives of the Pacific Island communities, currently underway, is to provide training in management and administration.
Mahi A Iwi	<ol style="list-style-type: none"> 6. All staff were informed of the results of the evaluation: <ul style="list-style-type: none"> • each of the five field staff involved in the evaluation discussed the results of the evaluation with their regional teams; • each regional office received a full copy of the evaluation report • summaries of the evaluation report were sent to all staff 7. Feedback hui were held for Maori in South Auckland and the Hawkes Bay to validate the findings and to get their response to the report. Feedback indicated that Maori viewed the report as taonga (a treasure) because of the honest approach CEG had taken to examine the delivery of their services to Maori. 8. CEG staff were encouraged to compare their own practices to the trends outlined in the report and to discuss the findings in regional meetings. 9. The evaluation has provided a basis for further research and evaluation. For example, the following has been incorporated into the 1997/98 research plan: <ul style="list-style-type: none"> • research into how groups access CEG's services and implications of this for CEG, • a project to monitor community employment outcomes for three years to assess the short, medium and long term outcomes of CEG's work - the project will include Maori and Pacific peoples communities, • a project to develop and then refine new database measures for project outcomes (to capture a wider range of community employment outcomes than is currently the case), • scoping of a follow up to the Mahi A Iwi evaluation, and • the women's strategy evaluation focusing on outcomes (the next stage of CEG's evaluation strategy building on the work of the two exploratory evaluations).

Department of Social Welfare (DSW)

The two evaluations selected were:

- Compass, 1995
- Training Incentive Allowance, 1995

DSW had some documented evidence of action taken as a result of the 1995 Compass Evaluation. The findings of the Compass evaluation were reported to the Cabinet Committee on Employment, Education and Training (ETE) and the Cabinet Committee on Expenditure Control and Revenue (ECR) and Cabinet [CAB (95) M 27/5Dii refers]. The Compass evaluation did not formally list recommendations for action, rather it responded to each of the stated evaluation objectives listing findings. The recommendations to Cabinet about the future of the programme were based on a combination of the successful outcomes found by the evaluation and policy objectives. The evaluation also involved researchers 'identifying ways in which the operation and targeting of the programme could be improved'. Research for this objective involved qualitative interviews to ascertain the views of staff and customers. These recommendations fed into wider policy and operational delivery decision making rather than being specifically implemented. This was to ensure that changes are in line with Government's direction and limitations.

The same approach has been taken for more recent evaluations of the Compass programme. No information was provided on the evaluation of the Training Incentive Allowance.