

Working Paper Number 8

**MEASURING SOFT OUTCOMES AND DISTANCE
TRAVELLED: A METHODOLOGY FOR DEVELOPING A
GUIDANCE DOCUMENT**

**Measuring Soft Outcomes and Distance
Travelled: A Methodology for Developing a Guidance
Document**

**A study carried out on behalf of the Department for Work and
Pensions**

By

**Richard Lloyd and Fionn OSullivan
GHK**

© Crown copyright 2003. Published with permission of the Department for Work and Pensions on behalf of the Controller of Her Majesty's Stationary Office.

The text in this report (excluding the Royal Arms and Departmental logos) may be reproduced free of charge in any format or medium provided that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Crown copyright and the title of the report specified. The DWP would appreciate receiving copies of any publication that includes material taken from this report.

Any queries relating to the content of this report and copies of publications that include material from this report should be sent to: Paul Noakes, Social Research Division, Room 4-26 Adelphi, 1-11 John Adam Street, London WC2N 6HT

For information about Crown copyright you should visit the Her Majesty's Stationery Office (HMSO) website at: www.hmsso.gov.uk
First Published 2003

ISBN 1 84388 104 7

ISSN 1476 3583

CONTENTS

1 INTRODUCTION	1
2 LITERATURE AND PRACTICE REVIEW.....	3
2.1 Literature review.....	3
2.2 Practice review.....	4
2.2.1 Common elements	5
2.2.2 Requirements of different systems	6
2.2.3 Use of IT	6
2.3 Details of Books and articles reviewed.....	6
2.4 Internet searches.....	7
3 SUMMARY OF POSTAL SURVEY FINDINGS	9
3.1 Introduction	9
3.2 Postal survey.....	9
3.3 Response rates	10
3.4 Key Findings	11
3.4.1 Projects Measuring Soft Outcomes/Distance Travelled	11
3.4.2 Projects Not Currently Measuring Soft Outcomes/Distance Travelled ..	13
4 RESULTS OF THE CASE STUDY FIELDWORK	17
4.1 Introduction	17
4.2 Key findings.....	17
4.2.1 Target groups and context.....	17
4.2.2 Approach development.....	19
4.2.3 Approaches followed	20
4.2.4 Sophistication and complexity	21
4.2.5 Operational Context.....	25
4.2.6 Evidence vs. judgement	26
4.2.7 Resource issues	27
5 CONCLUSIONS AND IMPLICATIONS	29
5.1 Conclusions.....	29
5.2 Implications	30

1 INTRODUCTION

This report details the background research undertaken to produce a guidance document for ESF projects that are interested in monitoring soft outcomes and distance travelled (Measuring Soft Outcomes and Distance Travelled: A Practical Guide, 2003).

The research consisted of three main stages:

- Project commencement – including a review of relevant literature, and interviews with a range of stakeholder organisations (including those with an interest in, and experience of, measuring soft outcomes) to provide a review of current practice.
- Survey work with projects – consisting of two components, a postal survey of 1,500 ESF-supported projects, and case study visits and interviews with a further 10 projects.
- The development of the guide – based on the findings of the postal survey and case study research. Early drafts were circulated for comment to the steering group and the projects contributing to the research, and reviewed in a workshop event.

Following this structure, this report has three chapters, as follows:

- Chapter 2 summarises literature on soft outcomes measurement and current measuring practice as revealed in interviews with a range of “stakeholder” organisations.
- Chapter 3 provides background on, and the findings of the postal survey to assess the extent of current practice in soft outcomes measurement by ESF Objective 3 projects.
- Chapter 4 provides the findings from case studies of ten ESF projects and organisations whose responses to the postal survey suggested they had particularly distinctive approaches or significant experience of measuring soft outcomes.
- Chapter 5 provides summary conclusions and implications for the production of the main project output, the guidance documents, based on the postal survey and case study fieldwork.

2 LITERATURE AND PRACTICE REVIEW

This chapter details the results of a search and review of the literature and existing practice in the use of measures of soft outcomes and distance travelled. It examines nine models identified as part of the initial review. These included stakeholder organisations from the voluntary and public sectors already doing work in this area, a prior Institute of Employment Studies report on the same subject, participants at a DfEE seminar on soft outcomes in 2000 and Internet searches.

2.1 Literature review

In spite of extensive searching, the literature review in fact revealed very little on measuring soft outcomes or distance travelled. Almost no reference to the issues can be found in a search of the classic texts on evaluation, including relatively recent discussions (Chelimsky, 1997; Freeman & Rossi, 1993; Shaw, 1999). Similarly, recent “how to” evaluation guides for projects published in both this country and the United States, make no reference to the issue (DETR, 2000; Kellogg Foundation, 2000).

This contrasts with the various practical approaches to measuring distance travelled described below, suggesting that the academic and policy research literature on the subject has yet to catch up with current practice in project administration. However, there are some exceptions which are worth briefly discussing.

Collins (2000) discusses the use of software systems in the measurement of soft outcomes by non-profit organisations in the US. The article reviews the use of six specific IT systems by voluntary sector organisations to record and track soft outcomes and to produce analytical reports. It notes the substantial problems of doing this and therefore the importance of technology solutions in reducing time and other constraints.

A DIY guide to implementing outcome monitoring (Burns, 2000) is published by Alcohol Concern, an organisation that has carried out sustained work on both hard and soft outcome monitoring since the early 1990s. This includes discussion on outcome monitoring in general, common problems and pitfalls, and how to avoid them. It also contains discussion of soft outcome measures and how to formulate them, including guidance on how they can be developed by individual projects. The guide raises two interesting questions.

Firstly, with reference to the extent to which “off the shelf” or generic systems are suitable for a range of different ESF projects. The Guide suggests that individual projects should develop their own outcome indicators (both hard and soft) for the following reasons:

- Even if dealing with a similar client group (i.e. those with alcohol-related problems) each project is likely to have different aims, and furthermore, different stakeholders whose concerns must be addressed in designing its monitoring systems.
- The process of developing indicators itself is important in gaining compliance and in maintaining quality assessment standards.
- There is no essential need for projects to use exactly the same indicators, e.g. so that the outcomes of different projects can be compared. Due to the inherent subjectivity of systems for monitoring soft outcomes it is anyway mistaken to use them to compare the performance of different projects.

Secondly, the Guide categorises the benefits of using soft outcome indicators for three groups:

- Clients (e.g. by illustrating to individuals that they have made progress while on the project).
- Project managers and project staff (e.g. by enabling an agency to “stand proud” and state what they have achieved, or allowing them to fine-tune their services).
- Funding bodies (e.g. by providing a clearer idea of what a programme is achieving beyond the more easily demonstrated hard outcomes).

It is notable that the benefits are different for each stakeholder group. This issue becomes clearer through reviewing the contributions of Davey (2001) and O’Neill (2002) to the overall debate on the appropriate use of performance indicators. Both note the drawbacks of using performance indicators to drive “continuous improvement” in the voluntary sector, and that there is a risk that inappropriate use of such measures may perversely end up undermining those organisations whose performance they seek to improve. The implication is that funding bodies should avoid using soft outcome measures to compare performance between projects. This is particularly the case for those measurement approaches that rely heavily on subjective judgements. Imposing a particular standard of measurement on projects may also obscure more than it reveals, as different projects may have different aims and may also work with clients with very differing needs.

2.2 Practice review

Nine organisations were identified as having particularly relevant experience of measuring soft outcomes and distance travelled, and their approaches were reviewed through a series of interviews.

The approaches examined varied in terms of their sophistication and complexity, as well as other variables such as target groups and the range of indicators followed. The approaches included those developed by both

individual projects and commercial organisations. A series of issues were identified as part of the practice review, as summarised below.

2.2.1 Common elements

There are close similarities between all of these systems for measuring soft outcomes and distance travelled, and most include the following elements:

- **A set of target indicators** relating to the soft outcomes that the programme or agency wants to track. These can be more or less detailed, but they essentially break down a broad target area into a list of component parts (e.g. examples given above: exercising self-discipline, or assuming responsibility for oneself at work). One of the issues here is to ensure that the sub-indicators are comprehensive, and capture all the intended benefits of the project, as there is a tendency for effort to be directed to what is being measured.
- **A scoring system.** This is usually in the form of a scale in varying degrees of complexity, (e.g. ranging from three to ten point scales). The way scores for particular indicators are recorded is important. In some cases the client themselves directly records the score using a paper or computer-based system. Elsewhere an interviewer, rather than the client, does the scoring and assesses distance travelled.
- **Baseline and subsequent interviews** are used to assess progress. The variation between different systems here lies in the intensity and frequency of the process, which will depend on the client group and the nature of the programme. In some cases very substantial periods of time (e.g. several weeks) are taken to make an initial assessment, and carry out daily monitoring. For others, time and resource constraints mean that initial assessment needs to be done relatively quickly, and that subsequent assessment will be less frequent.
- **A system for reporting results.** There is some variation here, for example, between tables, computer-generated charts or graphs. Again, this varies according to client groups and the nature of the programme; graphic representation may be preferable to writing for some clients or programmes.
- **Training for staff using the system** is essential for using all systems to ensure quality control and consistency, but will depend on the complexity of the system. All stress the importance of organisation-wide commitment to, and understanding of, the systems used.

2.2.2 Requirements of different systems

Some of the systems reviewed above come with supporting literature that outlines the specifications with which the system was designed. The good practice points that these identify suggest that systems for measuring soft outcomes should:

- Identify barriers to learning / personal development;
- Focus on variables that can lead to changes in behaviour;
- Have valid measures;
- Be reliable – so that the similar results can be produced in similar situations;
- Include multiple variables and multiple sources of information to get an even balance of indicators;
- Be part of a wider evaluation process, building on existing assessments and information and contributing to existing plans;
- Be relatively simple and cheap to administer.

2.2.3 Use of IT

Several systems use IT packages (including web-based systems) to collect and present the assessment results. Others are developing an electronic aspect to their systems. The main advantages to doing this are that it makes it much easier for information on distance travelled to be collected and analysed, and also that it enables the results to be presented in graphical format, which is especially preferable for client groups with literacy problems. IT issues should be considered throughout the design stage of the framework. An IT element need not necessarily be complex, for example, with the framework developed with an accompanying Excel spreadsheet. It is important to note that this may raise certain legal data protection issues about how information is recorded that will need to be considered at the same time.

2.3 Details of Books and articles reviewed

Burns, S (2000) *A DIY Guide to Implementing Outcome Monitoring Alcohol Concern*, London;

Chelimsky, E (1997) *Evaluation for the 21st Century* Sage, London;

Collins, M (2000) *Using software systems to measure non-profit programme outcomes: assessing the benefits and barriers to strategic management*, Kennedy School of Government
(http://pages.prodigy.net/michael_collins/outcomes_IT.htm);

Davey, B (2001) *The Measurement of Performance in the Community and Voluntary Sector*, Nottingham MSF;

DETR (2000) *New Deal for Communities – Outcomes & Milestones*, DTLR, London;

Freeman, H and Rossi, R (1993) *Evaluation 5 – A Systematic Approach* Sage, London;

Kellogg Foundation (2000) *The Kellogg Foundation Evaluation Handbook*, (www.wkkf.org/pubs/Pub770.pdf);

Institute for Employment Studies (2000) *Guide to Measuring Soft Outcomes and Distance Travelled*, IES, Brighton;

O'Neill, O (2002) *Called to Account BBC Reith Lecture No.3*, (<http://www.bbc.co.uk/radio4/reith2002/3.shtml>);

Ian Shaw (1999) *Introducing Qualitative Methods*, Sage, London.

2.4 Internet searches

Searches were conducted using words “soft outcomes” and “distance travelled”. The most useful site is the Yorkshire/Humber Connexions website, which details over 15 assessment tools (the most relevant ones to the project are discussed above). See www.getting-on.co.uk/toolkit/index.html. Otherwise, the web search did not come up with anything very useful. There is a vast amount of literature on evaluation in general, but nothing specifically about measuring soft outcomes/distance travelled.

3 SUMMARY OF POSTAL SURVEY FINDINGS

3.1 Introduction

This chapter describes the key findings from the postal survey of ESF funded projects undertaken in April and May 2002. The aim of the survey was to identify the extent to which soft outcomes/distance travelled were being measured by projects funded under ESF Objective 3, as well as:

- Gaining an insight into the nature and degree of sophistication of the approaches followed where identified;
- Identifying barriers to measurement, and assessing the demand for guidance on developing new approaches amongst projects who were not currently measuring.

The findings were also used to identify a series of 'case study' projects to be visited and their approaches examined in detail. This paper summarises the method followed, the survey sample and response rates achieved, and an overview of the key findings.

3.2 Postal survey

Following the development and agreement of a suitable project sample and postal questionnaire, the survey consisted of an initial and follow-up mailing to 1,533 projects. The first mailing took place on 3 April, with an initial deadline for responses of 19 April, with a second mailing taking place shortly afterwards.

The **questionnaire** was designed to identify if projects were measuring soft outcomes/distance travelled as part of their work, and if so with which beneficiary groups. Projects were then asked to summarise their approaches including: the soft outcomes being measured, how the relevant data is collected, how baselines are established, how the resulting data is used. Where projects were not measuring soft outcomes/distance travelled, they were asked if they planned to do so at some stage, and if not would they consider doing so if an appropriate approach was available. This group of projects were also asked what they considered to be the main barriers to/difficulties in measuring soft outcomes.

The **survey sample** was initially intended to be based on ESF Objective 3 projects who had described measuring soft outcomes/distance travelled in their interim monitoring returns. However, as only 372 projects submitted returns describing such measurement approaches, the sample was enhanced with the inclusion of projects from policy fields 1, 2, 3 and 5.

A series of filters were applied to the additional projects to ensure they focussed on supporting individuals (and individuals without qualifications in field 3), and to ensure projects with a disability focus were not contacted (to avoid duplicating the work of the Disability Research Forum). Once these filters were applied, a sample was drawn randomly.

The **distribution of the sample** is shown by policy field in Table 3.1 below. The projects shown under policy field 4 were drawn solely from projects describing measuring soft outcomes in their interim monitoring returns.

Policy Field	No. Projects	% of Sample
1 – Active labour market policies	468	30.5%
2 – Social inclusion	625	40.8%
3 – Lifelong learning	308	20.1%
4 – Adaptability and entrepreneurship	31	2%
5 – Gender equality in labour markets	101	6.6%
Total	1,533	100%

3.3 Response rates

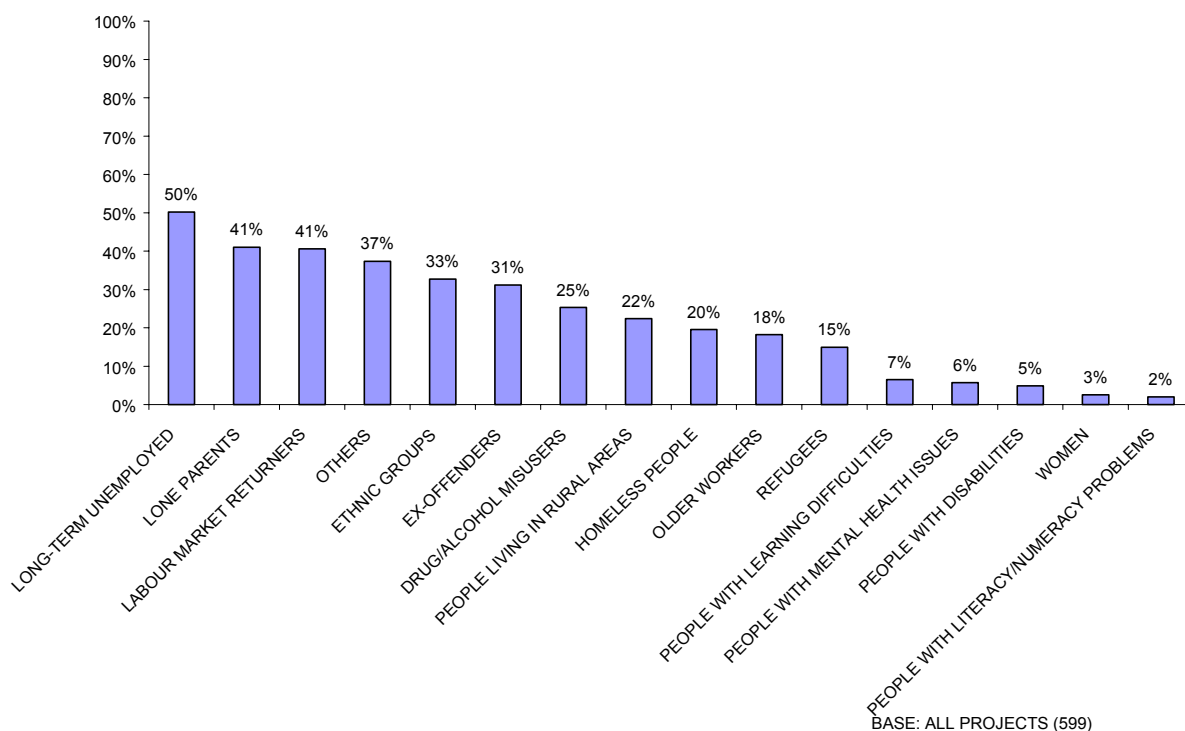
The survey achieved a high response rate for a postal approach, with 599 responses being received from the initial sample of 1,533 projects (a response rate of 39%). The response rate is more impressive given that in a number of cases organisations submitted responses on behalf of more than one ESF project – so in reality the coverage by ‘project’ was greater than the 39% described. This included responses describing ‘national’ approaches from organisations such as Fairbridge and RPS Rainer, who are operating multiple ESF funded projects from different centres across the country.

Table 3.2 below shows the distribution of 599 responses by policy field. As described previously, a number of responses referred to more than one project, and so the totals in the column total more than 599/100%.

Policy Field	No. Projects	% of Respondents
1 – Active labour market policies	160	27
2 – Social inclusion	266	44
3 – Lifelong learning	156	26
4 – Adaptability and entrepreneurship	29	5
5 – Gender equality in labour markets	55	9
Not provided	16	3

Figure 3.1 below shows the distribution of respondents in terms of the beneficiary groups served. Again, as a number of projects described working with multiple client groups the percentages add to over 100%.

Figure 3.1: Beneficiary Groups Served by Survey Respondents



3.4 Key Findings

The key findings of the survey are summarised below, for projects measuring soft outcomes/distance travelled and for those that were not.

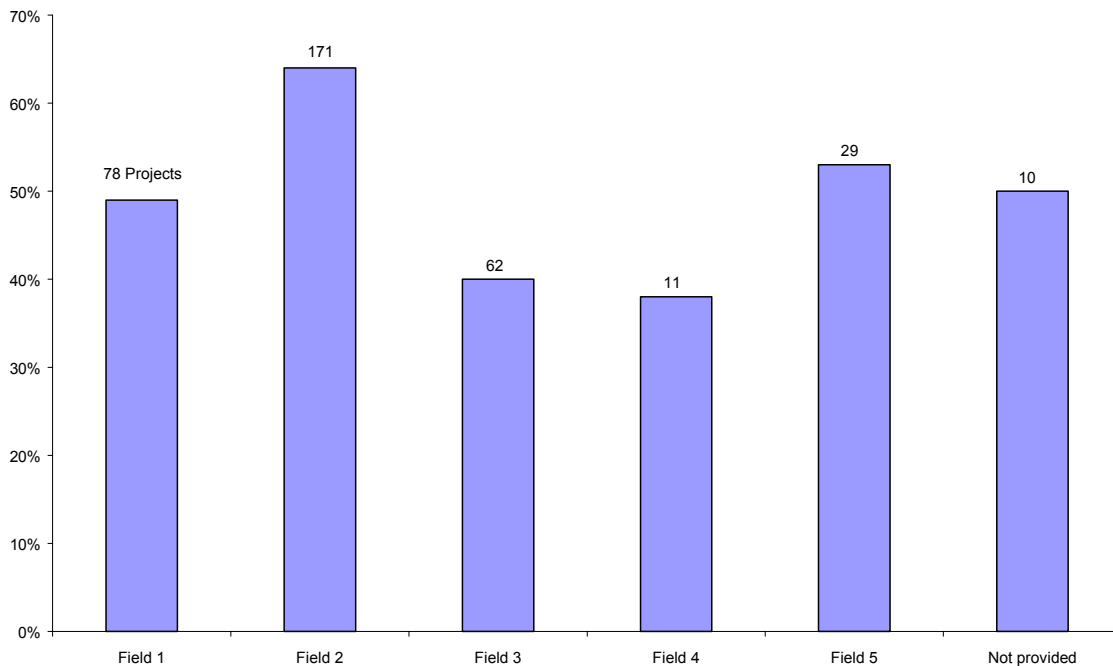
3.4.1 Projects Measuring Soft Outcomes/Distance Travelled

The key findings from the respondents describing measuring soft outcomes are described below.

Numbers Measuring

Of the 599 responses received, **311 (52%) described currently measuring soft outcomes and/or distance travelled** as part of their projects. The distribution of the respondents, by policy field, is shown in Figure 3.2 below. The figure shows the numbers of respondents measuring soft outcomes, and as a percentage of the respondents in each policy field.

Figure 3.2: Respondents Measuring Soft Outcomes by Policy Field



As described previously, a number of responses came from organisations involved in multiple ESF projects, in effect increasing the response level and coverage but importantly suggesting that a number of ‘common practice’ approaches are in place. This was supported by the identification of a number of common approaches by projects (such as the commercially produced Rickter Scale and Quality of Life models), and derivatives of other approaches (for example, a number of approaches described as being ‘based on the Bridges project’).

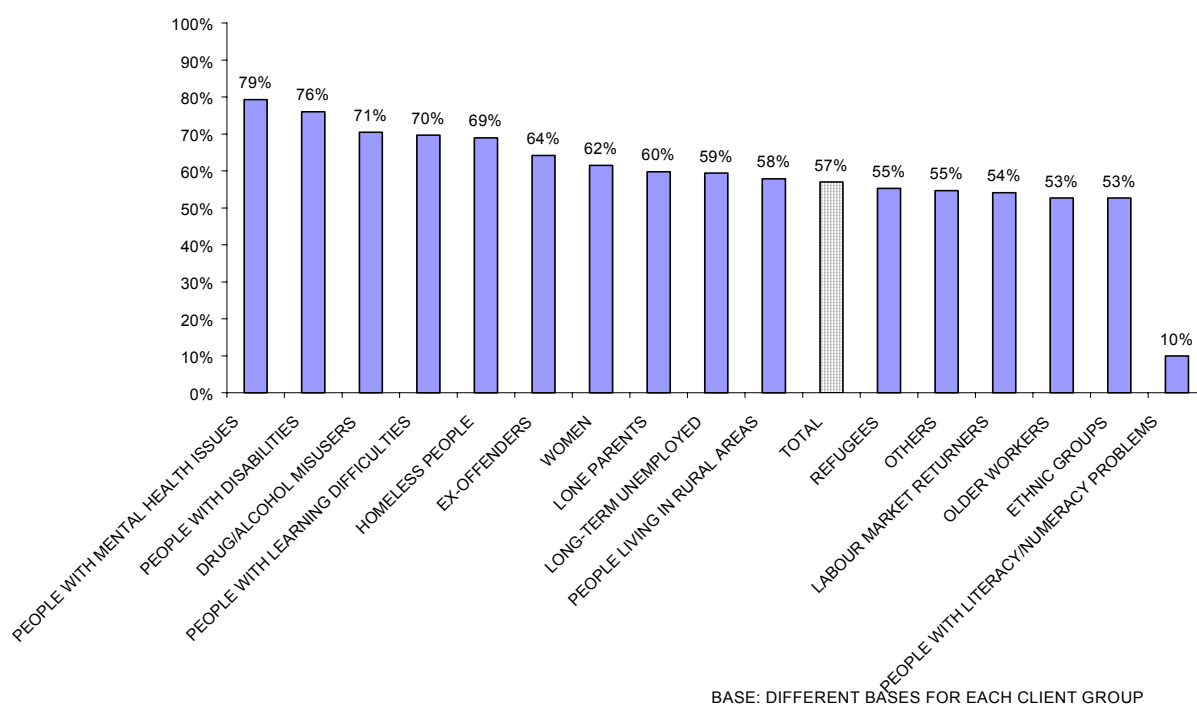
Measurement by Beneficiary Groups

Having estimated the scale of measurement across ESF Objective 3 projects, the next consideration was the distribution of measurement approaches by beneficiary groups. Clearly the survey sampling approach is relevant to the findings, as are the precise nature of the target groups for the individual respondents. Importantly, the survey sample sought to avoid projects with primarily disabled beneficiaries, as approaches to measuring soft outcomes for this particular client group were being examined as part of a parallel research study.

Figure 3.3 below describes the share of projects measuring soft outcomes for different beneficiary groups, expressed as a percentage of those working with each group (for example, 79% of all respondents working with people with mental health issues were attempting to measure their soft outcomes). However, it should be remembered that the base for different client groups will vary – and can refer to low numbers of respondents. This may account for the

apparent low share of measurement with people with literacy/numeracy problems, a beneficiary group which were described by just 2% of all respondents (10 respondents), of which one described measuring soft outcomes.

Figure 3.3: Respondents Measuring Soft Outcomes, by Beneficiary Group



While Figure 3.3 shows the extent of coverage by individual beneficiary type, the survey also identified that **measurement approaches rarely focused on single beneficiary groups** and commonly covered multiple beneficiary groupings (although in the survey this is clearly a function of the range of beneficiary groups individual projects are working with).

Sophistication / complexity

The questionnaire also allowed projects to provide summary details of their approaches, to provide some measure of their sophistication. This was particularly helpful in supporting the process of case study selection, although it showed that in many cases ‘measuring soft outcomes’ was more of an intention than a systematised component in the projects’ approaches.

3.4.2 Projects Not Currently Measuring Soft Outcomes/Distance Travelled

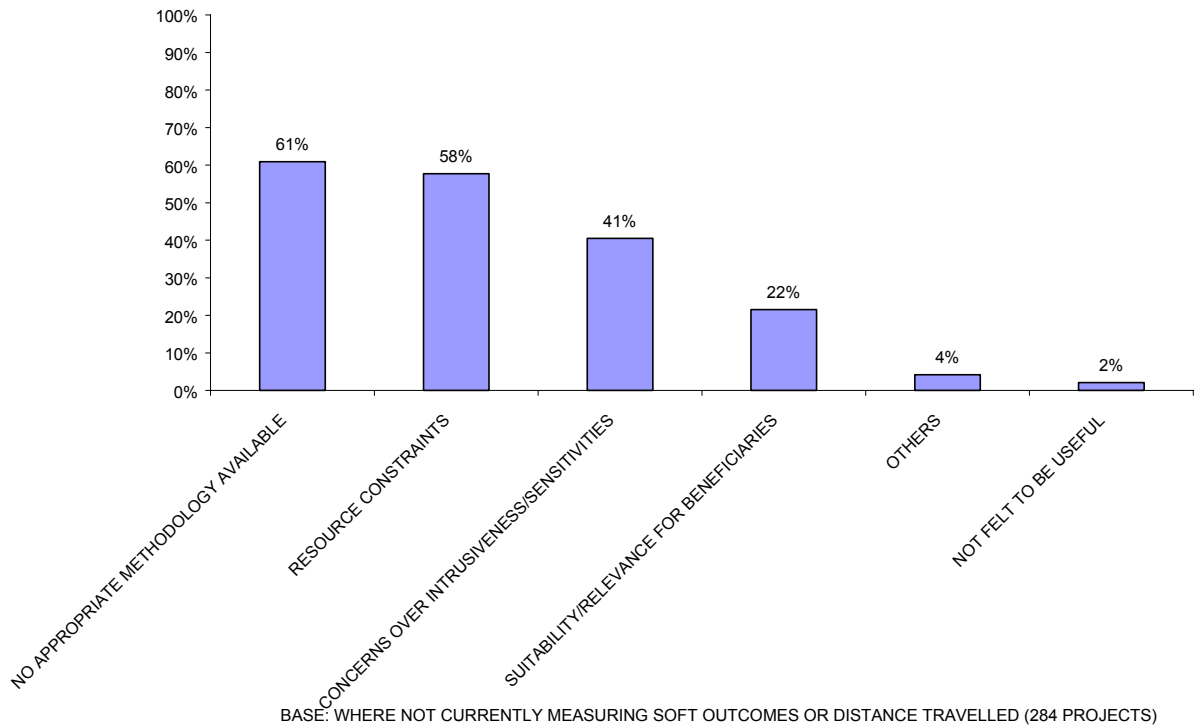
While 52% of the respondents described measuring soft outcomes/distance travelled, 48% (288 respondents) were not. The characteristics of this group were broadly similar to the ‘measuring’ respondents, in terms of distribution by policy field and the beneficiary groups they are working with.

Key Challenges/Barriers to Measurement

Respondents not measuring soft outcomes identified what they considered to be the main barriers to/difficulties in measuring soft outcomes. The distribution of responses is provided in Figure 3.4 below, which shows that.

- The most commonly cited barrier was the absence of an appropriate methodology, cited by 61% of respondents.
- Resource constraints were also an important consideration, a close second to methodology issues and cited by 58% of respondents.
- The issues of intrusiveness and concerns over client sensitivity were cited by four out of ten respondents, illustrating the level of concern over such issues in measuring soft outcomes with their target groups.
- Of less concern, but still cited by over one in five respondents, was the perception of the suitability and relevance of measuring distanced travelled with their particular target groups.

Figure 3.4: Perceived barriers to measuring soft outcomes for projects



Importantly also, the fact that only 2% of respondents (representing five projects) felt that measuring soft outcomes was not useful to them provides evidence to the perceived usefulness of the information such approaches can collect.

Future Plans

Respondents were also asked if they were planning to develop and implement systems in future, and if not whether they would consider doing so if an appropriate approach was available.

Positively, **one third of the respondents not currently measuring described plans to introduce systems in the future**, broadly distributed across a range of beneficiary groups. While this is a measure of intention only, and would need to be tested over time, it suggests an interest in developing approaches to measuring soft outcomes, the realisation of which would be assisted by some form of model approach or guidance.

This notion was supported by the responses to developing a system if an appropriate approach was available. Again, while still a measure of intention, **80% of the respondents said they would be interested in measuring soft outcomes/distance travelled if an appropriate approach was available**. While this further evidences the interest in measurement and the demand for guidance, the findings on perceived barriers, notably resource issues, mean any approach will need to consider the other cited constraints.

4 RESULTS OF THE CASE STUDY FIELDWORK

4.1 Introduction

This chapter summarises the findings of the fieldwork undertaken with the ten project case studies. It is also informed through contact with stakeholder organisations, as well as other relevant studies and approaches such as the Employment Service's Client Progress Kit.

The case studies were selected on the basis of their postal survey responses, and consequently focused on measuring soft outcomes relating to, and distance travelled towards, employability. An additional Wales-based project was included in the final sample. The case studies were selected on the basis of distinctive and reasonably well-developed and tested systems being in place, as well as providing balanced coverage by regions, urban and rural areas, and beneficiary group. The stakeholder organisations were selected on the basis of their experience and the insights they could provide on the development and use of relevant systems.

In broad terms, the approaches comprised a series of similar components, reflecting the principles for soft outcomes measurement systems identified in the initial literature review. These included the establishment and use of indicators relating to the outcomes to be measured, the assessment of progress through baseline setting and subsequent review, the use of a scale to track progress (including numerical scoring systems; a "traffic lights" system to indicate progress, stability or regression; or a set of "pit stops" along the course of a race track), and providing a formal reporting process. Beyond these basic principles, however, there was substantial variation between the approaches, in terms of the indicators used, data collection approach followed, and overall complexity. The case studies also provided a series of important process lessons, and key success factors for effective development and implementation.

4.2 Key findings

4.2.1 *Target groups and context*

The case study projects each worked with a range of target groups, although "disaffected" 14 to 25 year olds emerged as the most common 'main' target group. The characteristics of this group and the challenges they face mean that they fall into different "beneficiary groups" - for example having literacy or numeracy problems, histories of offending, and/or substance abuse problems.

However, the use of the systems in all but one case is not confined to young people, but also applied across a range project beneficiaries, sometimes with minor modifications. For example:

- One approach was originally developed for young people with very severe problems but a subsequent version has been developed which focuses specifically on employability issues for people of all ages.
- Another approach included the use of a scratch card with a range of clients, whose unifying characteristic is long-term unemployment.
- A third featured different checklists for clients aged “over 25” and “under 25”, although the differences are slight and the process by which they are used is the same.

A lesson from the case studies is that there is **not necessarily a rigid link between different measurement approaches and particular target groups**, more that elements of some approaches were of particular relevance to certain groups, and that most can be adapted for a different group if necessary. One organisation, for example, used a routed questionnaire including questions relating to homelessness and substance misuse. The questionnaire featured skips where the response to an introductory question suggested the topic was not relevant for the individual beneficiary.

The case studies also identified that **not all approaches will be used with all clients, on an individual rather than beneficiary group basis**. There were two main reasons for this:

- **First, where there are specific reasons for not monitoring certain client groups/individuals**, for example if the target group or individual is sufficiently disaffected/distanced from society that assessment would be perceived as unacceptably intrusive, and risk their withdrawal from the service or damage to a carefully nurtured relationship. While this was often more associated with the timing of initial baselining, examples were cited where the process of measurement would be considered as a risk to subsequent delivery.
- **Second, where attendance at projects is erratic, insufficiently frequent, or of too short a duration for either measurable change to result or measurement to be practicable**. In some cases this ‘filtering’ of use was based on experience. For example, one project described moving from 100% client coverage to exclude ‘drop-in’ sessions where inputs were limited and subsequent attendance sporadic. This ‘filtering’ was further refined to differentiate between clients requiring simple signposting to support their job-hunting efforts (for whom distance travelled was not measured) and those facing identifiable issues where long-term, intensive support was required. Another organisation described not attempting to measure change unless clients are attending for a minimum of three months, because of clients’ complex needs and

the necessity of having sufficient knowledge of these in order to work effectively with them.

More broadly, the very nature of the topic area (such as personal issues, or areas of personal awareness/self-review that may be new to the beneficiaries) has implications for the most appropriate approach, timing and context for measurement. One frequently stressed message was the need to allow sufficient trust to develop between the beneficiary and the assessor, to ensure compliance and honest reporting (though this is less of an issue in evidence-based system were discrepancies between what clients say and what they can actually do should be immediately clear). This timing issue will have implications for baselining.

4.2.2 Approach development

The motivation for the development of the different case study approaches was almost unanimously the same. It stemmed firstly from the wish to assess clients' initial needs and draw up action and personal development plans. Secondly it arose out of the desire to show both participants and project workers that progress is being made. This reflects the fact that nearly all the case projects deal with clients who are often some way from achieving hard outcomes. Interestingly and importantly, this emphasises that the approaches examined were not developed/implemented solely as 'monitoring' tools, but also to have an active role in both in initial diagnostic and on-going motivational support.

The developmental experiences of the case studies (and the stakeholder approaches) provided a series of lessons of relevance to the assignment, including:

- Systems for measuring distance travelled take a long time to develop, and are frequently refined and improved over time. A number of approaches have taken years to develop, pilot and refine, with one organisation describing how they did not consider the approach ever being a 'finalised product'.
- There are considerable benefits to be gained from periodic review and revision in order to fine-tune the approach to the specific needs of a project. Annual reviews were suggested, in order to capture potential improvements from project staff and to assess suitability in light of any changes in the client group.

A key finding was that **the development and implementation process itself can have a considerable influence on the effectiveness** (in the widest sense) of the resulting approach. Where measurement approaches were 'dropped into' a project, with limited involvement of operational staff and other stakeholders, adoption was more likely to be limited. Conversely, one good practice principle related to the development or introduction/tailoring process. This suggested the involvement of both management and operational staff,

other stakeholders/partners and ideally beneficiary representatives made for a more relevant, appropriate and fit for purpose system that all parties could 'buy in to'. While this process must be resourced and well managed, the long-term benefits will be worthwhile.

This collaborative approach was mentioned frequently, and had become particularly evolved amongst the stakeholder group. One organisation described undertaking reviews of their approach in this way, including 'brainstorming' sessions to identify suitable indicators. For their next review, they are planning to involve specialists in alcohol, drugs and mental health areas to help ensure the questions they ask are most relevant.

Even where shared development work has taken place, buy-in across users is not guaranteed. In some cases management interviewees expressed differing views of the usefulness of the approaches to operational staff. This suggests the importance of the review process both to foster a continuous improvement approach, and to ensure any issues reducing coverage and benefit are addressed.

4.2.3 Approaches followed

At the overall approach level, a number of key points emerged:

- The approaches varied considerably in their **sophistication and complexity**, defined by the nature and range of the indicators used, baselining and data collection approaches, data collection tools employed and the frequency/nature of assessment followed.
- The **operational context** within which the measurement process was set was also important, both in terms of the characteristics of the approach but more importantly the benefits realised for all parties and commitment resulting.
- The nature of the evidence collected also varied, in terms of **evidence vs. judgement based approaches**. The distinction was made between 'hard' and judgemental (often, but not exclusively, self-reported) evidence. Most frequently pragmatism and practicability led to judgement based approaches – with resulting issues around validity and accuracy.
- The amount of **time/resources required by each approach also varied considerably** – again as a function of the sophistication and complexity of the approach followed, although this remains a key constraint for measurement approaches.

Insights into the development approaches followed by the case studies also identified **different development trajectories** – either from initially complex systems which were revised and contracted to become more focused and practicable, and initially simple systems which expanded as their benefits (and

shortcomings) became apparent. This again emphasises the importance of the development and review process.

4.2.4 *Sophistication and complexity*

Each of the approaches comprised a series of common features, namely a set of indicators to establish baselines and measure subsequent progress, a series of data collection tools (including hard copy forms and checklists, other more visual recording systems, and IT based systems), and an overall data collection, analysis and reporting strategy.

While these features were broadly similar, they differed widely in terms of detail. The paragraphs below provide an overview of the different approaches.

Indicators

An indicator is a measure used by a project to assess the extent to which soft outcomes have been achieved. Indicators used ranged in terms of numbers and focus – from banks of indicators to just six questions in the simplest case. Examples of indicators used by some of the projects are described in Table 4.1 below.

Most, but not all, of the assessment approaches used some form of **scale or scoring systems**. In some cases diagrams are used to record when a task has been completed.

The focus of the approaches was also different, in terms of the extent to which the indicators used were common across the beneficiary group, or selected on a more individual basis, and whether they addressed skills or specific tasks.

Table 4.1: Exemplar Indicators

Project	Sub-heading / standard	Examples of indicators
Organisation 1 (Clients are scored on a scale between 1 - 6)	Reliability: the young person is able to: Motivation / Attitude	<ul style="list-style-type: none"> - Be punctual - Attend regularly in accordance with agreement - Notify absence appropriately - Manage time effectively - Demonstrate honesty - Work without supervision when appropriate - Arrange external appointments in agreed time - Demonstrate a positive attitude towards a challenge - Work co-operatively as part of a team - Work co-operatively with those in authority - Demonstrate enthusiasm and enjoyment - Sustain agreed activity / activities
Organisation 2 (Client completes form saying if they strongly agree to strongly disagree)	Communication issues Presentation issues	<ul style="list-style-type: none"> - I find completing forms easy - I sometimes need help with reading instructions - I often worry about handwriting/ spelling - When I need to check change I do it in my head - I can use a computer - I often have difficulty understanding when people speak English - I have somewhere to wash - I am always punctual e.g. work/the course/appointments - I always telephone with a reason if I know I'm going to be late - I often have to leave early due to my situation
Organisation 3 (Clients are graded A (above standard), B (satisfactory), C (below standard) or D (unacceptable))	<ul style="list-style-type: none"> - Work output - Work quality - Attendance - Timekeeping - Communication - Attitude - Appearance - Conduct - Team working - Job search 	<ul style="list-style-type: none"> - Completed set tasks within timescales - Tasks did not need redoing - Was either 100% or there is a satisfactory reason for absence which is properly documented and communicated to Personnel - Was consistently punctual. Returned from breaks without prompting - Maintained clear, effective and appropriate communication with all - Worked willingly, receptive to instruction - Was neat and presentable. Wore clothing suitable to the job. Observed corporate / safety clothing requirements - Maintained a positive, professional, mature and helpful attitude with all - Worked as a team member, wasn't selective over work tasks. Shared workload and responsibility - Made consistent, conscientious and measurable attempts to find work - reflects individuals job search strategy
Organisation 4 (clients indicate between always and rarely)	Achievements	<ul style="list-style-type: none"> - I am aware of my strengths and weaknesses - I can control my finances - I don't give up easily on things - I am able to work on my own - I can concentrate for 30 minutes - I can complete tasks on time - I find it hard to ask for help

The nature of indicators was of particular relevance in baselining, with issues around the indicators used and measurement approach followed described below.

Issues around Baselining

Baselining refers to the process of finding a starting point specific to the individual client against which their progress in the areas the project plans to work with them can be measured. A number of views and issues around this process emerged in the research. First, the process is clearly central to measuring distance travelled, although it must acknowledge the importance of elucidating honest responses and recognising the sensitivities given the nature of the data. A number of approaches were described, with the key issues being:

- **Timing** – a key point – some projects felt that if not done at the start of the intervention, baselines will by definition be inaccurate. For example, one organisation described having an intensive introductory week, when they would expect some change, but rarely being able to baseline in this period. Hence are their baselines valid, or will they always undervalue their initial work?

- **Interpretation** – projects also identified issues around how baseline data is interpreted. In self-assessment processes, clients may rate themselves initially highly then find ratings drop, as they become self-aware for the first time, or because positive initial feelings about the project drop off. This drop may not necessarily represent regression, but the way in which such ratings are reported to the beneficiary will be important to their continued motivation and engagement.

- **Accuracy/Validity** – in describing their experience of baselining, one project described how individuals are asked to rate the importance of their appearance to them, and later their satisfaction with their appearance. The project manager described how a client's appearance could be a barrier to employment, but that they could report high importance and satisfaction with it at the baselining stage. Clearly this was useful in identifying an issue to work with, but how useful was it as a baseline against which to measure progress? This in turn raised the issue of the extent to which isolated self-reporting can always provide clear baselines.

The importance of design (in terms of data collection tools and the representation of change achieved) was stressed by a number of projects, as summarised below.

Design Issues

The importance of design and layout issues was stressed by projects, both in the general design of the system and in the presentation of results. One of the disadvantages of paper-based assessment systems is that they can be off-putting to clients with literacy problems. Two organisations use a CD ROM-based system, which have screen pictures as well as voice-over options reading the questions aloud. It is notable, however, that each of these tools were developed on a commercial basis.

Several interviewees stressed the importance of good basic design principles, such as keeping forms short, avoiding over-crowding on the page and over-

small text. They also stressed the need to be careful in the use of language, being careful to avoid technical or ambiguous terms.

Data collection approaches

A range of data collection approaches were described from the simple completion of a short checklist, to more complex/detailed approaches. The frequency and nature of follow-up assessment depended mainly (but not solely) on project context and delivery approach. Their key features and issues commonly identified are summarised below.

Data Collection Approaches

Projects used either hard copy/paper questionnaires and checklists or IT based approaches, each having specific advantages and disadvantages. These included:

- **Paper-based systems** – advantages include: being straightforward to operate, cheap to develop, and cheap to implement (with the exception of any support needed to help completion). Disadvantages include requiring clients to be literate (although some ways round this), and the need to be consolidated by hand/input separately into an IT system for analysis and reporting.

- **IT-based systems** – advantages include avoiding paper work, the ability (in some cases) to provide verbal instructions and detect unreliable response patterns, attractive graphics/ animation, and simplified and rapid storage, consolidation, reporting and analysis of data. There is evidence that these systems are also attractive to certain groups, particularly young people. Disadvantages include cost/time/skills to develop, requirement for IT system may be off-putting to the IT shy, and can be subject to development and operational problems (e.g. must complete whole assessment before saving, problems with data transfer, etc). In addition, data security issues must also be considered.

Only one system used a wholly IT-based approach (with beneficiaries inputting their responses directly via computer), although others have supplemented their paper-based approaches with IT packages to store, analyse and present records of progress. IT is however not the sole means of presenting assessments - other approaches included a variety of charts and graphics to represent progress/distance travelled.

Data sharing issues

Each of the case studies were aware of the legal and data protection issues surrounding the information they collect, with most having organisation-level protocols and rules on keeping records private. These generally include a verbal explanation to participants that all information will be confidential unless written consent is provided, with the exception of information indicating dangers to children, which all agencies have a legal duty to disclose to the Police.

Apart from legal/moral data security issues, some projects and beneficiaries have found it very useful to have results that can be presented to employers or

other agencies. For example, one project participant took a chart showing the progress he had made to a housing agency in support of his claim.

4.2.5 Operational Context

Clearly any measurement approach must bear some relationship to the aims, objectives and delivery activities of the project it is used in. Importantly, in a number of cases the measurement of distance travelled was not viewed as an isolated, solely 'performance measurement' activity, but as a central component of their diagnostic and delivery processes.

This could take a number of forms, the most common being an initial diagnostic/needs identification process featuring a baselining component, which would also lead to the targeting or emphasis of certain delivery activities. For example, one project uses an initial baselining process to identify general areas that clients need to work on to develop their employability. However, the specific tasks that clients then work towards are individually chosen according to their needs and work ambitions. For example, these range from "taking a bus to work unaided" (for a client with learning difficulties) through to "learning a computer package" (for a client with mental health issues, but otherwise highly competent to work in a range of jobs).

Elsewhere the initial assessment/baselining work was an important element of 'individual development plans', which were themselves an individualised approach to identifying goals, steps towards, and a means of quantifying progress. For example, one approach links soft outcome assessments with progression with individual development plans for clients attending the project.

In addition to these more 'formalised' approaches, it was clear that the process of baselining could provide important additional information to ensure the benefits resulting from clients' involvement with a service, were maximised. The accuracy/validity example on baselining described, for example, showed how the process could lead to the identification of important information on how specific areas needed to be worked on.

The integration of systems for measuring distance travelled into operations did not stop with the baselining stage, but could extend throughout the individuals' involvement with the programme. This allowed the follow-up measurement of progress to inform continued service delivery, potentially re-tuning emphasis or indicating when a client was ready to progress to the next 'stage' in working towards their employability. One of our key findings was that **when the measurement of soft outcomes/distance travelled was an intrinsic part of delivery, the process was more valued**. This simple lesson has real significance for attempts to promote the assessment of distance travelled, by evidencing the benefits such assessment can offer to both projects workers and their clients.

4.2.6 *Evidence vs. judgement*

One of the key features of the different approaches was the extent to which they were 'judgement' or 'evidenced' based. The majority of case study approaches (and indeed approaches to measuring distance travelled more widely) were based on the judgement of the (most commonly) individual client, although some included third party assessments. Only two projects included evidence-based approaches, and a series of perceived barriers were cited against the suitability of such approaches with typical client groups. The distribution of approaches is summarised below:

- **Judgemental approaches - beneficiary** – five approaches were based primarily on beneficiary judgement, most commonly through self-reporting approaches. A key risk associated with this approach is the influence of the state of mind of the beneficiary on a particular day. Attempts were described to allow for this, for example asking the client to think about a particular situation (e.g. “how confident would you feel if you were at a college for the first time?”), rather than “how confident do you feel today?”. There are, however, disadvantages in relying solely on self-assessment approaches.
- **Judgemental approaches – project worker/third party** – three projects described approaches that rely in part or wholly on the opinions of supervisors/third parties. However, there may be issues of subjectivity and fairness, which can be aggravated by the tensions of working with “difficult” client groups in some situations. For example, one project mentioned examples of supervisors using assessment forms to criticise clients, due to frustration, friction or anxiety (especially where clients' motivation for attending the course is questionable).
- **Evidence based** – two projects include the collection of hard evidence of progress (e.g. completing a task with a concrete output, such as preparing a CV). While these approaches were useful, other projects described a series of barriers suggesting appropriate evidence will not always be available/applicable. One project noted that some of their partners do not have a culture of collecting evidence, which would be time-consuming/difficult to change. In such circumstances, a judgement-based approach may be the best alternative.

Clearly the project's context (in terms of aims, objectives and activities) will be key in whether evidence or judgement-based approaches are suitable. Projects need to consider if measuring perceptions is enough to satisfy their aims (whether 1st or 3rd party, or both). Ideally evidence would be used to support measurement, given the risks associated with approaches based solely on beneficiary judgement (which rely on respondents' understanding of the questions/principles of the approach, their willingness/ability to answer honestly, and to take part in the process in the first case). The beneficiary's state of mind at the time of measurement is also important, although attempts can be made to account for this as described above. Third party judgement

can play a useful part, but as also shown above may itself be subject to issues of subjectivity and equity. While the example cited above was of negative assessment, there may be cases where third parties inadvertently give clients 'false positive' ratings based on emotional/attachment issues.

While the use of judgement or evidence-based approaches may depend on a number of factors, from the relevance/suitability of different 'measures' of progress to the delivery culture within which a project and its partners is set, a key consideration is the level of resource available to the project.

4.2.7 Resource issues

In many cases, it is clear that 'one-off' projects will have neither the time or the finances to develop a distance travelled measurement system from scratch. Indeed, one of the common factors amongst the case study and stakeholder approaches studied is that they were either 'proprietary'/commercially developed systems or had been developed over multiple funding rounds.

It was impossible to provide an estimate of the cost of developing the different systems examined, although it was clear that the investment in their development varied widely. Some approaches required considerable investment in developing both the indicator framework and the IT platform on which it sits. Even the most simple system featured considerable design, development and piloting costs. However, as discussed previously, the involvement of a wide range of staff in system development and piloting can be a key success factor, and ensure that any such 'investment' will be worthwhile.

The introduction and initial interpretation of a new system will also have resource implications, as the summary box below describes.

Introduction and training issues

Many case study projects referred to the time taken to introduce systems to staff and train them in their use. Some also noted resistance by staff who initially saw the administrative work required as taking time away from their work with participants, or even dislike of an approach that could be seen as devaluing clients by assigning a score to them. (It is notable though that there was almost negligible opposition to assessment systems by the various project workers interviewed). ESF projects introducing systems for the first time need to be aware that additional time will be required to introduce, train and address issues raised by staff. An organised piloting period is essential, with a senior manager acting as a champion. One approach was introduced in the following way:

- A general introduction was held at a meeting for all project workers.
- A pilot was conducted with four individual clients before the system was used by other staff.
- For the first six weeks, fortnightly meetings were then held to discuss issues arising and to answer staff questions.

While this example referred to a relatively complex approach, simpler systems may require less time or effort. Nonetheless, a planned roll-out of any new monitoring system will be needed.

In addition to the time required for introducing the system, further time may need to be devoted to ensuring consistency in its use by different project workers, and to ensure that the same evidence is not interpreted differently. An internal verification system may be required, with a senior project member acting as a quality manager. This is particularly important in multi-site projects, and should be applied to new staff as they join the project.

Having been developed and introduced, the systems also varied considerably in the amount of time required for implementation. For example:

- One IT-based system was found to take between half and one hour to complete, depending on the beneficiaries' levels of literacy and understanding. This would be repeated every two to three months with each client, and required the presence of a project worker to offer help if needed.
- Another project uses an intensive system which requires beneficiaries to fill in assessment forms before and after every training session (though each session may last one or more days).
- A third project's interviews could take up to two hours, reflecting in part the fact that the client group includes those with learning difficulties who may require substantial guidance.
- The least time intensive approach to deliver required beneficiaries to answer six questions, taking a minimum of time to complete.

5 CONCLUSIONS AND IMPLICATIONS

This final chapter provides the main conclusions of the postal survey and project case studies, and their implications for the guidance produced to complete the assignment. Both are provided in summary format below.

5.1 Conclusions

- The high response rate for the postal survey, alongside subsequent practitioner contacts, illustrated a **considerable interest in the measurement of soft outcomes and distance travelled** amongst the projects and promoting organisations. This interest is shared across the respondent group, be they measuring soft outcomes or not.
- The postal survey also identified a **larger share of respondents describing measuring soft outcomes than initially expected**, with over half describing current measurement approaches. The postal survey also identified a number of **common approaches** – both single systems used by a range of respondents, and organisations using their own systems across a range of projects - many of which may be firmly embedded in operational approaches. The case study fieldwork allowed this issue to be examined in greater detail.
- The measurement of soft outcomes was found to be **fairly evenly distributed by policy field and beneficiary focus**, and there was **no apparently link between the propensity to measure soft outcomes and project policy focus/beneficiary targets**.
- Importantly, the postal survey responses showed that projects measure soft outcomes **across a range of beneficiary groupings, rather than concentrating on single beneficiary groups**. This suggests that in measurement terms beneficiaries are seen primarily as individuals progressing towards employment, rather than by their beneficiary status, and that approaches are designed to be used across a range of groups. This finding was echoed in the case study fieldwork.
- The approaches to measurement described in the postal survey **varied considerably, in terms of both content and credibility/sophistication**. This suggests that some respondents **would benefit from guidance to inform the further development of their current approaches**. The case study fieldwork allowed this view to be confirmed, and also identified a demand from even those following sophisticated and well developed approaches for further information on wider practice.
- Where soft outcomes were not being measured, the **key barriers were the lack of an available methodology, resource constraints, and (to a lesser extent) concerns over sensitivities/intrusiveness and benefits for beneficiaries**. This stresses the importance of any

guidance considering these issues, as well demonstrating potential benefits for beneficiaries. Positively, only 2% of respondents to the postal survey (five projects) felt that measuring soft outcomes offered no benefits.

- Finally, while one third of the postal respondents described plans to introduce measurement systems in the future, **eight out of ten would do so if a suitable approach was available**. While these findings refer to intentions only, they further illustrate the strong demand for guidance in this area.

5.2 Implications

- The interest in the measurement of soft outcomes, both by those currently measuring them and those who are not, **showed that strong demand existed for additional guidance**.
- **This guidance needed to be of relevance to two ‘user groups’** - organisations who are seeking to develop approaches from scratch, and those currently measuring soft outcomes but who could benefit from additional information and practice examples.
- The proliferation of different approaches followed, and the extent to which they are spread both across and within organisations, had implications for the nature of the guidance produced. That is, it may be risky seeking to ‘impose’ an approach or range of approaches on an existing active community.
- Given the interest in soft outcomes, it was considered that a **guidance document, setting out a range of practical examples from approaches currently in use as well as guidance on building a system from scratch, would best meet the needs of the two ‘user groups’** described above.
- While the case studies allowed these initial findings to be expanded and refined, the postal survey suggested that the guidance materials needed to:
 - Provide clear examples of approaches operating in the field – including examples of indicators, data collection approaches and tools, and examples of how information can be used.
 - Describe approaches which are achievable within projects’ resource parameters.
 - Show how approaches can be delivered with minimal intrusion, and avoid any features which may be off-putting to beneficiaries.
 - Illustrate the benefits of measuring soft outcomes/distance travelled for beneficiaries and delivery organisations.