

# Feasibility Study for an Evaluation of Aimhigher: P4P

A report by the National Centre for Social  
Research to the HEFCE and the LSC

## **Foreword by HEFCE and the LSC**

HEFCE and the LSC are pleased to receive this report from the National Centre for Social Research (NatCen). We would like to thank all those who gave their time to feed into the study, particularly the four Aimhigher: P4P areas who agreed to be case studies. We would also like to thank those who attended the seminar to discuss the initial recommendations in July 2003.

This report makes recommendations to us about the evaluation of Aimhigher: P4P. We are aware that this initiative will become part of the overall Aimhigher programme from April 2004. We do not propose to continue with two separate evaluation strategies, but rather to develop one single coherent Aimhigher evaluation.

This Aimhigher evaluation will be guided by the following principles. It should:

- be robust
- support practice
- have support from and be useful to a range of stakeholders (including practitioners and policy-makers)
- build on what exists, including existing Aimhigher: P4P evaluations and the Excellence Challenge evaluation
- recognise the burdens we already place on institutions in terms of collecting data and information, and avoid unnecessary bureaucracy
- not be imposed 'top-down' but be developed as part of an on-going dialogue with the sector
- operate at sub-regional, regional and national levels, and be coherent across the different levels
- support the case for investment in the initiative
- take account of equality and diversity
- not attempt to evaluate all activity which relates to progression or widening participation but only that funded by Aimhigher
- be a planned and coherent whole.

A sub-group of the Aimhigher Transition Task Group has been set up, consisting of representatives from HEFCE, the LSC, the DfES and Action on Access. This is considering NatCen's recommendations, and using them to inform a proposed monitoring and evaluation strategy for Aimhigher. This will be discussed with the regions in autumn 2003, and will feed into the guidance for Aimhigher planning.

The sub-group would welcome comments on NatCen's final recommendations. These should be sent to Fiona Reid (f.reid@hefce.ac.uk, tel 0117 931 7115).

## Contents

<b>1</b>	<b>INTRODUCTION</b> .....	<b>1</b>
1.1	Aimhigher .....	1
1.2	Objectives and scope of the feasibility study .....	1
1.3	Conduct of the study .....	3
	Familiarisation and scrutiny of documents .....	3
	In-depth interviews .....	4
	Examining the potential use of existing or planned survey and administrative data .....	5
	Applying expertise to methodological issues .....	5
	Presentation of early findings .....	5
1.4	Structure of the report .....	6
<b>2</b>	<b>CONTEXT</b> .....	<b>7</b>
2.1	Structure of Aimhigher: P4P .....	7
2.2	The regional partnerships .....	8
	Diversity of activities .....	8
	Targeting .....	8
	Monitoring .....	9
	Evaluation .....	9
2.3	Widening participation beyond Aimhigher: P4P .....	9
<b>3</b>	<b>VOLUME</b> .....	<b>11</b>
3.1	Components of volume .....	11
3.2	Measuring volume .....	12
	A typology of activities .....	12
	Activity-level monitoring .....	14
<b>4</b>	<b>OUTCOMES</b> .....	<b>18</b>
4.1	Which outcomes? .....	18
	‘Increasing’ and ‘widening’ participation .....	18
	Intermediate outcomes .....	19
4.2	Measuring outcomes .....	20
	Participant-level monitoring .....	20
	The problem of ‘intensity’ .....	24
	Links with administrative data .....	26
	National tracking surveys of participants .....	27
<b>5</b>	<b>IMPACT</b> .....	<b>30</b>
5.1	Observing change in participants .....	30
5.2	Comparison groups .....	31
	A question of targeting .....	32
	Where to find a comparison group .....	33
	Opportunity and participation .....	34
5.3	Conclusion .....	35
<b>6</b>	<b>OTHER RESEARCH: ADDING DEPTH AND DIVERSITY</b> .....	<b>36</b>
6.1	A ‘menu’ of research .....	36
	Identifying research questions .....	36
	Identifying methodologies .....	37
	How to write the menu .....	38
	How to do the work .....	39
<b>7</b>	<b>NEXT STEPS</b> .....	<b>42</b>

# 1 INTRODUCTION

In May 2003, the National Centre for Social Research (NatCen) was commissioned by the Higher Education Funding Council for England (the HEFCE) and the Learning and Skills Council (the LSC) to undertake research into the feasibility of evaluating Aimhigher: Partnerships for Progression (P4P). This report of the findings was written by Alice Bell, Andrew Shaw and Tim Knight in August 2003.

## 1.1 Aimhigher

Aimhigher: P4P is part of a multi-faceted increasing and widening participation programme aimed at both achieving the Government's target that, by 2010, 50% of those aged 18 to 30 should have had the opportunity to benefit from higher education (HE), as well as ensuring that substantially more young from under-represented groups enter HE. Initially, the HEFCE and the LSC made available funds of £60 million over the three years 2003-2006 to nine regional partnerships across England, which comprised HE institutions, further education (FE) colleges, schools and other parties wishing to contribute within this programme. These funds will be used to deliver a range of regionally-determined activities, including summer schools, mentoring schemes, staff and student visits to schools, colleges and employers and a variety of other events aimed at disseminating information, advice and guidance to a wide range of target groups in order to encourage and support them in continuing their education.

In July 2003, it was announced that an additional £18 million per annum would be made available through Aimhigher: P4P, enabling existing and planned programmes of work to expand. In addition, it was announced that, during 2004, Aimhigher: P4P was to be merged with its sister initiative, Aimhigher: Excellence Challenge, under a single brand entitled 'Aimhigher'. It has subsequently been announced that, following the merger, there will be a single evaluation of the unified Aimhigher initiative.

## 1.2 Objectives and scope of the feasibility study

It is important to understand that this report presents recommendations for consideration by the HEFCE and the LSC. It does not, therefore, represent an agreed framework for evaluation. Moreover, it must be remembered throughout that we were not asked to consider a single evaluation of the unified Aimhigher initiative. The announcement of this merger, as well as of additional funds for Aimhigher: P4P, did come in time for recommendations to be adjusted in recognition of the vastly greater scale of the merged programme. However, it falls to the HEFCE, the LSC and the Department for Education and Skills (DfES) to determine how to take forward this work in relation to the ongoing evaluation of Aimhigher: Excellence Challenge.

The overall aim of this study has been to establish whether it is feasible to estimate robustly, at a national level, the *volume* of activities taking place using Aimhigher:

P4P funds and the *net impact* of these activities, particularly on HE participation rates. The study was required specifically to:

- explore definitions of groups which are under-represented in HE and make recommendations on operationalising definitions within regional and national evaluations;
- review the evaluation activities planned by the regional partnerships in order both to assess the potential for synergy between the regional and national evaluations and to make recommendations on the conduct of regional evaluations and how these might best be supported;
- consider in detail the objectives of the national evaluation, with regard not only to the main aim of measuring volume and impact but also to the potential value of evaluating other factors such as the relative merits of different strands of activity, the process of delivering Aimhigher: P4P and intermediate outcomes for young people; and
- make recommendations for the scope and methodology of a national evaluation, including data collection methods, sampling strategies and sizes, timetable and approximate costs.

A feasibility study was required partly because it was not apparent whether a reliable measure of net impact could be achieved either at all or, even if possible, at anything like reasonable cost and/or by the time results are needed in 2006. There was no prospect that provision of the Aimhigher: P4P service would be withheld from some young people in order to facilitate experimental evaluation (that is, allocation of individuals to a 'control group' was precluded). Moreover, there were recognised to be a multitude of other influences on HE participation including demographic change, development of the 14-19 curriculum, changes in funding (e.g. the introduction of Education Maintenance Allowance and tuition fees), the broadening of HE opportunities (Access Courses, foundation degrees) and reforms of admissions policies and practices. In addition, Aimhigher: P4P is far from being the only or even the largest initiative seeking to widen participation in HE.

As the study evolved, the aims were addressed through four key questions or areas which together define the parameters for an evaluation of the programme:

- How much activity is taking place (a question of **volume**)? See Chapter 3.
- What results follow this activity (a question of **outcomes**)? See Chapter 4.
- What *difference* does this activity make (a question of **impact**)? See Chapter 5.
- What other aspects of the programme merit research and how might this research be organised (a question of adding **depth and diversity**)? See Chapter 6.

The first and third questions are re-statements of the core aim of the study. The second question, that of outcomes, is in a sense a stepping stone within impact measurement: to assess impact one has to define and measure the outcomes of interest among the group who participate. However, there is significant demand to collection information on outcomes even if impact cannot be measured, as well as

debate on whether this is worthwhile. Thus it seemed sensible to examine discretely the question of outcomes.

The fourth question recognises that to maximise the utility of national and regional evaluations for partnerships, the Funding Councils and policy makers, research may well be needed into matters such as how effectively and efficiently the programme is delivered, who becomes involved and on what basis, and who the programme fails to engage.

### **1.3 Conduct of the study**

There were four key elements to this study:

- familiarisation and scrutiny of documents;
- conducting in-depth interviews;
- examining the potential use of existing or planned survey and administrative data;
- applying expertise to methodological issues.

These are described below. These elements enabled us to present preliminary findings at an Action on Access Evaluation seminar, the response to which provided a further source of useful information for this report.

#### ***Familiarisation and scrutiny of documents***

The starting point for the research was the scrutiny of existing documentation and literature relevant to Aimhigher: P4P. This allowed the research team to become familiar with the development of policy and practice to encourage wider participation in HE and to gain a broad understanding of what the nine regional partnerships proposed to do with their Aimhigher: P4P funding.

Initial familiarisation involved reading core documents, such as the call for strategic plans (HEFCE November, 2002/49, LSC circular 02/23) and key reports on earlier widening participation (WP) activity, including an earlier study to which NatCen contributed.<sup>1</sup> The research team also benefited from personal briefings provided by the HEFCE and the LSC.

The strategic plans produced by the nine regional partnerships were also studied by the research team in order to develop questions for partnerships and others and to inform recommendations on the conduct and potential value of regional evaluations, with particular regard to the opportunities for synergy with research at a national level.

---

<sup>1</sup> *Evaluation of the HEFCE Widening Participation Support Strategy*, a HEFCE-commissioned report authored by the Higher Education Consultancy Group and the National Centre for Social Research, 2002. Available at [http://www.hefce.ac.uk/Pubs/rdreports/2003/rd11\\_03/](http://www.hefce.ac.uk/Pubs/rdreports/2003/rd11_03/)

### ***In-depth interviews***

The central element of the research was a series of in-depth interviews with a range of personnel involved with Aimhigher: P4P in various capacities. Twenty-two interviews were carried out in total, 13 by telephone and nine face-to-face.

Interviews were carried out with the following:

- Representatives of the regional partnerships. Four partnerships were selected, with the aim of reflecting different characteristics (in terms of geography, regional structure, planned evaluation activities, and existing HE participation rates) across the nine regions as a whole. Interviews were conducted with a range of people involved in the partnerships, including regional planners, individuals with responsibility for sub-regional planning, people delivering WP activities and people who had particular experience and/or expertise in evaluating such activities. In one region, representatives of Connexions and UCAS were interviewed on the basis of their involvement in some aspects of planned evaluation activity.
- Representatives of local LSC offices.
- Representatives of the team involved in the evaluation of Excellence Challenge.<sup>2</sup>
- Policy experts from both the HEFCE and the LSC.
- Representatives from Action on Access, the national co-ordination team for the over-arching WP programme.

Several of our interview respondents, including some of the representatives of Action on Access, the HEFCE, the LSC and the regional partnerships, were members of the Aimhigher: P4P National Steering Group and Regional Monitoring Groups and were consequently able to comment on the issues from more than one perspective. In addition, in order to avoid exclusion of non-selected regions from the research process, an email was sent out to all partnerships explaining the nature and purpose of the study and inviting them to send any comments or queries to the research team directly.

Interviews were conducted in accordance with NatCen quality standards, making use of as many qualitative research tools as time and resources would allow. Detailed topic guides were prepared in advance of the interviews, their content informed by discussions with the HEFCE and the LSC, and also by the initial scrutiny of relevant documentation. A general topic guide was developed for interviews with the representatives of regional partnerships and Action on Access, whilst for other respondents it was often necessary to tailor the topics discussed within the interviews to their particular role in relation to Aimhigher: P4P, and their particular areas of knowledge and expertise. Each interview was tape-recorded (with permission) to allow the research team to refer back to the tape to clarify any key points raised in the interview. Comments made within the interviews have been treated as confidential, and the names of those interviewed are not included in this report. The research team made notes on each interview, making use of a

---

<sup>2</sup> We acknowledge that Excellence Challenge has switched to use of the Aimhigher brand, but for brevity and to avoid potential confusion in this report, we henceforth use 'Excellence Challenge' to identify this programme and evaluation.

spreadsheet, or 'framework', to summarise and group comments under a series of topic headings.

### ***Examining the potential use of existing or planned survey and administrative data***

An important consideration within the study was the potential value of using existing or planned survey and administrative data in an evaluation of Aimhigher: P4P. An initial desk-based review of the most promising sources was carried out, looking at administrative data such as the National Pupil Database and major surveys including the forthcoming Longitudinal Study of Young People in England (LSYPE), the Youth Cohort Study (YCS) and the Youth Panel of the British Household Panel Survey (BHPS). This was followed up by further 'in house' discussions with colleagues possessing expert knowledge of some of the sources. NatCen has contributed to the development of the LSYPE and conducted numerous surveys within the YCS.

### ***Applying expertise to methodological issues***

The evidence collected in the first three elements of the study that are described above was examined by the research team to draw out key findings and identify factors to be considered in proposing a methodology for the national evaluation. During initial discussions of the information collated, the methodological expertise of the research team was brought to bear in order to derive potential sets of recommendations relating to the study's objectives.

Tentative recommendations were subject to further discussion and expert scrutiny within NatCen in order to focus and refine the research team's thinking. Ivana La Valle, a co-author of the 2002 report for the HEFCE on widening participation, played an important role in this process. A statistician from NatCen's Survey Methods Unit also contributed to the development of sampling strategies and sizes for potential surveys.

### ***Presentation of early findings***

In addition to the four elements of the research described above, there was a further source of input into the study. On 14<sup>th</sup> July, members of the research team gave a presentation of preliminary findings at an Action on Access Evaluation Seminar, to an audience made up of representatives from regional partnerships and other interested bodies including the LSC, the HEFCE and the Excellence Challenge evaluation team. The feedback received from those attending was valuable, both in highlighting issues that needed to be explored further by the research, and in crystallising some of the ideas that had already been formed.



## **1.4 Structure of the report**

Chapter 2 outlines the context for the study through a brief sketch of Aimhigher: P4P and the broader widening participation picture. Chapter 3 addresses the question of Volume, proposing both a typology and method of recording activities. This is followed by a discussion of Outcomes in Chapter 4, which supports the case for tracking certain participants over time in order to determine their destinations. Potential mechanisms for measuring Impact are considered in Chapter 5, while Chapter 6 outlines how research projects into other topics might be identified and undertaken in order to add depth and diversity to national and regional evaluations. A summary of the Next Steps which might be taken to progress the evaluations is the concluding chapter.

## **2 CONTEXT**

An initial aim of the research was to understand the context in which Aimhigher: P4P operates, and in which any evaluation would have to operate. This chapter highlights the organisational structure of Aimhigher: P4P, the nature of Aimhigher: P4P activity within regional partnerships, and other widening participation initiatives that exist alongside Aimhigher: P4P.

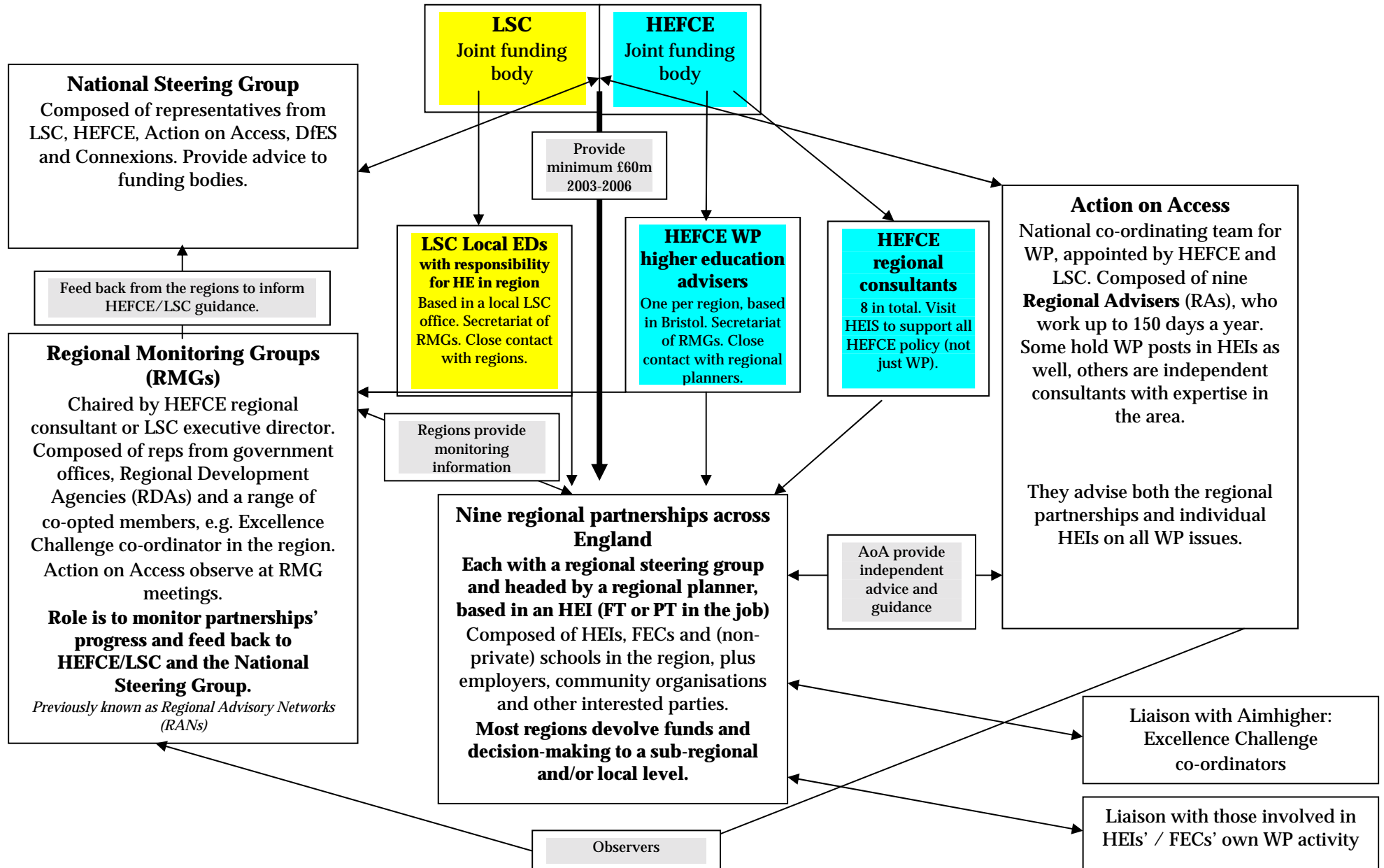
### **2.1 Structure of Aimhigher: P4P**

Figure 2.1 illustrates the structure of Aimhigher: P4P. It shows that Aimhigher: P4P is a complex and multi-faceted initiative operating within a context of widening participation activity that is even more diverse and wide-ranging.

**Figure 2.1      Structure of Aimhigher: P4P (as at 14<sup>th</sup> July 2003) [overleaf]**

# Structure of Aimhigher: P4P

As at 14/07/03



## 2.2 The regional partnerships

Through reading the strategic plans and the subsequent interviews with representatives of four of the partnerships, some key lessons were learnt about the ways in which regional partnerships are working. There are significant differences between the regions in a variety of respects, and those that appear to be particularly relevant to evaluating Aimhigher: P4P are discussed here.

### *Diversity of activities*

During our research, we observed a rich diversity in Aimhigher: P4P activities, in terms of who they were aimed at, what form they took, their duration, frequency, and intended numbers of participants. The majority of activities were aimed at young people participating in education, but there were also activities for the parents of young people, older people, disabled people of all ages, and teaching or admissions staff within schools, FE colleges, HE institutions and other organisations. There was even greater variety in terms of the different forms activities took, ranging from mentoring schemes, organised visits and tours around HE institutions, to leaflets and websites promoting participation in HE. Some activities took place on a single day whilst others were spread across a number of days, either concentrated into a small number of weeks or spread out over a term or school year. The intended number of participants also varied from small cohorts of learners from a single environment to wider projects aimed at potential HE students across a designated area in a region.

### *Targeting*

The research quickly established that there were differences between partnerships in terms of how they were targeting Aimhigher: P4P funding within their regions. Whilst there was a common emphasis across all partnerships on targeting areas of low existing HE participation, there were quite a few variations in the measures and indicators being employed to identify these areas. In some cases, basic measures such as free school meal allocations and the percentage of pupils staying on after GCSEs were used to identify areas for Aimhigher: P4P funding. Other factors such as take-up of benefits and a 'rurality' measure have also been used to prioritise areas. In a number of regions POLAR data were being used as the basis for targeting, and elsewhere statistical mapping models had been developed locally, drawing on a range of social, economic and educational indicators.

There were also variations in how Aimhigher: P4P funding is being targeted at the individual level, that is how it is decided which individuals are given the opportunity to participate in activities funded by the initiative. The most common criterion used was parental (or familial) participation in HE, with activities being intended primarily for individuals whose parents had not participated in HE. However, it appeared that in practice this was not being enforced consistently, with institutions to which learners belonged often playing a large part in deciding which individuals participated in activities. The issue of targeting has crucial implications

for feasibility and potential design of an evaluation, and a means of recording the details of targeting is proposed in Section 3.2.

### ***Monitoring***

Partnerships are required to provide basic information on all Aimhigher: P4P funded activities to the HEFCE using a standard monitoring form. Beyond this there was limited evidence of more detailed monitoring of activities. For example, there was no consistent practice of recording the numbers of individuals participating in each activity, or how successful activities were in attracting individuals from target groups. There are three particular issues that appear to be important in considering the feasibility of monitoring Aimhigher: P4P activities in a more detailed, systematic way:

- Burden on those delivering activity. More detailed monitoring would require further input from those delivering activity to keep and dispatch records of those participating.
- Difficulty of monitoring more diverse activities. There are a number of potential difficulties in monitoring activities such as poster campaigns, or information disseminated through the media, in terms of recording which individuals they may have influenced.
- Issues around data protection. There are potential issues around young people disclosing personal information, and the conditions of data protection legislation. Any monitoring system that required this kind of information from participants would have to explore these issues which may, for example, mean that parental consent would be needed.

### ***Evaluation***

All partnerships have plans for evaluation of Aimhigher: P4P in their regions, though the scope of what is planned and the preparation which has been undertaken differ greatly. There are plans in some regions for forms of evaluation which go well beyond those used in relation to the preceding WP projects – most notably plans to track people who take part in Aimhigher: P4P activities through their participation in FE and ultimately HE. In other cases, adopting or amending existing tools, such as post-event questionnaires, seemed a favoured option.

Overall there was a strong demand for greater clarification and guidance in this area. Although some thinking and planning had been undertaken, no partnership indicated at the seminar on evaluation that they had progressed so far in their own right (for example in terms of contractual commitments with regard to evaluation) that they would find it difficult to engage with the recommendations of this study.

## **2.3 Widening participation beyond Aimhigher: P4P**

An important consideration for the feasibility of evaluating Aimhigher: P4P was that it is one amongst several potential influences on participation in HE. These include:

- social and cultural factors;
- changes to funding, curricula – tuition fees, Educational Maintenance Allowance, foundation degrees;
- other WP initiatives – include HEIs’ own ‘inhouse’ efforts/policies, European Social Fund (ESF) funding, and most notably Excellence Challenge, a geographically focused but much larger programme of funding.

It is clear that these influences do not operate in isolation, with an inevitable degree of interaction between them. This complicated picture presents a major challenge to the feasibility of evaluating Aimhigher: P4P in isolation. This challenge is enhanced by the fact that different funding streams are often used to support the same activities, and that there is not always a distinction made between them by those involved in delivering the activities. Even though Aimhigher: P4P partnerships have been established relatively recently, some close working with Excellence Challenge was already reported.

Excellence Challenge focuses mainly on widening HE participation among young people aged 13-19 in areas of disadvantage. Excellence Challenge partnerships work directly with young people and provide funds for HEIs to enhance their outreach in these areas. There is also a marketing campaign targeted towards communities and families with little or no tradition of involvement in HE. The ongoing evaluation of this multi-strand initiative is itself multi-dimensional. It includes surveys of teachers and tutors, interviews with Excellence Challenge co-ordinators, discussion groups with parents and case and area studies. Most importantly, though, over 60,000 young people in both Excellence Challenge and comparison areas have participated in longitudinal surveys, the data from which is to be linked to administrative information. In this way, the evaluation aims to measure the impact of activities funded by Excellence Challenge, though not to isolate precisely the impact of Excellence Challenge funding from all other WP initiatives.

Whilst the integration of Excellence Challenge and Aimhigher: P4P means that some activities funded partly by both will now be funded by from a single source, there will still be an issue of attributing impact for activities that are also partly funded by the other funding streams operating in the WP field.

## 3 VOLUME

### 3.1 Components of volume

The question of how much Aimhigher: P4P activity is actually taking place is not intrinsically evaluative; rather, it might be viewed as an essential pre-requisite to evaluation. It will not be possible, for example, to evaluate whether a regional or sub-regional plan represents value for money without an accurate measure of exactly how that money is being used, that is a measure of the volume of activity being resourced in the area.

So what exactly do we mean by ‘volume of activity’? Clearly, it is not simply a matter of counting up the *number* of Aimhigher: P4P-funded activities taking place, as such a count would mask a multiplicity of differences between individual activities across a range of dimensions. Therefore, a meaningful measure of volume will need to provide answers to all the following questions:

- How many activities are taking place?
- What are these activities?
- When are they taking place?<sup>3</sup>
- How long do they last?
- Where are they taking place?<sup>4</sup>
- Who is participating in them?

Fortunately, most of the questions required to measure volume of activity are relatively straightforward to answer. This does not, of course, imply that there exists a satisfactory method for *obtaining* the answers, however, and we will come on to tackle that issue below when we discuss the recording of data for monitoring purposes (a measurement of volume being one key objective of monitoring). For the moment, however, we need to address the one question in the list which is somewhat more problematic to answer, namely: what **are** these activities?

The interviews that we conducted with regional and sub-regional representatives highlighted a degree of concern regarding the ways in which Aimhigher: P4P activities are defined and described (and this concern was affirmed by subsequent feedback at the presentation of early findings of the feasibility study). The ways in which activities are referred to in, for example, strategic plans or monitoring documents, do not always allow for easy and accurate interpretation; this, in turn, inhibits the possibility of making statements of the form: ‘activity A is of the same

---

<sup>3</sup> This will need to be recorded in order to measure the volume of activity within a specified time period.

<sup>4</sup> It may not be necessary to know where activities take place for a national measure of volume, but it would certainly be vital for obtaining inter-regional or inter-sub-regional comparisons.

type as activity B (they are both student mentoring schemes)' or 'activity A is not of the same type as activity B (because, although they are both student mentoring schemes, activity A is an e-mentoring scheme whereas activity B is not)'.

In order to make an accurate assessment of volume, it will be important to develop a method for generating transparent and easily comparable descriptions for individual Aimhigher: P4P activities. Ideally, this would take the form of a typology, whereby each activity can be assigned a code that defines it clearly, classifying it together with activities that are the same or very similar but also taking account of differences between relatively similar activities where appropriate. We make recommendations to this effect.

### **3.2 Measuring volume**

We make two proposals of ways to obtain an accurate and meaningful measure of the volume of Aimhigher: P4P activity taking place in regions and sub-regions. The first recommendation addresses the preliminary question, expounded above, of how to generate transparent and easily comparable descriptions of activities of different types. The second recommendation goes on to describe a way in which all the monitoring data necessary for measuring volume might be collected at individual activity level.

#### ***A typology of activities***

Our research has indicated that it should be possible, perhaps with some further investigative work, to compile a typology of Aimhigher: P4P activities which is both sufficiently comprehensive as to allow for an accurate measure of volume and sufficiently sensitive as to represent the nature of individual activities with an acceptable degree of precision. In practice, it will be important to build the typology from the 'bottom up', initially by consulting the descriptions of activities given in the regions' revised strategic plans and subsequently by seeking to clarify any descriptions which are not deemed sufficiently detailed for this purpose through consultation with regional and sub-regional representatives. We would also strongly recommend that representatives of all interested parties be allowed the opportunity to comment on the typology at key junctures in its development. At the presentation of initial findings of the feasibility study, the idea of creating a typology of activities received strong support from regional representatives.

We believe it will be possible to devise a typology which is easily comprehensible and straightforward to use. A skeleton design is proposed, with examples, at Figure 3.1. The level of detail required would need to be determined according to the level of detail at which it is considered appropriate to make assessments of volume, which will be a matter for the funding bodies and partnerships to decide. Our initial view, however, is that it would be worth classifying activities more precisely than as simply 'a summer school' or 'an open day' and also that it would be worth, where possible, including a description of the target group or groups for each activity that takes place. At the same time, it would be important to keep descriptions sufficiently general so as to avoid making the typology overly long or complex. To this end, it would also be important to emphasise to those delivering activities that the code



assigned to their activity was intended to group or classify it, rather than to describe it with absolute accuracy.

Finally, it is also worth noting the value of including an ‘other’ category within the typology in order to accommodate both the inevitable outliers and any innovation that takes place during the course of a single funding period. However, as Figure 3.1 suggests, it may be necessary to put some measures in place to restrict the use of the ‘other’ category in order to ensure the robustness of the volume measurement obtained. This would form part of a package of training materials designed to promote both reliable coding of activities and comprehensive completion of these sheets.

**Figure 3.1 Proposed skeleton design of an Aimhigher: P4P typology of activities (with example entries not comprehensively descriptive of any category)**

**(i) Activity code**

Category	Code	Description
<b>1 MENTORING</b>	1.1	Mentoring scheme using face-to-face HEI student mentors
	1.2	Mentoring scheme using email HEI student mentors
<b>2 TASTER DAY</b>	2.2	Subject taster day in HEI
	2.3	Subject taster day in FEC
<b>3 STUDENT AMBASSADOR SCHEME</b>	3.3	Student ambassador visit to school
	3.4	Student ambassador visit to workplace
<b>0 Other</b>	00	[NOTE: Activities should only be coded 00 where practitioners are adamant no other code is applicable. A full description of any activity assigned a code 00 should be forwarded to the HEFCE and the LSC to inform revisions of the typology.]

**(ii) Target group code**

Code	Participant profile	Criteria
A	Year 9 students	No family history of HE attendance
B	Year 9 students	Learning difficulties
C	Work-based learners	Studying for Level 3 qualifications
D	Work-based learners	Attending FECs
E	Non-learners	Possessing Level 3 qualifications
Z [Other]	[NOTE: Target groups should only be coded Z where practitioners are adamant no other code is applicable. A full description of any target group assigned a code Z should be forwarded to the HEFCE and the LSC to inform future revisions of the typology.]	

According to the above structure, all activities should be assigned a two-digit code denoting the type of intervention plus an unlimited number of ‘letter codes’ indicating target group(s). All codes would be ordered in the most logical sequence possible and the process of determining the correct code would be as follows:

Step 1: Identify the category of activity, e.g. ‘1’ for mentoring

- Step 2: Within that category, identify the appropriate two-digit activity code, e.g. 1.1 for face-to-face mentoring using HEI student mentors.
- Step 3: Identify the target group(s) for the activity by looking first for the profile of participants (e.g. Year 9 students) and then for the criteria necessary for participation (e.g. no family history of HE attendance). Be sure to include in the code all letters denoting appropriate target groups.

So, for example, a face-to-face mentoring scheme using HEI students and targeting Year 9 students with no family history of HE attendance would be given the code 1.1A (if it also targeted Year 9 students with learning difficulties, it would be given the code 1.1AB). In a rare case where both the activity type and the target group concerned were not listed, the coder would enter 00Z; if one were available but not the other, they should enter, e.g. 1.1Z or 00F.

Ultimately, the processes of data collection and consultation with the regions described above will help determine whether a structure like this will be workable. However, at this stage it is most important to understand the overall task: to create a typology such that (nearly) every Aimhigher: P4P activity can be assigned a code that describes it accurately and to a sufficient level of detail to meet national, regional and sub-regional data needs.

### ***Activity-level monitoring***

Assigning every Aimhigher: P4P activity a code from the above typology will be a highly informative exercise in its own right, as it will yield comprehensive data on the range of activities of different types taking place. However, as mentioned earlier, in order to obtain a full measure of volume, it will be necessary to obtain answers to a number of further questions. To this end, we believe it will be necessary for all individuals with main responsibility for delivering an activity to complete a short monitoring form. An initial design for such a form – which we have called an ‘activity summary sheet’ – is provided at Figure 3.2.<sup>5</sup>

---

<sup>5</sup> We would like to thank those respondents in the regions who allowed us access to their own activity-level monitoring forms, from which our design has extensively drawn.

**Figure 3.2 Initial design for an Aimhigher: P4P 'activity summary sheet' (activity-level monitoring form)**

**AIMHIGHER: P4P  
ACTIVITY SUMMARY SHEET**

**PLEASE WRITE CLEARLY IN CAPITAL LETTERS**

Name or title of activity/event/scheme .....

.....

Location: .....

Activity code .....

DAY      MONTH      YEAR

Dates : First or only day

--	--	--

Final day

--	--	--

Duration : Total number of hours

Participants :

Total number	
--------------	--

Target groups (specify)	
1.	
2.	
3.	
4.	

Age range	
-----------	--

Funding : Aimhigher: P4P Percentage

If other funders (P4P < 100% funding), please write in who

.....

**ACTIVITY LEADER(s) / ORGANISER(s)**

Name(s)

Organisation(s)

Address (inc. postcode) .....

Phone number(s) ..... *end of sheet*

It should be noted that the activity summary sheet would remain distinct from the sub-regional and regional monitoring returns currently required by the funding bodies.<sup>6</sup> Completing it will represent an administrative burden on Aimhigher: P4P practitioners and attention should therefore be paid to ways of minimising this additional burden.<sup>7</sup> One important element of this will be to maximise synergy both between monitoring forms at activity, sub-regional and regional level and also, where feasible and appropriate, between Aimhigher: P4P documentation and monitoring returns required by other funding bodies (e.g., ESF). A second way in which the level of bureaucracy associated with administering activities might be minimised is by allowing maximum flexibility with regard to means of data collection; we would strongly advocate making the activity sheet available to practitioners both as an electronic document (e.g. in Microsoft Word) and online.

The question of whether it is feasible to make the completion of the activity summary sheet a requirement for all those delivering Aimhigher: P4P interventions depends to a great extent on whether it can be designed in such a way as to apply to the range of different activities taking place (or, alternatively, whether a number of variations on the sheet, all collecting the same basic information, can be designed to accommodate this diversity). For example, it might be difficult for those practitioners delivering very long-term activities to specify duration as a number of hours or to pinpoint start and finish dates. In some cases, new monitoring requirements might effectively enforce changes in process, for example the need to specify the number of participants and their age range might require some practitioners to begin recording details that they have not recorded in the past.

A change in practices is particularly likely to be required in order to facilitate responses to the funding question on the activity sheet; our research suggested that the level of financial awareness among those delivering activities did not, on the whole, extend to a knowledge of the proportion of resources being channelled through Aimhigher: P4P. Therefore, the ways in which information about such matters is disseminated may have to change, especially as, even under the integrated 'Aimhigher' initiative, it seems likely that many activities will continue to draw a proportion of their funds from other sources (HEIs' widening participation allocations, for example). In practice, we suspect that such implications of substantive change are likely to lead to a 'weighing up' of the importance of the funding question, as well as other items on the activity summary sheet, against the cost and effort involved in collecting relevant data as well as the overall value of making available a robust and detailed account of the volume of activity taking place. Moreover, owing to the fact that those completing the sheets may be required to adapt to practices substantially different from any they have undertaken before, it may be advisable to put in place measures for back-checking completed activity

---

<sup>6</sup> Our research showed that, in practice, these monitoring returns rarely elicited the depth of activity-level information required to measure volume. This appeared to be the result of a combination of factors including a focus on demonstrating that plans have been pursued (hence the A-H structure) and the fact that these monitoring forms are typically completed by personnel who are one or more steps removed from the point of individual intervention delivery.

<sup>7</sup> Although it is important that the administrative burden on practitioners be minimised, we are optimistic about the level of willingness in the regions to complete something resembling the activity summary sheet, not least because practitioners have an interest in proving that their work represents value for money in order to secure future funding.

summary sheets, at least during the first phase of implementation, to ensure the data collected are of sufficient quality.

## 4 OUTCOMES

### 4.1 Which outcomes?

The second key research question that an evaluation can address is: ‘what results follow the initiative?’. In the case of Aimhigher: P4P, outcomes will be wide-ranging and measuring them is likely to involve consideration of several different subject groups: young people, adult learners, parents, communities, employers, ‘gatekeepers’, institutions, staff and potentially others. Outcomes for all of these groups provide potential for research, and it is our aim in Chapter 6 to present a model that allows flexibility for regions and sub-regions to decide the most salient groups on which to focus.

Here we are focusing upon outcomes for participants in Aimhigher: P4P activities. A range of outcomes may result following these activities and these outcomes may, in turn, vary in terms of salience for evaluation. Therefore, those commissioning and designing research will need to determine carefully which potential outcomes merit measurement. A good place to start may be to gain an understanding of intended outcomes which, at least in most cases, will be salient to measure, but unintended outcomes may often be of interest too (e.g., a summer school may not aim explicitly to attract young people into higher education by virtue of the extra-curricular activities available in HE institutions, but it may be salient to know whether young people’s attitudes to HE are altered in this way nonetheless).

While Aimhigher: P4P activity may have many different outcomes, its over-arching aim is to increase the rate of participation in higher education among 18-30 year olds to 50% by the year 2010 (through targets for HE participation rate among young adults). HE participation will therefore certainly be one outcome to measure. Whilst this may be a relatively straightforward task in practical terms, there are nevertheless two potential challenges worth highlighting.

#### ***‘Increasing’ and ‘widening’ participation***

Several of our interview respondents raised concerns regarding a tension between the aims of ‘increasing’ and ‘widening’ participation. This is a fundamental issue which logically precedes the development of Aimhigher: P4P policy, and it was not our brief to assess the policy objectives. However, the question of whether, at root, Aimhigher: P4P aims at a straightforward increase in HE participation rates or whether, on the contrary, the overall increase is of secondary importance relative to the aim of increasing participation amongst currently under-represented groups, has at least one important implication for the feasibility of evaluating the initiative.

It is probably uncontroversial to suggest that a lack of consensus about the aims of an initiative is likely to be linked to the ways in which that initiative is implemented on the ground. If those delivering Aimhigher: P4P activities or selecting participants are

not clear about the relative priorities of the distinct aims of increasing and widening participation, they will also, by implication, be unclear about what their particular activities ought to be aiming to achieve. In such cases, the most likely reaction will be some form of attempt to address both aims simultaneously, but in practice it will never be possible to give the two aims equal priority because practitioners will always have to decide whether or not to restrict participation to individuals from under-represented groups (thereby giving 'widening' highest priority) or to allow participants from outside these groups (giving 'increasing' highest priority). Potentially, this could undermine an evaluation in two possible ways, both related to the measurement of intended outcomes:

- (i) The evaluation might apply an inappropriate intended outcome measure, that is, apply a measure of increase where widening was the aim, or vice versa.
- (ii) It might not be possible to ascertain whether the intended outcome measure is appropriate or not because the tension between increasing and widening participation has led to a lack of clarity about the aims of the activity on the part of the practitioner and therefore information about intended outcomes is unavailable.

Of course, it may be that the intentions of those delivering activities are not salient in many cases, and therefore this problem – which relates only to measuring intended outcomes, not to measuring outcomes more generally – will lose its force. However, on the assumption that it might, at least in some cases, be important to be in a position to isolate intended outcomes, the tension between increasing and widening participation could yield a significant challenge.

The lack of clarity and consensus regarding the aims of Aimhigher: P4P activity represents a problem for evaluation that extends well beyond the perceived tension between increasing and widening participation. Our research uncovered widespread variation in practices aimed at identifying and targeting particular groups for intervention which are closely connected to variations in the intended outcomes of individual activities. This has implications for the measurement of impact, which we discuss in the next chapter.

### ***Intermediate outcomes***

A more straightforward, but equally challenging, issue concerning outcomes stems from the HEFCE/LSC requirement for evaluation outputs by the summer of 2006. It was immediately apparent when we began our research that Aimhigher: P4P participants come from a variety of age groups and, indeed, the HEFCE and the LSC have formally sanctioned the targeting of individuals outside the priority 13-30 age range where the activities concerned form part of a comprehensive programme. Many activities target young people in Years 8 and 9 and sometimes younger. (Opinion varied amongst our interview respondents regarding the optimum age at which to initiate widening participation intervention, but at least one well informed respondent pointed to mounting evidence that, to be effective, intervention was needed before aged 13). For this reason, it quickly became clear that data on HE entry (or lack of HE entry) would not be available for a great number of Aimhigher: P4P participants by the 2006 deadline. Therefore, it is important to consider the potential

for obtaining 'proxy' outcome data that could address the need to assess progress towards the crucial 50% target at an early stage.

The kinds of outcomes that might serve this purpose can broadly be classified as 'intermediate'. In other words, we would be looking for outcomes which would be available prior to the participant completing a Level Three qualification and which would signify, in a consistent and maximally reliable manner, the likelihood of a respondent entering HE in the future. Under the remit of this relatively brief feasibility study, it has not been possible to look in detail at which intermediate outcomes might best meet this criterion. However, generally speaking, it would be worth considering various 'hard' outcomes such as GCSE results and Level Three entry and retention as well as a range of 'soft' outcomes, e.g. changes in attitudes and aspirations concerning HE.<sup>8</sup> The likelihood is that some of these intermediate outcomes, or a combination of different ones, could provide a satisfactory proxy for HE entry data in order to meet the early data deadline. In order to assess their relative usefulness, it would be worth consulting statistical data regarding the correlation between, for example, GCSE results and HE entry and, if appropriate, some psychological literature concerning the processes and chronology involved in career decision-making. In addition, results from the Excellence Challenge evaluation are expected to be informative in this regard.

## 4.2 Measuring outcomes

It is worth stating at the outset that the recommendations in this chapter are designed to yield data about outcomes only in the first instance; the question of whether they can also yield impact data depends on the feasibility of identifying a comparison group, which is discussed in the next chapter. Our view is that there is considerable merit in collecting robust data on outcomes and related factors, even if impact measurement cannot be attained. The majority of our interview respondents expressed support for the notion of tracking participants in order to measure outcomes. The methods proposed for doing this met with a positive reaction when outlined in draft form at the aforementioned presentation of early findings.

### *Participant-level monitoring*

In order to measure the outcomes (and impact) of Aimhigher: P4P activities for their participants, it will first be necessary to devise a method for recording which individuals participate in which activities. That is, to measure outcomes one needs to know who took part and have a means of tracking them over time. We propose to do the former by means of a form, which will be filled in at the point at which participation commences as part of a 'registration' procedure (although the precise arrangements for this would depend on the nature of the activity). An initial design of a participant-level monitoring form – which we have called a 'record of participation' – is shown at Figure 4.1 (and continues over two pages).

---

<sup>8</sup> Our research found that data on attitudes and aspirations was being widely collected in the regions and we would emphasise the value of such data when obtained by means of a well-designed research instrument.



**Figure 4.1 Initial design for an Aimhigher: P4P 'record of participation' (participant-level monitoring form)**

**AIMHIGHER: P4P  
RECORD OF PARTICIPATION**

**PLEASE WRITE CLEARLY IN CAPITAL LETTERS**

Name or title of activity/event/scheme .....

.....

	DAY	MONTH	YEAR
Date of event (start date if more than one day).....	<input type="text"/>	<input type="text"/>	<input type="text"/>

Activity code (from typology).....

**YOUR DETAILS**

First and middle names

Surname

Female <input type="checkbox"/>	OR	Male <input type="checkbox"/>	Date of birth	DAY	MONTH	YEAR
				<input type="text"/>	<input type="text"/>	<input type="text"/>

Address .....

Postcode.....

Phone number(s) .....

Name and postcode of your school/ college/ employer /other institution

If you are not sure what to enter here, please ask your activity organiser;  
if you are not in education or work, please write 'None'.

Name of institution

Postcode.....

**NOW PLEASE TURN OVER** →

**BACKGROUND INFORMATION**

We ask you to answer two questions to help us monitor and plan our activities. But if you prefer not to answer, that's OK.

1. Did your parents – including step-parents or guardians – study for a degree or equivalent higher education qualification (at a university or other institution)?
- |                       | ‘Mother’                 | ‘Father’                 |
|-----------------------|--------------------------|--------------------------|
| Yes                   | <input type="checkbox"/> | <input type="checkbox"/> |
| No                    | <input type="checkbox"/> | <input type="checkbox"/> |
| Don't know / not sure | <input type="checkbox"/> | <input type="checkbox"/> |
2. Do you think **you** will study for a degree or equivalent higher education qualification at any time before you are 30 years old?
- Definitely.....
- Probably.....
- Probably not.....
- Definitely not.....
- Not sure.....

**RESEARCH**

The information you have provided about yourself will be treated as confidential. We may wish to contact you as part of research into Aimhigher: P4P or to inform you of other Aimhigher activities. It would always be your decision whether or not to take part in any research or other activities. We may also wish to link your details to administrative data for research purposes only. Your details will NOT be used for commercial purposes. For example, we will never pass information on to anyone who wants to sell you something. If you are not willing for your details to be used for research, please tick this box

**SIGNATURE**

Please sign and date this form to confirm your participation in Aimhigher: P4P.

**Signed**

**Date**

.....

**\*\*\*THANKS FOR COMPLETING THIS FORM\*\*\***

The first section of the record, including the name, date and code of the activity concerned, could be completed either by the participant or, more probably, by the practitioner responsible for delivering the activity.<sup>9</sup> We envisage that measures could be taken to make the completing of this section as simple and quick as possible: for example, the record could be made available as an electronic document so that the practitioner could fill in these preliminary details on a computer and then print out the number of copies required with the first section pre-filled. It may also be worth considering making the record available online, although this would require participants to have internet access prior to the commencement of the activity (our design does not allow for completing the record after the intervention takes place) and extra measures might need to be taken to make sure that all participants did log on to complete their sections, particularly if they were to do so from home.

The section entitled 'your details' requires participants to provide some basic information about themselves. Most of the questions here are straightforward although the activity organiser would be well advised to have the addresses of the participants' institutions/employers to hand as many participants may not know these. The record of participation has been designed on the basis of a typical event targeting students at schools or colleges or other individuals taking part with the consent of their employer. However, Aimhigher: P4P funds can, as we have already mentioned, be used to target other groups, e.g. parents, institutional staff and community groups. It may be that, for these groups of participants, the design of the form would need adjusting. For example, it would not be appropriate to ask parents for the names and addresses of their 'institutions', but it might, instead, be worth asking for some details of their children's school(s) or college(s) so that parent and child data could be matched at a later stage. For such reasons, the record of participation we have proposed should be viewed as a template to which amendments will be required to render it applicable to all Aimhigher: P4P participants.

On the second side of the record of participation, we included two 'background information' questions which, again, will be appropriate to 'learner' participants but not to all other groups. The first question aims to collect a basic core indicator of the participant's target group status by asking whether either of his/her parents have studied for a degree-level qualification. These data can then be used to assist the sampling of participants for more detailed tracking (although further data from other sources are also likely to be required for this purpose). This question will elicit data on participants' perceptions of parental HE attendance rather than parental HE attendance *per se*. It is hard to estimate what the level of accuracy will be here, but that need not be of grave concern as it seems likely that a participant's perceptions will be at least as good an indicator of an individual's pre-intervention disposition towards HE as actual parental attendance in HE.

The second question in the 'background information' section will collect a relatively basic indicator of participants' views regarding the likelihood of their entering HE before turning thirty. This question will provide basic baseline data which can be compared with data collected at a later stage by means of follow-up tracking research. Having answered this question, the participant is offered the opportunity

---

<sup>9</sup> The activity code would be drawn from our recommended typology, described in Section 3.2 above.

to withhold his/her permission for the details entered on the form to be used for research purposes.<sup>10</sup> Finally, the participant will be asked to sign and date the form to confirm the verity of the information they have provided.

Further consultation is required on the contents and administration of the record. For instance, ethnic group is not on the example presented, though we recognise the strong case for including this. The form also needs to be piloted among participants and, preferably, tested cognitively among a small but diverse sample. Detailed guidance on the purpose and administration of the form would also need to be developed.

### ***The problem of ‘intensity’***

Though the completion of the record of participation will clearly represent an additional burden on both practitioners and participants, our research did not uncover strong or widespread resistance to the idea of individual-level monitoring for the purpose of measuring outcomes (and potentially impact). Nevertheless, there will be some challenges associated with the record of participation, stemming from the diversity of the work funded through Aimhigher: P4P. It seems likely that variations on the template record would be required to accommodate the range of potential participants in Aimhigher: P4P activity. In addition, there is a further, potentially more serious, problem for the record, concerning not the diversity of participants but the diversity of activities themselves.

Whilst it may be possible to design different versions of the record of participation to cope with the wide range of people taking part in Aimhigher: P4P activities, it would be harder to design a form that could collect participant-level data for every activity that takes place. Put simply, the problem here is one of ‘intensity’: participants in activities which might be classified as ‘low-intensity’ (e.g. open days, ambassador visits) are likely to be harder to monitor than those taking part in ‘high-intensity’ activities (summer schools, taster events, mentoring schemes and so on) for the simple reason that it is typically more difficult to record details of participants in the low-intensity cases.

The problem of intensity is different to the problem of diverse participants because, rather than having implications for the *design* of the record, it brings into question whether, in some cases, the record can be administered at all. Three possible ways of responding to the problem of intensity are explored below.

### **Response one**

We accept that the record of participation can only be administered for a sub-group of Aimhigher: P4P activities, namely those classified as ‘high-intensity’. In other words, participant-level monitoring information can only be collected for those activities where it is practically feasible to ask every participant to complete a form

---

<sup>10</sup> Under normal circumstances, we would expect such an ‘opt-out’ clause to be sufficient for the form to comply with the Data Protection Act. However, the situation may be slightly different when collecting information from minors, and it is always advisable to have any such forms approved by a lawyer prior to implementation. We do not, however, anticipate that data protection issues would prohibit the administration of such a form altogether with respect to any groups of Aimhigher: P4P participants.

before the activity commences. We do not attempt to obtain participant-level data for 'low-intensity' activities.

**Advantages:**

A clear response to the problem of intensity, avoiding the need to try to collect detailed participant data for low-intensity activities. Contains the administrative burden.

**Disadvantages:**

Any data obtained by means of participant tracking would only be able to yield findings on the outcomes of high-intensity intervention. Our research (particularly the feedback we received at the presentation of early findings) indicated that most practitioners would consider this an incomplete evaluation. Aimhigher: P4P is as much about long-term, cumulative programmes of low-intensity activity (the 'drip-feed' approach) as it is about one-off, high-intensity interventions such as summer schools: to fail to include the former would be to mis-measure the initiative.

**Response two**

We find ways of obtaining a completed record of participation from every participant in every Aimhigher: P4P activity, regardless of the level of intensity.

**Advantages:**

Yields a comprehensive set of participation data from which realistic and wide-ranging findings concerning outcomes can be drawn.

**Disadvantages:**

In the best-case scenario of low-intensity activity (imagine, say, an open day), it might be possible to ensure that every participant completes a record of participation and here the disadvantage is confined to a (potentially huge) additional administrative burden. However, in the worst-case scenario (a leaflet, say, or a website), it seems unlikely that every 'participant' might complete such a form.

**Response three**

We accept that the record of participation can only be administered for a sub-group of Aimhigher: P4P activities, namely those classified as 'high-intensity', but attempt to obtain participant-level data for 'low-intensity' activities by other means, primarily through retrospective questions in a follow-up survey (e.g. "have you ever attended an open day at a university or other higher education institution...?", "have you ever seen an advertisement [like this] for 'Aimhigher'...?").

**Advantages:**

Yields a relatively comprehensive set of participation data from which relatively realistic and wide-ranging findings concerning outcomes can be drawn whilst limiting the administrative burden on practitioners and participants.

**Disadvantages:**

Relies on respondent recall regarding participation in low-intensity activities, which may lead to under- or over-estimation of levels of participation. Does not allow for sampling on the basis of participation in low-intensity activities.

Overall, we believe that Response Three is the most practical and pragmatic route to take. It represents a realistic assessment of what constitutes a reasonable burden for participants and practitioners whilst acknowledging the vital role played by low-intensity activities in the overall Aimhigher: P4P picture. In view of the level of support for some kind of outcome measurement relating to participants, Response

Three seems to us to entail the lowest level of sacrifice necessary to reach this goal. In addition, the attractiveness of this or any option for obtaining participant-level data depends to a large extent on where you draw the line between 'high-intensity' and 'low-intensity' activities. If the majority of the most common activity types can plausibly be classified as 'high-intensity', the sacrifice involved in accepting that the record of participation cannot be administered for low-intensity activities might in fact transpire to be very small. Indeed, on the basis of our research, our view is that it will, given a certain amount of flexibility and willingness on the part of practitioners, be feasible to ensure that participants in most types of Aimhigher: P4P activities complete a record of participation. Moreover, a well-designed follow-up questionnaire should be capable of collecting data on the remaining (low-intensity) activities to a satisfactory level of reliability and precision. The question, therefore, of how attractive Response Three is ultimately deemed to be will rest largely on the definitive answer to the question of where to draw the line. Compiling a typology of activities as we have proposed ought to make it easier to reach such an answer.

There is scope, therefore, for flexibility in determining how the record of participation might best be implemented. However, once decisions on this are reached and if it is agreed to proceed, administration of the forms has to be mandatory across all partnerships. Selective, partial administration and, hence, significantly incomplete data would undermine the objectives of what we acknowledge to be a burdensome exercise.

These records will provide a database of participants who then require tracking over time in order to observe outcomes. This can be done through linking administrative data and/or conducting surveys. Linking with administrative data may also offer the potential to add further baseline information on participants.

#### ***Links with administrative data***

The record of participation will, as we have shown, collect data about Aimhigher: P4P participants' perceptions of parental HE attendance, one core indicator of an individual's target group status. However, it would help to know more about participants' characteristics and background. In addition to improved descriptive information on participants, additional data would i) help with structuring the sample for the tracking survey (see below) and ii) improve the analysis of outcomes which are measured by identifying more sub-groups of interest. The nature of the data required for each of these purposes will be determined through discussion about and clarification of the definitions of target groups. The list will be likely to include ethnicity, disability and parental occupation. If it were possible to obtain this kind of information for Aimhigher: P4P participants, these details could be matched to those given on the record of participation (or to multiple records in cases where an individual had participated in more than one activity) to give a much richer participant profile than is yielded by the record alone.

Although our investigation into possible sources of data for matching has not been detailed, we are fairly confident that there are a number of potentially fruitful sources available, at least for the main groups from which Aimhigher: P4P participants will be drawn. One promising source of data on young people, for example, is the National Pupil Database, held centrally by the DfES, which contains

information about all pupils on maintained school rolls in England, combining the Pupil Level Annual Schools Census (PLASC) and relevant records maintained in schools themselves. The National Pupil Database certainly contains several data items which might assist in the sampling of Aimhigher: P4P participants for tracking and other follow-up research. Other databases that might merit further investigation include those compiled by Connexions. However, through our limited research into these, we perceived some potential drawbacks. It was not entirely clear whether Connexions would hold information on every pupil in a maintained school, and there was also the problem of a lack of uniformity across the Connexions services in England (although Connexions databases may become standardised in the future).

Some investigation into accessibility of administrative data sources held outside the sphere of the HEFCE and the LSC will be necessary in order to assess the viability of this kind of data matching, and the data protection implications of any information transfer will need to be carefully considered as well. However, the greatest area of difficulty associated with data matching is likely to centre around finding suitable available sources for all the diverse groups of Aimhigher: P4P participants. Outside schools, the Individual Learner Record (ILR) contains much detailed information about individuals taking courses through FE colleges, and would definitely be worth consideration, but Aimhigher: P4P is likely to reach groups who are not connected with particular schools or colleges, and sourcing appropriate administrative data for these wider groups will almost certainly pose more of a problem. Information about learners who are taking part-time courses, for example, or those who are not studying towards recognised qualifications, seems, from our initial research, to be less readily available; data on those not currently learning may not be available at all. For such reasons, possible methods for obtaining a rich profile of all Aimhigher: P4P participants will be a key consideration in planning a programme of evaluation research.

These databases might also be used to obtain intermediate outcomes for participants. Other sources, such as UCAS data, also come into play in this respect. Though the potential for administrative data linking is vast, this has to be set against drawbacks with regard to coverage (e.g., UCAS data only covers HE applicants for full-time courses) and considerable practical obstacles to bringing the data together. Therefore, while we recommend pursuing these opportunities, we also believe that a survey of a representative sample covering all types of participants is necessary.

### ***National tracking surveys of participants***

We strongly recommend that a sample of Aimhigher: P4P participants be selected for follow-up survey interviews on the basis of information given on a record (or several records) of participation plus, if possible, additional data sourced from administrative databases. There are several reasons for carrying out such a survey, in addition to measuring outcomes:

- To obtain additional data to enrich the profile of participants further and to enable a greater number of sub-group analyses.
- In particular, to enrich the profiles of those participants (if any) for whom administrative data are not available for matching.

- To obtain the sort of data that will not be available from administrative sources for any participants, e.g., a depth of data regarding attitudes and aspirations and feedback on the programme.

A major advantage of the survey method based on sampling from participant records is that it can encompass all types of participants, regardless of their environment, because it depends neither on links with other data nor on access through institutions. However, in order to yield statistically robust findings for key sub-groups a follow-up survey would probably have to be fairly large. A representative, unclustered achieved sample of 4,000 participants at first interview would both give quite precise estimates for the whole Aimhigher: P4P population and support a good deal of reliable sub-group analysis, for example by type of activity or ethnic group. For the whole sample, 95% confidence intervals would be less than  $\pm 2\%$ ; for sub-groups of 1,000 confidence intervals for estimates would be up to  $\pm 3\%$ , while for groups of 400 the 95% confidence intervals would be  $\pm 5\%$  or less. The sample size per region would average nearly 450, sufficient to provide broad feedback across the range of a region's activities but probably insufficient for reliable figures on different activities and certainly insufficient for sub-regional or robust inter-regional analysis.

There is, therefore, a case for a larger sample of, say, 9,000 interviews in the first round, which translates, of course, into an average of 1,000 per region. As we recommend in Chapter 6, we feel each region should have the option of funding a larger sample to obtain detailed analysis in its own right. The question here is whether the Funding Councils might wish to commission a sample of this size to ensure that such analysis and, hence, inter-regional comparison is facilitated across the country.

To achieve this level of precision it would be important to have an unclustered (or at least only modestly clustered) sample, given that Aimhigher: P4P provision varies substantially according to geography. For this reason and for cost-effectiveness, we recommend the survey be conducted by telephone. Hence, the quality of telephone number collection on the participation form, together with participants willingness to help with a phone survey, are critical. We judge that, for this population of young people, data collection by telephone will prove relatively effective and certainly cost efficient. We would hope a response rate of 60% or above could be achieved. Clearly the representativeness of that achieved sample is crucial to the quality of the estimates. There are likely to be some biases, but the information available from the record of participation should help with detecting and, possibly, correcting these.

Survey interviews could sensibly be conducted approximately nine months after the sampled participation in Aimhigher: P4P (ignoring any subsequent involvement in other activities). An alternative to consider at the design stage would be to interview all respondents in October or November, in order to maximise the collection of relevant outcome data.

Participants in this first survey would then be 'tracked' over time. A second approach would be made one or two years later to update outcome data. We would expect that about 3,000 respondents could be re-contacted and would agree to a short second interview.



It may be possible to add value to the survey by linking data with one or more existing or planned quantitative projects. The most promising such survey looks to be the Longitudinal Survey of Young People in England,<sup>11</sup> which will track an initial sample of some 15,000 young people selected from the National Pupil Database from Year 9 at least until they turn 25 years of age. There is, we understand, unlikely to be space to expand the questionnaire for the first 'wave' of the survey, which is scheduled to take place in the spring of 2004 and will comprise a young person interview as well as an interview with each young person's parent or guardian. However, questionnaires for later waves may well have room for questions to be added and, owing to the size of the sample, would be likely to capture information about a substantial number of Aimhigher: P4P participants<sup>12</sup> and about young people in any comparison group that might be identified. Moreover, the subject matter of the LSYPE – which explores, among a wide range of other topics, the young person's attitude towards education and future plans – is likely to provide a suitable context for the kinds of questions an Aimhigher: P4P evaluation would require.

In addition to the LSYPE, there are a number of other ongoing survey programmes onto which a follow-up of Aimhigher: P4P participants might 'piggy-back'. These include the Youth Cohort Study survey series (DÆS) and the Youth Panel of the British Household Panel Survey. However, as with the issue of identifying administrative data sources, we suspect it will be more difficult to identify surveys pertinent to participants who are not in the 'youth' bracket or who do not possess some kind of affiliation to a school or college. Further investigation will therefore be needed here, and it may transpire that the only way of reaching some groups will be to design one or more smaller-scale parallel surveys from scratch.

---

<sup>11</sup> The invitation to tender for Waves 1 to 3 of the LSYPE had been issued by the DÆS at the time of writing this report, but a research contractor had not yet been confirmed.

<sup>12</sup> Let us provide an *illustrative* example of the numbers which cannot, at this stage, be known. In 2001, there were 660,000 10 year olds in England. So the Aimhigher: P4P targets to raise the young participation rate by about 5 percentage points by 2010 imply roughly 33,000 extra HE entries in this cohort. How many of these young people would Aimhigher have to engage to achieve this? Maybe four times that if one assumes half of those engaged entered HE and a realistic 50% deadweight (see next chapter for definition of this term) amongst this half. On the other hand, some impact is likely to be indirect, through teachers, parents, FEC and HEI staff etc. Even so, involvement of less than 66,000 seems minimal if the target is to be achieved. That equates to 10% of the age group, which implies about 1,500 respondents in the LSYPE, if this group is represented proportionately. This is a fair base for analysis, with quite a large sample within which to find comparable young people.

## 5 IMPACT

The question of impact is closely related, but not identical, to the question of outcomes. In order to measure impact, it will be vital to obtain outcome data, but it will also be necessary to go a step further by asking the question: *what would have happened otherwise?* To put it slightly differently:

**the impact of X is those outcomes that would not have occurred without X**

So if, for example, an Aimhigher: P4P activity inspires an individual to enter HE who would not have done so had s/he not participated in this activity, then that extra HE entrant can be counted as contributing to the impact of Aimhigher: P4P. If, however, that participant was already destined to enter HE prior to the intervention then clearly this should not count as an impact of the scheme; or, in technical terms, all such instances comprise the 'deadweight' associated with the programme. The net impact (or 'additionality') attributable to Aimhigher: P4P equals the successful outcomes minus this deadweight.

The previous chapter illustrated the value of collecting outcome data, and this is not affected by the feasibility or desirability of attempting to measure impact. The above example concerning HE entry provides a clear illustration of the special challenge associated with the further step of introducing an impact measurement to the evaluation, namely: how can we possibly know what would have happened otherwise?

### 5.1 Observing change in participants

One method of measuring impact involves collecting data about the group of interest before and after the intervention takes place. The idea is that you measure, say, an individual's attitude towards HE prior to intervention and again after the intervention is complete and infer that the change (or 'distance travelled') constitutes the impact of the intervention.

In principle, it might be possible to use this 'before-after' method for measuring the impact of Aimhigher: P4P on participants in activities although, given the long-term cumulative effects of (particularly low-intensity) activities, it could be difficult to pinpoint the times at which it might be most appropriate to collect the relevant data. There are, in addition, two further substantial problems with this type of design. The first is the difficulty of disentangling the impact of the intervention from change which might anyway have occurred between the two points in time at which data are collected. In other words, this method is only really a solution if one can be confident that the measurements taken beforehand are either unlikely to have changed in the absence of the intervention or to have changed in predictable ways. The gravity of this problem depends partly on the length of time that elapses between the stages of data collection. If data were collected, say, before and after a summer school took place, the degree of natural change would probably be negligible or, at least, easily controlled for in analysis. However, this assumption almost certainly would not be

reliable over any significant period for young learners whose preferences and decisions with regard to their future are likely to be in flux. Given, as mentioned above, the aim is to measure impact over a longer period of time, this fundamentally undermines this approach.

The second problem is more practical in nature. We have already touched on the diversity of groups and individuals likely to participate in Aimhigher: P4P activity and, for this reason, it might be difficult to source 'before' data that could be considered sufficiently robust as to yield a meaningful measure of impact. The example cited above referred to attitudinal data and it is feasible that these could be collected from most participants, regardless of their profile (although the practicalities of this would also depend on the nature of the activity). However, in order to obtain a robust measure of impact, it would probably be necessary to make use of 'hard' data as well. Given the diversity of those who participate in the initiative it is unlikely such 'hard' measures would be universally available (consider, for example, those returning to learning after a lengthy absence from education).

For some participants it may be possible to get round these difficulties by making use of predictions of future attainment made before participation in Aimhigher: P4P. Specifically, if GCSE predictions could be shown, in the absence of participation, to correlate well with actual grades for similar pupils, then any divergence from this correlation amongst a group of participants might provide a valuable intermediate measure of impact. This is one of a number of 'partial' measures of impact that might be pursued should a global net impact estimate be judged unattainable.

## **5.2 Comparison groups**

In our view, the most favourable method for estimating the counterfactual would involve identifying a 'comparison group' (a term preferred to 'control group', which carries the implication of randomised allocation of individuals within a trial, which is not possible in this context). This will entail finding a group of individuals who possess similar key characteristics, and whose circumstances are similar in a range of important respects, to the group with which we are concerned (the 'action group'). Some further investigation might be needed to determine the criteria for comparison group eligibility for an evaluation of Aimhigher: P4P, but the list would likely include key measures of propensity towards HE entry such as parental occupation and the presence or absence of a family history of HE attendance.

### ***A question of targeting***

The feasibility of using a comparison group method for measuring the impact of Aimhigher: P4P depends to a great extent on the degree and nature of activity targeting. (Of course, Aimhigher: P4P as a programme is designed to target sub-groups of potential HE entrants from within the whole population; the issue here is targeting – or rationing – of provision within these sub-groups because resources do not permit opportunities to be offered to all potential participants). In order to identify a comparison group, it is vital to have some method of determining which individuals are eligible for participating in Aimhigher: P4P activities and which, despite similar characteristics, are not. Our research has shown this to be a major hurdle facing the feasibility of measuring impact.

The targeting of Aimhigher: P4P activity appears typically to take place in two stages. First, an Aimhigher: P4P practitioner will identify a set of institutions (e.g., schools or colleges) to target and then, through negotiation with personnel at these institutions, a set of individuals will be identified according to certain criteria and offered the opportunity to participate. Our research showed that, at both stages of the targeting process, certain core concepts tended to be in play. For example, when identifying institutions, the HEFCE POLAR data and details of free school meal allocations were regularly used. Similarly, the criteria by which institutional personnel were encouraged to select individuals for opportunities often seemed to focus on parental HE attendance or occupation. However, crucially, there was also a great deal of evidence of variation in the criteria used for targeting, the methods by which these criteria were operationalised and the extent to which desired criteria were in practice fulfilled. To take the individual-level example: at the least serious end of the scale, it might be the case that, say, one sub-region was asking institutions to select young people whose *parents* hadn't attended HE whereas another was asking them to select young people who had *no family members* with a history of HE participation. At a more serious level, however, an Aimhigher: P4P practitioner might have encountered such difficulty in persuading the institution concerned to apply any criteria whatsoever to their selection of young people that, in the event, potential participants ended up being selected on a simple 'first-come-first-served' basis.

For evaluation purposes, the targeting of Aimhigher: P4P activity need not be standardised, in the sense of all regions targeting similar groups, but similarly described target groups should be consistently and transparently defined and operationalised in order that like-for-like comparisons can be made. Our research highlighted a need for considerable work in this area, which would have to be prioritised in advance of launching an evaluation involving a comparison group and which might require both further investigation into regional and sub-regional practices and, ultimately, a national directive of the sort that has not, thus far, characterised the management of the initiative by the funding bodies. For a national measurement of impact, it would be necessary to know how both institutions and individuals were being targeted for activity; in other words, it would be necessary to understand both the criteria and the methods used for targeting, and also to be confident that the chosen criteria were being fulfilled.

On a more general level, the very notion of targeting Aimhigher: P4P activity inspired a degree of resistance amongst some interview respondents, who felt that it was not a concept appropriate to the initiative, which they viewed as 'filling any gaps that needed filling'. Other respondents felt that quite stringent targeting was required to make effective use of the committed resource. Our view was that, in practice, high intensity activities would have to be targeted, while recognising that partnerships may be committed to universal provision of some less intensive elements of their programme. Of course, all these views pre-date the confirmation that substantial additional resources will be forthcoming. It is too soon to know how this will affect provision, but the availability of comparison individuals is bound to be diminished. Before returning to this key factor, let us pursue the discussion of the comparison group methodology.

### ***Where to find a comparison group***

The next question to tackle would be where to find a sufficient number of sufficiently similar individuals to select as a comparison group. There are various options here, including:

- Selecting the comparison group from non-targeted institutions.
- Selecting the comparison group from a non-targeted age cohort within targeted institutions, e.g. from the year group preceding the first year group to be offered the opportunity to participate in Aimhigher: P4P activities.
- Selecting the comparison group from the same age cohort(s) within targeted institutions but only including those individuals who did/will not participate owing to the rationing of opportunity.

Selecting a comparison group from similar yet non-targeted institutions is preferred, as it is least likely that non-participants might, indirectly, have been impacted by the programme. The number of such institutions thus becomes critical. Geographically focused initiatives, such as Excellence Challenge, are especially likely to ‘run out’ of comparison institutions. It seemed unlikely that Aimhigher: P4P would be able to cover all or nearly all low participation areas but perhaps this will now be possible.

If so, one could turn to the second option of, effectively, seeking a comparison among those who are a little older than those to whom the extended Aimhigher provision will be offered.<sup>13</sup> In this case, however, it would be advisable to bear a couple of points in mind. First, the age span and range of learning environments of participants suggests this would be a hugely complex exercise. Secondly, selecting a comparison group from a non-targeted age cohort introduces a risk of contaminating the impact measurement with unrelated systematic differences between the action and comparison groups. For example, if there had been some significant change to funding or curricula which might have influenced a person’s propensity towards entering HE, and which occurred at a point in time that meant it may have affected the two age cohorts in systematically different ways, then the scope for attributing differences in the two groups’ outcomes to Aimhigher: P4P would be considerably curtailed.

For the third option, we are not concerned with the question of who would be eligible for targeting (because the comparison group would, by definition, have to be eligible in principle), but rather with the question of whether there would be a sufficient number of eligible individuals within targeted institutions who did not participate in Aimhigher: P4P activities due to a rationing of places. This question (which is akin to the first question of whether a comparison group could be found in non-targeted institutions) brings us back to the complex area concerning the extent to which activity is targeted and the ways in which targeting varies according to type of activity. There is also the issue of indirect effects on non-participants in organisations

---

<sup>13</sup> At the time of this report, the research team working on the evaluation of Aimhigher: Excellence Challenge were considering alternative ways of sourcing a comparison group owing to the extension of the programme, including selecting individuals from within targeted institutions.

with participating individuals. In view of these factors, the third option looks unpromising.

However, further research could usefully discover whether there are significant differences between those individuals who are offered and take up the opportunity to participate in Aimhigher: P4P activities and those who turn down such opportunities or are otherwise inhibited from taking part. This will involve not only consideration of the characteristics of the two groups, but also some investigation into the degree of choice potential participants have over whether or not to participate. Fortunately, the set of factors influencing take-up of opportunities is likely to be a topic of interest in its own right, and the methodological recommendation we go on to make in Chapter 6 acknowledges the potential need for answers to questions of this type.

### ***Opportunity and participation***

Throughout our discussion of comparison groups, we have talked about attempting to identify a set of people who, for various reasons, had not been (or would not be) given the opportunity to participate in Aimhigher: P4P activities. The assumption has been that such a group could legitimately be compared to an action group made up of individuals who had participated (or would participate) in activities, so long as they were sufficiently similar in a range of key respects. There is, however, a potential pitfall here, which concerns the distinction between those who are given the opportunity to participate and those who actually take up that opportunity (assuming that these two groups are not identical, that is, that participation is not universally mandatory). In order to be able to draw conclusions about the differences between an action group and a comparison group, clearly we must be as confident as possible that we are comparing like with like and, for this reason, if there would be significant differences between an action group made up solely of participants and an action group that included some individuals who were given the opportunity to participate but, for whatever reason, did not take it up, then this must be mirrored in the make-up of the comparison group.

To put it another way, if there is something systematically different about people who say 'yes' to the opportunity and people who say 'no', then we have two alternatives:

- (i) We include some of the 'no' people in our action group and stick with a comparison group made up of individuals who are similar to those targeted but not targeted themselves.
- (ii) We restrict our action group to 'yes' people whilst applying extra criteria to the selection of our comparison group to try to identify only people who, had they been given the opportunity to participate, would have been likely to take it up.

Option i) here implies the need to know who is given the chance to take part; either records of opportunity (rather than solely of participation) would be required or screening tools would need to be administered to identify such individuals. Both seem impractical.

Option ii) is challenging but there are methods such as propensity score matching which are used to enhance the reliability of the comparative method. This requires identification of a large group of young people with characteristics apparently similar to Aimhigher: P4P participants, who would then be sampled and surveyed in order that sophisticated matching of this sample with members of the participants' survey could then take place. The comparison survey needs to be at least as large as the participants' survey, preferably larger to allow for some redundancy in the matching process.

If a well-matched sample could be produced, one would be confident of detecting even a fairly modest impact of Aimhigher: P4P. More specifically, if the successful outcome (HE entry) rate in the comparison group were, say, 25% and the true success rate in the whole Aimhigher: P4P group equal to or greater than 30%, then it is highly likely that this survey design would record a statistically significant difference between the two groups.

If the impact of Aimhigher: P4P were much greater, as it probably needs to be to meet the core participation target, then the question becomes not so much one of detecting an impact but rather how precisely this impact is measured. The 95% confidence interval for the net impact would be in the order of +/- 2.5 percentage points. That is, if the success rate on a chosen outcome measure among the Aimhigher: P4P survey respondents was 35% as against, again, 25% in the comparison group, then clearly the central estimate of impact is 10 percentage points and, importantly, the 95% confidence for this estimate is approximately 7.5 to 12.5 percentage points.

However, well-matched samples are difficult and relatively costly to produce and, even then, prone to a degree of uncertainty about just how well-matched are the two groups.

### **5.3 Conclusion**

On present knowledge, attempting a robust measurement of the global net impact of Aimhigher: P4P does not seem justified. The resources required, the risks involved in the estimation process and the likely margins of error seem to outweigh the potential value of the information.

We would suggest, rather, that Aimhigher: P4P awaits the substantive and methodological results of elements of the Excellence Challenge evaluation involving comparison sample designs. In light of these findings, it might be worth pursuing partial measures of impact, that is, for some activities and/or groups.

## **6 OTHER RESEARCH: ADDING DEPTH AND DIVERSITY**

So far, our recommendations have focused on measuring the volume of Aimhigher: P4P activity taking place, the outcomes of those activities for participants and potentially their impact (on participants) as well. The final stage of our proposed research strategy aims to provide scope for adding depth to the investigation of these three issues and for encompassing a range of other topics which may merit research, including matters of process and other salient issues. The suggested approach aims both to broaden the scope of the evaluation in order to gain a fuller, more inclusive understanding of the initiative and to allow more study of activities and outcomes where this is required. Moreover, our proposed strategy for providing depth and diversity of research aims to be as flexible as possible in order to accommodate the different and wide-ranging priorities of various parties involved in the initiative and to allow responsiveness to the changing shape of Aimhigher: P4P over time.

### **6.1 A 'menu' of research**

In order to accommodate the variety of areas in which research and evaluation might be deemed appropriate, we propose that research projects should be 'purchased', potentially both nationally and regionally (or sub-regionally), from a centrally-held 'menu' of research. The idea here is that the menu will include a set of potential research projects considered to be both feasible and, in principle, desirable for understanding and evaluating Aimhigher: P4P. The level of detail included in the menu is likely to depend both on the extent to which it is considered appropriate to determine the precise nature of the research projects prior to purchase and on the level of understanding of research methodology that can be assumed to exist in all regions. However, we would envisage a brief description of the key characteristics of each research item including, for example, basic methodology, proposed sample composition, potential outputs and approximate unit costs.

#### ***Identifying research questions***

Clearly, the menu of research could include as many different kinds of project, targeting as many different population samples, as met the criteria of feasibility and desirability. The first step in determining the set of projects which met these criteria would be to identify a list of research questions to which those involved in Aimhigher: P4P would like answers. As mentioned above, it may well be the case that there are additional questions relating to volume, outcomes and impact which could not be answered by the research strategies proposed so far, and one key function of the menu would be to allow scope for pursuing these to the depth desired. For example, thinking about outcomes, it should be acknowledged that there are a huge number of ways in which Aimhigher: P4P can bring about change, which would not necessarily be explored by tracking participants alone. For example, the influences of Aimhigher: P4P activities on parties other than participants could range from something as specific as a curriculum addition in a college or HEI through to effects as general as an overall cultural shift within a



community, workplace or school, and these kinds of outcomes seem likely to merit research over and above the outcomes we have referred to earlier in this report.

In addition, there are likely to be a number of salient questions not relating to the three major issues of volume, outcomes and impact, and this is where the role of the menu in facilitating diversity of research comes into play. For example, many of the partners we spoke to expressed a desire for information relating to what might broadly be called 'process': what is actually happening in the regions?, what's working and what's not?, how are partnerships functioning? Partners are interested in finding out about structures and methods utilised in other parts of the country and in sharing best practice across all aspects of process, ranging from recruiting partners to staff development to matters of delivery and administration. Finding out what works, over and above outcomes and impact as they relate to activity participants, is clearly at the forefront of their minds.

Finally, the menu of research projects is expected to include some strategies aimed at investigating a set of 'why' questions: why does someone take up an opportunity to participate in Aimhigher: P4P?, why do some people get the opportunity to participate and not others?, why do some partners get more involved than others?, and so on. Answers to such 'why' questions are instrumental in the success or otherwise of the initiative and it will therefore be important to consider these as part of any attempt to 'prove or improve' the work that partners undertake.

### ***Identifying methodologies***

Having identified a list of salient research questions, the next step will be to consider what types of research methodology will best facilitate the acquisition of relevant data. As a starting point, it seems likely that the menu will include the potential to expand or deepen existing national projects aimed at measuring volume, outcomes and/or impact. For example, it should include an option for regions to 'purchase' larger sample sizes within a national tracking study in order to enable a more detailed, robust analysis or to obtain greater focus on particular activity types or target groups of learners. In addition, there are research methodologies which we have not yet mentioned but which emerged from our research as being strong candidates for the menu; these are listed below.

- **Feedback forms:** forms completed by participants immediately before and after interventions to measure the immediate reaction to a single activity. Feedback forms were already being used for several activities in the regions we spoke to.
- **Case or area studies:** focused, in-depth, qualitative investigations. Case studies could be used to look at a range of issues including the processes involved in delivering particular activities; outcomes relating to particular institutions/employers/communities; and reasons underlying involvement or lack of involvement of particular partners.
- **In-depth interviews and focus groups:** qualitative exercises focused on particular individuals or groups. Again, these methods would be appropriate to researching a variety of topics including reasons underlying take-up of opportunities to participate in activities; views and experiences of partners; and outcomes for 'gatekeepers'/parents/communities/employers/institutions.

Could also help to add depth to the understanding of outcomes for participants that will have been gained from the quantitative work.

- **Additional surveys:** We have already mentioned that survey work is likely to be necessary for measuring the outcomes and impact of Aimhigher: P4P for participants (especially those for whom administrative data are not available for matching), but surveys could also be useful for investigating other parties where there is an interest in quantifying outcomes, attitudes, experiences and so on. It might be appropriate, for example, to survey institutions, employers or community groups as well as staff delivering Aimhigher: P4P.

### ***How to write the menu***

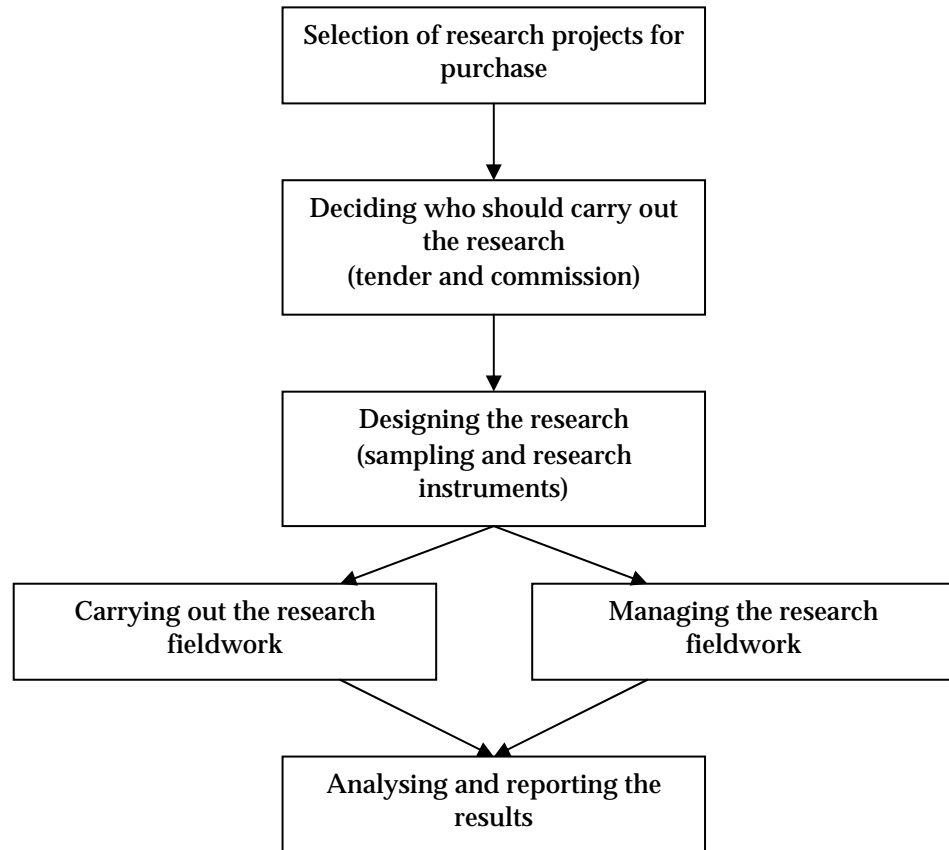
We have seen that compiling the menu of research will require, first, the identification of a set of salient research questions and, second, consideration of what methodologies might best be deployed to answer these. The resulting menu should include several research projects which are both feasible and potentially worthwhile. The best way of ensuring it does so will almost certainly involve approaching its compilation in a 'bottom up' manner. In the first instance, using the revised regional plans to compile a fresh list of existing and planned evaluations in the regions and sub-regions is likely to be helpful. It should not automatically be assumed that all existing and planned projects will ultimately be eligible for inclusion in the menu, however, and this gives rise to the important question of how to deal with existing and planned projects that fall short of the menu's requirements. Given the co-operation which will be required of partnerships to implement the recommended evaluation strategy, it seems unlikely the Funding Councils would wish to proscribe such initiatives. Equally, to facilitate these projects with the research management function envisaged for agreed projects might undermine the concept of a structured, quality assured menu. Perhaps advice on the merits and conduct of 'DIY' research could be offered, while making it clear that responsibility for such expenditure remained with the partnership concerned.

The second stage of development of the menu will involve consultation with regional and sub-regional representatives, both about the kinds of topics they would be interested in researching and the kinds of research outputs they would ideally be seeking to obtain. In carrying out this feasibility study, we have gained the impression that levels of understanding of research principles and methodologies at regional and sub-regional level are rather variable. Some regions have the benefit of considerable expertise, perhaps even employing a research officer of some kind, whilst others seem to be without a source of substantial knowledge in this area and, in some cases, we even saw evidence of a lack of basic understanding, for example, of the difference between evaluation and monitoring. The state-of-play in individual regions and sub-regions seems, in some cases, to be dependent simply on the personnel involved in managing and working within the partnership. For this reason especially, it may be worth considering supporting the consultation process aimed at developing the menu by providing guidance on evaluation research to help inform and enrich discussion. It may also be necessary, in the final stages, to supplement the menu compiled by means of this 'bottom up' process with some further suggestions generated at a national level.

### ***How to do the work***

Once the menu has been compiled, a set of procedures will need to be in place for managing the various stages of the research process. Figure 6.1 shows the main tasks involved:

**Figure 6.1**      **Flow chart: the research process**



In order to devise a set of procedures for implementing the menu of research, a decision will need to be taken regarding where responsibility for each stage of the above process should lie. One way of approaching this decision is to view the possibilities as existing along a continuum ranging from ‘overall national control’ at one end to ‘overall regional control’ at the other. Of course, when we talk about ‘overall control’, we do not intend to imply that, even at the most extreme points on the continuum, either the national or regional perspective would be sidelined altogether: if it were decided that overall control of the research menu ought to be nationally maintained, we would still expect regional representatives to be given the opportunity to comment and provide feedback on national plans; equally, if the regions were to take responsibility for most elements of the research process, we would wish to highlight the importance of some nationally-driven guidance and staff development activity.

Ultimately, the decision of how to implement the menu will have to be arrived at through more extensive consultation with regional and sub-regional personnel than

we have been able to undertake. However, on the basis of our research, we would favour an approach that probably sits slightly further towards the 'national' end of the continuum of possibilities (though certainly not at the extreme). The main reasons for this are, first, the variable levels of research expertise in different regions we have already mentioned and, second, our impression that many practitioners working in the regions, whilst interested in and willing to contribute to evaluation activity, do not view it as their priority and do not necessarily wish to dedicate a great deal of their time and energy to research (especially not if they were required to do so at some cost to intervention delivery and other kinds of 'outreach' work).

Our initial recommendation is as follows:

### **Selection of projects**

Regions (or sub-regions) select projects for purchase on the basis of a range of factors including:

- their particular research priorities (topics and methods);
- perceived usefulness of outputs;
- desire to co-ordinate with other regions' selections in order to obtain comparative data (this may require some processes to be put in place for inter-regional consultation);
- costs and budget.

We would anticipate that there would be a second stage to this selection process, whereby the funding bodies would review the regional selections and decide whether to supplement these with additional national funds, either to fund certain projects in regions that have not chosen to purchase them in order to obtain more extensive comparative data or to fund projects that have not been selected by any region because it is deemed that particular outputs would be useful at a national level.

### **Tendering and commissioning work**

Selected research projects would then be put out to competitive tender by the national funding bodies. This would reduce the burden on regions at this stage and also ensure a certain level of quality and robustness from the outset. It would also provide an opportunity for research teams based in the regions (or consortia based across the relevant set of regions) to tender for work, perhaps even on the basis that regionally-based bids would be viewed favourably owing to the value of localised knowledge in guiding the design and execution of projects. Regions and sub-regions would be kept fully informed throughout the tendering process, particularly with regard to any amendments to predicted costs.

### **Designing the research**

Sampling strategies and research instruments would be designed by successful contractors. It would be important for each contractor to liaise with a named Aimhigher: P4P contact at the national level, but we would also recommend that representatives of the regions buying into the project(s) concerned be given the opportunity to comment in detail on draft questionnaires, topic guides and so on at

all key stages of their development. This could be facilitated by the formation of a steering group for each project, whereby all regional, sub-regional and national components would be represented at meetings taking place at key stages in the research process.

**Carrying out and managing fieldwork; analysing and reporting findings**

The latter stages of the research process would all be the responsibility of successful contractors, again liaising primarily with a research manager employed at the national level. During these phases, members of the steering group could expect to receive regular updates on progress with fieldwork, preliminary headline findings following the completion of fieldwork and the opportunity to comment on at least one draft report.

This kind of combination of national and regional/sub-regional responsibility should result in a situation whereby regions and sub-regions are heavily involved in choosing and contributing to research but are not required to commission, design, carry out or undertake day-to-day management of projects, leaving them free to focus primarily on delivering activities. Also, the 'menu' model ought to allow the level of flexibility required by the wide range of research interests and priorities associated with Aimhigher: P4P.

## 7 NEXT STEPS

The overall aim of this report was to explore the main issues and challenges facing the task of designing and implementing an evaluation of Aimhigher: P4P and to come up with a set of recommendations representing our best response to these. Having presented these recommendations, it just remains for us to outline what we consider to be the appropriate next steps, should the Funding Councils decide to take forward the recommendations.

The most pressing task will be to develop a single coherent evaluation of Aimhigher, since it will not be feasible to isolate Aimhigher: P4P from Excellence Challenge from April 2004. We believe that the recommendations in this report provide a valuable starting point for the development of such an evaluation.

In our view, it will also be necessary to undertake a range of preparatory tasks in the remainder of 2003 if the 2006 data deadline is to be met. The range of tasks that can feasibly be undertaken will clearly be governed partly by resources and availability, but we would suggest that the following ought to take priority:

- Compilation of a typology of Aimhigher: P4P activities, based initially on regions' strategic plans and subject to review by regional and sub-regional representatives.
- An audit of existing evaluation work at regional and sub-regional level, in order to kickstart the design of a 'menu' of research and to aid consideration of the short-term potential of existing work.
- Consultation on and amendments to the design of the 'activity summary sheet' and 'record of participation', followed by some kind of piloting to assess the workability of these on the ground.
- Consideration of the potential for linking participant data to administrative databases, in particular a close review of the work of the Excellence Challenge evaluators in linking to the National Pupil Database.
- Monitoring progress of the LSYPE, with particular regard to opportunities to propose questions for inclusion in the second and subsequent round of interviews.

If the activity sheet and record of participation are trialled successfully, they ought then to be rolled out as early as practicable in 2004, given that activity in 2003 will already have passed by before this recording system is in place.

Then, following the integration of Aimhigher: P4P and Excellence Challenge, survey design can begin, including, consideration of the period – that is, dates – of participation from which a survey sample should be drawn and the timing of first interviews. By that stage, it would also be desirable to have begun putting in place the national infrastructure necessary for administering the 'menu' system, so that centrally-driven national research projects could be supplemented by other work

aimed at achieving depth and diversity from the time of launching this new programme for evaluating Aimhigher.