Research Report DFE-RR035

Feasibility study for a quantitative survey: how activity at 16-18 impacts on future pathways

Final report

Cheryl Lloyd, Victoria Brown, Emily Tanner, Jon Hales and Rosalind Tennant

Department for Education

DWP Department for Work and Pensions

Contents

1 Summary

1.1 Background

The Government's commitment to raise the age of participation in education or training to 18 by 2015 has reframed the debate about the group of young school-leavers who are not in education, employment or training (NEET). Despite an increase over the past fifteen years in the proportion of young people who participate in full-time education, a persistent minority remain NEET.

While there has been abundant research on young people who are NEET in the first years following the end of compulsory education, less is known about the *medium* term impact of being NEET on the transition into adulthood. Research studies that *have* focused on unemployed people in their twenties have tended to lack background data on their activity status at younger ages and subsequently it has not been possible to explore the relationship between being NEET at 16-18 and subsequent activities. Following the recession, the policy interest in the future implications of being NEET at this age has increased further. This feasibility study assesses the potential of various data sources to provide a sample frame for investigating the medium term impact and addressing this gap in knowledge.

1.2 Aims of the research

This study was commissioned to assess the feasibility of undertaking a quantitative survey to measure the medium term impact of individuals spending a period of time not in education, employment and training (NEET) at ages 16-18 as compared with being in employment or education. This involved investigating suitable data from previous surveys which could serve as sampling frames and conducting a pilot survey.

1.3 Pilot survey

Sampling

Following an investigation of a number of data sources, the Educational Maintenance Allowance (EMA) pilots and Youth Cohort Surveys (YCS) were identified as the most suitable sample frames for the pilot. NatCen conducted these longitudinal surveys with young people some years back when they were aged between 16 and 20 and by January 2010, the respondents to the cohorts used were aged between 24 and 27. YCS and EMA collected information about young peoples' activities, background characteristics and attitudinal data.

In order to compare the trajectories of adults according to their activity status between 16 and 18, respondents were allocated to the following sampling groups on the basis of the YCS and EMA data.

1. Full-time education
2. Job with training
3. NEET with identifiable barriers
4. NEET

5. Job without training with identifiable barriers
6. Job without training

For the pilot survey, 90 respondents were selected. Six interviewers were each assigned 15 cases comprising a mixture of the different strata groups. Interviewers were asked to trace and make contact with each of their issued sample, and if possible carry out an interview. Interviewers were asked to aim to achieve five interviews each.

Tracing and response

The methods used for tracing and making contact with sample members included sending an introductory letter, calling at the address, telephone calls and making contact with 'stable address' contacts provided at the last interview.

Considering the time since the last interview and the limited time for fieldwork, the tracing of former respondents was successful and a good response rate was achieved.

Across a three week period, interviewers were able to trace and carry out interviews with 35 young adults from the issued sample of 90. With additional time and interviewer resources, it is possible that more interviews could have been conducted. A little under half of the issued sample could either not be traced, had deceased or did not wish to take part.

Based on feedback from interviewers it is estimated that in 76 of the 90 issued cases (around 4 out of 5) some information was obtained about the whereabouts of the young adults.

The interview

Face-to-face interviews were conducted with a computer-assisted questionnaire (CAPI). The interview covered the following topics:

- Introduction and activity diary
- Details of current activities – 'outcomes'
- Demographics and life events
- Details of activities at 16-18 – starting points
- Pathways: activity transitions and trajectories
- Advice
- Activities and achievements (looking back)
- Activities and ambitions (looking forward)

The median interview length (including the completion of the activity diary) was 64 minutes, ranging from 32 minutes to 109 minutes.

1.4 Sampling considerations

The analysis of the YCS and EMA longitudinal data sets highlighted a number of issues that would need to be considered in drawing a sample for a main survey:

- The approach for defining an individual's activity status between the ages of 16 and 18. For the pilot survey, this was based on main activity at the time of interview, but it is also possible to take into account the activities reported during the year. Related to this is the issue of how many waves of YCS/EMA data should be drawn on. For the pilot survey, the sample was based on final wave respondents, but further analysis of attrition between waves suggested that later waves of the surveys were less likely to include respondents who had barriers to employment, education or training or those who were potentially susceptible to deprivation. However, there are ethical considerations with regard to recontacting respondents from all eligible waves.

- Ideally, the young people classified as NEET should be divided into subgroups according to how long they remained NEET, but there are various ways in which this could be achieved.

- Power calculations were carried out to estimate the size of sample groups that would be necessary to detect statistically significant differences in the outcome variables of interest.

1.5 Feasibility and recommendations

The response achieved during the pilot survey suggests that it would indeed be possible to trace adults who had responded to previous surveys when in their teens and to collect detailed information about their activities since leaving Year 11 in a face-to-face interview.

Recommendations were made for sampling, fieldwork and questionnaire development on the basis of the pilot study and investigation of the sampling sources. Alternative data collection methods were also considered, including a qualitative study, with discussion of the relative merits.

2 Introduction

2.1 Background

The Government's commitment to raise the age of participation in education or training to 18 by 2015 has reframed the debate about the group of young school-leavers who are not in education, employment or training (NEET). Despite an increase over the past fifteen years in the proportion of young people who participate in full-time education from 57% to 64%, a persistent minority remain NEET[1]. In fact, despite the DCSF PSA target to reduce the proportion of 16-18 year olds NEET by two percentage points from 9.6% in 2004 to 7.6% in 2010, the latest evidence shows that the NEET rate had risen to 10.3% by the end of 2008[2], albeit with variations between local areas (Ofsted, 2010).

There has been abundant research on young people who are NEET in the first years following the end of compulsory education at 16 including longitudinal cohort surveys and evaluations of policy interventions designed to increase participation. The Youth Cohort Study (YCS) has covered this age group for more than 20 years with fresh cohorts of young people being recruited at two yearly intervals and has given rise to segmentation analysis which highlighted the diversity of NEET young people and the high rate of churn in and out of NEET status (Spielhofer et al, 2009). Evaluation studies, such as of the Education Maintenance Allowance (EMA) and the Activity and Learning Agreement Pilots, have tested approaches to re-engaging young people in education and training (Rennison et al, 2005, Tanner et al, 2009, Perry et al, 2009). Research has also explored barriers to participation (Spielhofer et al, 2009, EdComs 2007).

Less is known about the *medium* term impact of being NEET between the ages of 16 and 18 on young people as they make the transition into adulthood and move through their twenties. While a number of programmes and research studies have focused on unemployed people in their twenties, these have tended to lack background data on their activity status at younger ages and subsequently it has not been possible to explore the relationship between being NEET at 16-18 and subsequent activities. Following the recession, the policy interest in the future implications of being NEET at this age has increased further. This feasibility study assesses the potential of various data sources to provide a sample frame for investigating the medium term impact and addressing this gap in knowledge.

[1] DCSF: Participation in Education, Training and Employment by 16-18 Year Olds in England, June 2009
[2] *Op. cit.*

2.2 Aims and scope of the feasibility study

The feasibility study was commissioned to:

> *"… provide an assessment of the possibilities, limitations, estimated costs[3], and practicalities associated with undertaking a quantitative survey to assess the medium term impact of individuals spending a period of time not in education, employment and training (NEET) at ages 16-18, and an assessment of comparative outcomes for those in jobs without training, jobs with training, and full time education between age 16 and 18."*[4]

The key stages of the study were as follows:

Stage 1: An investigation of the candidate sampling frames, including considerations of the numbers of young people within different activity groups and sub-categories of NEET young people.

Stage 2: Methodological considerations, including contact methods, mode of interview and response.

Stage 3: A pilot survey to test the methods, response rates and the quality of information collected.

Stage 4: Recommendations for a full-scale survey including method and sample issues based on the above.

2.3 Structure of this report

The next chapter outlines the pilot sampling strategy, fieldwork procedures and findings from the pilot survey. The sampling chapter (Chapter 4) provides further investigation of the two suggested sampling frames by focusing on attrition between waves of the EMA and YCS studies, considering how NEET subgroups could be defined, and sample sizes. Chapter 5 outlines our recommendations for sampling, fieldwork and questionnaire development. Chapter 6 discusses the feasibility of carrying out a main survey with face-to-face or telephone interviews, and sets out an alternative qualitative study.

[3] Indicative costs were provided separately to this report.

[4] DWP/DCSF research specification.

3 Pilot survey

The pilot survey aimed to address the following questions:

1. To what extent is it possible to trace respondents from the contact details provided during their last EMA or YCS interview?
2. How willing are respondents to take part in this further research?
3. Of what quality is the information collected during the interviews?

The pilot also provided an opportunity to test using the activity diary as an aide to respondents recalling their activities since the end of Year 11 and also give an indication of further areas for questionnaire development.

The pilot interviews were carried out during a three week period in February – March 2010 by six experienced NatCen interviewers in six different areas.

3.1 Sampling and interviewer assignments

Sample sources

Following an investigation of a number of data sources, the Educational Maintenance Allowance (EMA) pilots and Youth Cohort Surveys (YCS) were identified as the most suitable sample frames for the pilot[5]. These data sources provide information about young peoples' activities, background characteristics and attitudinal data across a number of waves when the young people were aged 16-20 years[6]. The table below summarises the EMA and YCS surveys which could be used as potential sample frames, showing respondent ages at the time of previous fieldwork and the beginning of 2010.

[5] The Longitudinal Study of Young People in England (LSYPE) would also be a potential sample frame for a main survey (or indeed provide information by itself about the employment trajectories of young people). However, the cohort of young people participating in this survey would not be in the age of interest (aged 25 years or over) until 2016 which is the year that that the last planned wave of the LSYPE is due to take place.

[6] The EMA sample had the advantage of yielding a good number young people who had been NEET or in JWT during their teens since the fieldwork focused on local authorities that had relatively low rates of participation in post-16 education and training. While the sample may not be representative of all young people in their situation, it nevertheless spanned a range of areas and had the advantage of the efficiency achieved by such clustering.

Table 3.1 EMA and YCS Survey data

Survey	Wave	Year of interview	Respondent age at interview	Respondent age Jan 2010
EMA Cohort 1	1	1999/00	16-17	27-28
	2	2000/01	17-18	
	3	2001/02	18-19	
	4	2002/03	19-20	
EMA Cohort 2	1	2000/01	16-17	26-27
	2	2001/02	17-18	
	3	2002/03	18-19	
	4	2003/04	19-20	
YCS 10	1	2000	16-17	26-27
	2	2001	17-18	
	3	2002	18-19	
YCS 11	1	2002	16-17	24-25
	2	2003	17-18	
	3	2004	18-19	
	4	2005	19-20	
YCS 12	1	2004	16-17	22-23
	2	2005	17-18	
	3	2006	18-19	
	4	2007	19-20	

Activity groups for sampling

Preliminary analysis suggested that the EMA surveys and YCS cohorts 10 and 11[7] yielded sufficient information to categorise young people into activity groups based on their employment, education and training at a number of different points in time and background information to identify those who were at risk of being NEET or in JWT due to identifiable barriers.

Data about young peoples' main activities were used to allocate the young people to one of the following categories for sampling purposes[8]:

1. Full-time education
2. Job with training
3. NEET with identifiable barriers
4. NEET
5. Job without training with identifiable barriers
6. Job without training

Defining main activity status in a consistent way across the EMA and YCS samples was a challenging task as the above categories were not explicitly defined in either survey. For the EMA respondents,

[7] The data from YCS 12 was not available at the time of the feasibility study.

[8] These groups were developed based on key groups of 16-18 year olds identified in the specification:

1. Those who were NEET for some or all of this period,
2. Those who went into paid work with training,
3. Those who went into paid work without training,
4. Those who remained in full-time education, including vocational studies.

they were assigned to activity groups on the basis of their current activity status and training variables at the point of interview at each wave spanning the ages 16-20. NEET status took priority, so a young person was categorised as such if they were NEET in any of the survey waves.

Categorising YCS respondents was more straightforward as there was a main activity status variable in each of the waves that included the above categories. For the purposes of the pilot study, additional categories (part-time education, government supported training, other education and training and other) were excluded. Due to time constraints in preparing the sample and the fact that YCS was used as a top-up, YCS respondents were categorised according to their activity at the final wave when they were aged 18-20. So those who were NEET at the final wave of cohorts 10 and 11 (when aged 18/19 and 19/20 respectively) were allocated to the NEET sample group.

A limitation of the pilot sample design is that the final wave respondents were included in the main activity definitions for both EMA and YCS. This meant that in the case of the EMA survey and YCS 11 final survey, the activities of respondents aged 19/20 were included in the classifications. In a main survey, respondents should be classified according to their activities in the earlier waves of the survey only, in order to meet the specification of comparing the outcomes of young adults according to their activities at 16-18.

The NEET and job without training groups were divided according to whether the young people had identifiable barriers to being in employment, education or training in order to assess whether those with barriers were more difficult to trace, and therefore if they would need to be over sampled in a main survey. Young people were classified as having identifiable barriers according to certain characteristics recorded at wave 1: persistent truancy in Year 11, long-standing illness, disability or infirmity or being a lone parent. These characteristics (and others) have been shown in the research literature to be associated with reduced participation in employment, education or training (Spielhofer *et al* 2009; Nuffield Foundation 2008:22)[9]. Other key predictors identified in the literature (GCSE attainment, gender and location) were not used to identify the 'barriers' group as the number of young people falling into these categories was considered large enough that they would be represented in the other strata. Learning difficulties was not included in the 'barriers' strata as it was too difficult to identify this group consistently across the YCS and EMA surveys.

Pilot sample

In order to select the pilot sample, the EMA districts which had enough respondents in each of the strata were identified and then from this list, six were selected where there were available interviewers. The sample was primarily from the EMA data (81 of the 90 issued cases), supplemented with YCS cases.

Each of the six interviewers was issued with an assignment of 15 cases, comprising a mixture of the different strata groups. Interviewers were asked to trace and make contact with each of their issued sample, and if possible carry out an interview. Interviewers were asked to aim to achieve five interviews each.

[9] Note also that the Connexions Client Caseload Information System (CCIS) has been used to classify the NEET group into three categories; 'not really NEET', 'barriers' and 'no barriers'. The barriers group includes those with children, an illness or disability. For more information see CCIS Requirement v1.0 Dec 2008 (available at http://www.dcsf.gov.uk/everychildmatters/Youth/clientcaseload/ccis/) and the NEET Statistics Quarterly Brief February 2010 (available at http://www.dcsf.gov.uk/rsgateway/DB/STR/d000913/index.shtml)

The distribution of issued sample across the six sample groups is shown in Table 3.2. The NEET groups were intentionally over-sampled based on the expectation that they may be harder to trace and less willing to take part.

3.2 The briefing

Interviewers attended a briefing with researchers and a member of the operations team. Interviewers were given background information about the study, and briefed on the procedures for contacting and tracing respondents. Interviewers took part in a practice interview with researchers and completed an exercise to practice working through the activity diary.

Interviewers were asked to keep detailed records throughout the pilot and complete a debrief form for each of their issued cases. These forms included information about tracing activities, the use of stable address details from the sample, gaining co-operation and the interview. The findings below report on the verbal feedback received from interviewers during the debrief, the completed debrief forms and activity diaries.

3.3 Making contact and tracing respondents

The task of tracing and making contact with the issued sample was a key component of the assessment of feasibility. This section sets out the procedure followed in the pilot for locating the adults sampled for the study, encouraging participation in the interview and the response achieved.

Advance letter

The advance letter (example in Appendix A) was sent from the operations department at NatCen to all issued addresses a couple of weeks prior to fieldwork to allow time for the residents at the address to pass on the letter to the named respondent if they had moved. The letter provided information about the study, including the following:

- It explained what the study was about and that an interviewer would be contacting them to arrange a time for an interview.
- It reminded the respondent that they previously took part in a study conducted by NatCen and explained why we would like to talk to them again.
- It included a free-phone number and email address for the project team at NatCen.

Interviewers were provided with spare copies of the letter to give to respondents and a laminated copy of the letter to use on the doorstep. Interviewers said that they found it useful to be able to give out spare copies of the letter to residents in order to help gain co-operation in an effort to trace respondents.

In a number of cases where the respondent was no longer living at the issued address, the letter had been forwarded to the named contact, normally by family members still resident at the address.

Tracing respondents

Interviewers were asked to try to contact the adults in their issued sample as soon as possible to allow adequate time to trace respondents who were no longer living at the address recorded at the last interview (of the EMA or YCS surveys).

It was anticipated that a substantial proportion of the adults in the issued sample would have moved given the time lapse between the last interview when they were aged 18-20 and the pilot survey when they were in their mid twenties[10] . In many cases, the issued address was believed to be the address of the adult's parents giving scope for obtaining information about the adult's current whereabouts. If the adult no longer lived at the address, interviewers were asked to explain to the resident why they would like to speak to them, and ask if they could provide any up-to-date contact details.

If the adult lived at another address in the same area, interviewers were advised to try and make contact and carry out an interview. If they lived outside the area, their contact details were nevertheless collected so that in a main stage these cases could be issued to another interviewer working in their area.

Interviewers used a range of different strategies to trace respondents. Some interviewers initially attempted to make contact with respondents by telephone, and if this was unsuccessful follow up any stable contact telephone numbers in an effort to trace the respondent, then make face-to-face visits. Inevitably the success of this strategy was dependent on having good quality telephone numbers and interviewers noted that the quality was highly variable. A number of respondents (or respondent's parents) had kept their telephone number despite having moved home which helped with this task. Other interviewers started by making face-to-face visits to the issued addresses and either followed up on leads by visiting stable contacts nearby, or using any new telephone numbers they had been given.

Interviewers commonly spoke to one of the respondent's parents and in many cases the parent acted as a gate keeper to getting up-to-date contact details and securing the co-operation of the respondent. The reactions of parents and other family members, both at the issued address and stable addresses to being asked for contact details varied. Some family members provided new addresses and telephone numbers (normally mobile numbers) while others were more wary about doing so and either took the interviewer's contact details to pass on to the respondent or said the interviewer should call at a specific time when the respondent was expected to be visiting.

In cases where the respondent and stable contacts proved unfruitful, some interviewers spoke to neighbours and consulted local telephone directories, with varying success.

Encouraging participation

In many cases, interviewers needed to gain the co-operation of family members or other residents at the issued or stable addresses in order to get in touch with the respondent, and then encourage the respondent to participate in the interview.

[10] The pilot sample comprised young people who had participated in all waves of either the EMA or YCS surveys. Although the majority were allocated to sample groups according to their activities when aged 16-18, the contact details used for re-contacting them were drawn from the final wave when they were aged 18-20.

The family members/other residents the interviewers spoke to were generally helpful, and if they were wary about providing contact details normally agreed to give the respondent a message from the interviewer.

Interviewers generally used the wording from the letter or the project instructions to introduce the survey to try and gain participation. However, the feedback from interviewers was that this particular survey was difficult to 'sell' to respondents on the doorstep. In order to help maximise response to a main survey it would be helpful to develop key messages about the study which would increase the perceived relevance of the survey to potential respondents.

Generally, interviewers tried to personalise the survey aims to the respondents, for example by talking about what their aspirations or desires were when they were 16 and how these might have changed and what they are doing now. The £5 incentive voucher which was available after the completion of the interview was not felt to have made much difference to participation[11]. There was agreement amongst interviewers that this was a small amount so it did not appear to have much influence on the willingness to participate.

Interviewers were asked to code the reason why respondents refused to participate, where possible. The reason most commonly given was being too busy. Other reasons included lack of interest, concerns about privacy and other personal reasons.

Achieved Response

Interviewers were able to trace and carry out interviews with 35 young adults from the issued sample of 90, which is a good response considering the time since the last interview, the low relevance of the study for respondents and the relatively short fieldwork period (Table 3.2). Furthermore, there was evidence that the response rate for this pilot sample could have been exceeded in the context of a main survey. The contact details for a further eight respondents were obtained but since they fell outside the pilot interviewing areas, they were not followed up. With a longer fieldwork period, interviewers thought they would have been able to achieve a couple more interviews by following up some broken appointments and interviewing respondents who were willing to be interviewed at a later date.

As may be expected given the time which has elapsed since the contact details were collected, in around a quarter (24%) of cases no follow up address was obtained (including one person who was in prison). In a further 2 per cent no contact was made at the issued address and interviewers established that two respondents were now deceased. Twelve respondents (13%) said they did not wish to participate, either directly to the interviewer or indirectly via their parents.

Based on feedback from interviewers it is estimated that in 76 of the 90 issued cases (around 4 out of 5) some information was obtained about the whereabouts of the young adults.

The level of success in tracing respondents differed across the six strata groups (Table 3.2). The proportion of interviews achieved was lower amongst both NEET groups and the jobs without training group without identifiable barriers than those who were previously in full-time education, jobs with training and jobs without training with barriers. For a substantial proportion of the NEET with barriers group, a follow-up address could not be established.

[11] Incentive payments are discussed in Chapter 5.

Table 3.2: Pilot survey response

NEET Feasibility Survey

	Pilot strata: Activity group from EMA/YCS interviews							
	Full-time Education	Job with training	NEET- with barriers	NEET- no barriers	JWT- with barriers	JWT- no barriers	Total	
Outcome	n	n	n	n	n	n	n	%
Interview achieved	**6**	**6**	**4**	**11**	**3**	**5**	**35**	39
New address outside area	**2**	**0**	**0**	**5**	**0**	**1**	**8**	9
Refusal	**2**	**1**	**1**	**3**	**2**	**3**	**12**	13
Broken appointment	**0**	**0**	**2**	**1**	**0**	**1**	**4**	4
Ill at home	**0**	**0**	**0**	**1**	**0**	**0**	**1**	1
Named adult deceased	**0**	**1**	**0**	**1**	**0**	**0**	**2**	2
No contact with anyone at address	**0**	**0**	**0**	**0**	**0**	**2**	**2**	2
Contact made at address but not with named adult	**0**	**0**	**0**	**2**	**0**	**0**	**2**	2
Follow-up address not known	**2**	**3**	**5**	**10**	**0**	**2**	**22**	24
Unable to locate address	**0**	**0**	**1**	**0**	**0**	**0**	**1**	1
Residence demolished / derelict	**0**	**0**	**0**	**1**	**0**	**0**	**1**	1
Respondents	***12***	***11***	***13***	***35***	***5***	***14***	***90***	*100*

3.4 The interview

After initially planning to use a paper questionnaire due to the limited size of the pilot, the timetable was extended to produce a computer-assisted questionnaire (CAPI).

The questions used in the pilot were drawn from a range of other studies which have investigated the activities of young people including the Pathways Special Survey of 19/20 year olds, Activity and Learning Agreements Evaluations, Education Maintenance Allowance, Youth Cohort Study and European Social Fund Cohort Survey. The questionnaire was developed in consultation with DWP and DCSF and aimed to last between 45 and 60 minutes, although it was understood at the outset that this would vary depending on the number of jobs and courses the respondent has done since they left school.

A brief outline of each module used in the questionnaire is provided below with comments on how well it worked during the pilot. Recommendations on the questionnaire arising from the pilot are recorded in section 4.3

Introduction and activity diary

The interview began with a brief introduction and the completion of an activity diary sheet.

The activity diary sheet was used by interviewers and respondents to record the different activities that respondents had done since Year 11 (the final year of compulsory schooling). Interviewers were prompted to complete this sheet at the beginning of the interview, and to record any other information that they thought might be helpful to aide the recall of start and end dates of activities. For example, if

a job was held some years before the interview took place, the respondent could be asked at which address he or she was living at the time, or whether a specific child had started school when the job ended. This approach (which has been used to aide recall over long periods of time on the English Longitudinal Survey of Ageing (ELSA), carried out by NatCen) aimed to make it possible to provide detailed and robust information about the timing of each activity.

Following the approach used on other surveys, we attempted to ascertain the correct month for all dates given by asking for exact dates, based on the logic that the respondent is required to think about the date more carefully. If respondents were unable to recall the exact date interviewers were instructed to record the day as the nearest week (so day 7, 14, 21 or 28), and if they could not recall the month to code this based on the season and use the mid-season months. Interviewers reported that in most cases where respondents had participated in a variety of different activities, particularly when these overlapped, the respondent felt that they could not recall exact dates.

Feedback from interviewers indicated that they thought that the activity diary was a useful tool to help recall during the interview. However, there was a consensus that recalling dates was difficult, particularly in such detail and for respondents who had done many different activities since they left Year 11. Respondents sometimes became confused about the start and end dates of their activities and this led to them feeling frustrated during the interview and increased the interview length.

Details of current activities – 'outcomes'

This module of the interview included questions about the respondent's current main activity:

- Details of employment including questions to enable job classification (such as SIC and SOC) and about work-related training.
- Details of study leading to a qualification
- Details about activities and courses not leading to a qualification
- Details of voluntary work
- Details of other activity/activities including job search activities or reasons for not seeking a job.

This module of questions seemed to work well in the pilot, although the location of these questions in the questionnaire should be reviewed prior to a main survey. Whilst it was helpful to talk about the respondent's current activity first to get an idea of their situation, this may not have been ideal for the flow of the interview as the respondent was then asked about demographics and life events, followed by details of previous activities. It may be that asking about demographics and life events as the first module would be more appropriate in the main survey, from which the activity history could follow.

Demographics and life events

This section included questions about the respondent's age, ethnicity, marital status, health and disability status, changes in living arrangements, children and caring responsibilities. There was also a section on benefit receipt currently and since leaving school.

Generally the demographics and life events section worked well. However interviewers did make some suggestions to cut down the sections on living arrangements and benefit receipt. Some quite detailed questions were included about the different residences respondents had lived in since they were in Year 11 at school. The information from this was helpful in some cases to help recall the dates of other activities, but it was suggested that the level of detail needed to be reviewed. It may be that current tenure and who the respondent is currently living with would be sufficient for a main

survey. One approach might be to draw on information about changes in living arrangements as recorded in the diary solely as prompts for recalling other activity dates rather than recording it in detail in the programme.

Interviewers noted that it was often difficult for respondents to recall exact periods of benefit receipt and being asked for such detailed information was time consuming. An alternative approach worth considering would be to link information about benefit receipt from DWP administrative databases for those respondents who give their permission, which at the pilot was 74%.

For a main survey it would be important to consider how this section may need to be extended to include questions about other background factors which have been identified as having an impact on the medium term labour market destinations, or identify sources of this data in the EMA or YCS datasets. This would include information such as parent's education and occupation, attainment and experiences at school and would allow final analysis which could look at respondent outcomes whilst controlling for these background characteristics (e.g. using regression analysis)[12].

Details of activities at 16-18 – starting points

This module of questions included an outline of the respondent's situation during Year 11 and immediately after they left compulsory education. This revisited some subjects from the YCS or EMA interviews. However, this was believed to be helpful in providing some background (particularly about attendance in Year 11) and the opportunity to cover some extra topics around intentions after school and reasons for not pursuing preferred pathways:

- Attitudes to school in year 11, attendance during year 11
- What the respondent wanted to do when they left Year 11 e.g. if they wanted to go to college and did not, why was this?
- Parents' expectations for the respondent when they left Year 11
- Activity straight after leaving Year 11

Feedback from interviewers indicated that this module worked well. Some interviewers suggested including questions about bullying which were mentioned spontaneously by respondents in the context of attendance.

Pathways: activity transitions and trajectories

This section of the interview included details of the activity spells of respondents since they left Year 11. Interviewers were advised to begin with the activity straight after the respondent left Year 11 although it was possible to record activities in any order, working backwards in time from the present, forwards in time or a mixture of the two. Each spell of activity since the respondent left Year 11 was categorised into one of four activity types (listed below), and respondents answered questions covering a standardised set of information for each activity type:

[12] Other contextual information to take into account might include local labour market statistics such as overall unemployment rate and proportion of young people who are NEET. It may be that some people have experienced greater unemployment because of fewer local jobs available rather than just because of their own propensity to work.

1. Employment spells, including changes in job roles within the same employment spell, training at work, details for deriving occupational classification, full-time/part-time work, work-based training, including types of formal and informal work place training
2. Education and training: Full time courses, covering both educational and vocational qualifications started, including courses which were not completed
3. Voluntary work, including the reasons for doing this unpaid work and perceived benefits
4. Spells of any other activity, including caring for the home or family and unemployment. This module asked whether the respondent was actively looking for work and reasons for difficulty in gaining work (including personal and contextual), including any government-sponsored employment or training programmes.

Feedback from interviewers and a review of the data collected provided useful information about possible areas for future development on the activity transitions, or pathways sections of the questionnaire. Overall the structure of this section worked well, in particular the flexibility of recording activities in any order, recommendations for the employment, education and training, and other activity blocks from this section of the questionnaire are below.

Employment

The employment module asked detailed questions about every individual spell of employment/job held in order to provide a standard occupational classification and industry coding, pay details, hours of work and establish the level of both formal and informal training received relating to the job. Three main areas of development were identified for this module.

Firstly, the section on formal and informal training was felt to be very repetitive, particularly for respondents who did not receive any training for many of their jobs. The inclusion of these questions in a main survey would depend on how the data about this is intended to be used in the analysis. It is expected that such detailed information would be helpful to distinguish between jobs with and without training, and the piloted questions could be used as a basis for further development for the main survey.

Secondly, a number of the employment spells recorded related to work alongside another activity, typically higher education amongst pilot respondents. This meant that a considerable amount of interview time included asking detailed questions (including those about training) about employment spells that were typically short-term, low skilled work when for analysis purposes the respondent's main activity at the time would have been education. For a main survey it would be important to focus respondents (and interviewers) on the main activity at each given time, possibly recording brief details about other work while studying in order to reduce the questionnaire length and make this section more relevant.

Finally, it was noted that in a number of cases the respondent was not able to recall the amount of pay received but did know that this was the minimum wage. Therefore, any questions asking for pay levels should include an option for the minimum wage which could be linked to the minimum wage at the time of employment for use in analysis.

Education and training

The pilot questionnaire recorded detailed information about any spells of education or training leading to a qualification, including qualification types and levels achieved, funding arrangements and the place of study. Collecting all of this information was time consuming, particularly when a respondent gave information about qualifications studied concurrently separately which created a lot of duplication. Therefore, a main survey may want to limit this section to establish the respondent's highest qualification and recording more detailed information about this, then just recording key details for other spells of study such as start and end dates, qualification type/level and whether the qualification was completed.

Other activities

If the respondent was not in employment, on a course or doing voluntary work at any time they were directed to the 'other activities' module which aimed to collect information about spells of unemployment and economic inactivity. Information about claimant unemployment would be available from DWP administrative data if permission for data linkage was provided. However, as there is a significant difference between ILO and claimant unemployment figures it would be important to collect information on what non-claimants who were unemployed have been doing. In addition, the other data collected from this module would also be helpful in analysis of a main survey, such as having details about participation in courses not leading to a formal qualification and other learning or activities which would be viewed as positive steps towards employment. This, together with reasons for inactivity such as caring for children or family members, ill health or disability would be valuable. Data collection during the pilot, whilst covering some spells of these other activities, in comparison with the completed activity histories suggested that not all the spells of other activity were recorded separately. This may partly have been due to these activities typically being collected near the latter stages of the interview, at a time when respondents were fatigued. Therefore, a main survey should consider how best to ensure spells of activity not classed as employment, education or training are recorded. This may be by reviewing the preferred order of collecting activity data to prioritise other activities over some education spells and emphasising the importance of recording what the respondent says they were doing during NEET spells to interviewers.

Other areas for consideration for a main survey questionnaire would be guidance on recording promotions as separate spells of work (and allowing interviewers to record this easily without repeating questions) and resolving issues around the level of detail needed for start and end dates of activities.

Interviewers reported that respondents used a combination of methods to recall their activities; with some preferring to start with what they did after they left year 11 and move forwards in time, while others started by moving backwards from their current activity. Therefore, it is important that any CAPI program provides a lot of flexibility to record activities in whichever way the respondent finds easiest, including recording all of a specific type of activity before moving on to the next.

Advice

This module focused on the sources of advice and support respondents had used to make decisions about what activities to do or not since they left Year 11, and how useful these sources were.

Activities and achievements (looking back)

This module of questions asked about barriers to work and study since leaving Year 11, including:

- Whether respondents felt they had the qualifications they needed for the jobs or courses that they wanted to pursue before and after making the transition to paid work,
- Attitude to availability of jobs and courses in their local area,
- Perceived barriers to work and study that may have affected the individual's chances and decisions, including constraints linked to the individual's situation.

Activities and ambitions (looking forward)

This module asked about plans for the future and expected barriers to work and study, including:

- What they thought they would be doing in a year's time
- What considerations would influence the type of work or study they do in the future
- Whether they had plans to engage in education or training in the near future
- Whether they felt that they had enough information and advice make choices and plan for the future.

The three modules on advice and support, achievements and ambitions were quite short and straightforward with respondents finding these questions easy to respond to.

Final questions

This final module included:

- A question aiming to establish if the respondent was in one of the groups of young people identified by the government as at risk of social exclusion.
- Asking for permission to link data from the interview with DWP administrative records.
- Assurances about confidentiality.
- Thanking the respondent for giving up their time to help with the study.

A showcard was used to try and establish whether respondents considered themselves to be in one of the at risk groups of interest to the government. Respondents were just asked to indicate whether they were in one or more of the following groups on the card by saying yes or no:

- You have spent time in a residential care home
- You have a criminal record
- You have drug or alcohol problems
- You have a learning difficulty, disability or are in contact with mental health services

As this is a sensitive question, prior to the pilot it was unclear how respondents would respond and this was placed towards the end of the interview. Overall, 15 per cent (5 respondents) indicated that they were from one or more of these groups with no respondents refusing to answer the question or objecting to giving this information.

Respondents were asked for their permission to link data from their interview to administrative records held by the DWP and were given a data linkage sheet (see Appendix B) with details of this which they could keep. Interviewers reported that respondents seemed to understand what was being asked and were clear about whether they were happy to provide this permission or not. This question was asked at the end of the interview, and 74 per cent (24) of the respondents indicated that they would give

permission for their data to be linked. Declining to give permission for data linkage was normally due to a general reluctance about sharing personal information in this way with respondents not providing specific reasons for this.

Questionnaire length

The median interview length (including the completion of the activity diary) was 64 minutes, ranging from 32 minutes to 109 minutes, which was longer than intended. The longest interviews drew a negative reaction from respondents, expressing frustration or fatigue. Questionnaire development for a main survey should focus on reducing the questionnaire length and limiting the collection of very detailed data about dates to what is expected to be most useful for the analysis and reporting stage of the project.

3.5 Summary

The pilot survey suggests that it would be possible to carry out a survey to investigate the medium term impact of being NEET between 16 and 18 as compared with other activity groups. The EMA and YCS data sets proved adequate for providing the necessary background information about young people to categorise them into activity groups and for tracing a high proportion some years later.

Considering the constraints associated with pilot fieldwork (namely the limited fieldwork period and small pool of interviewers), the response rate was good. There was evidence to suggest that response would be exceeded in the context of a main survey.

On the whole, the face-to-face computer assisted programme worked well and detailed activity information was collected. The interview topics have been described in detail in this chapter. In chapter 6, recommendations for the method and interview programme for a full survey are made on the basis of the pilot.

4 Sampling considerations for a main survey

This chapter discusses a range of considerations associated with sampling for a main survey in light of the investigation of the sample sources and the experience of the pilot. First we consider how the issues raised by the pilot could be used to inform the sampling for a main survey. Secondly, we investigate the attrition in the EMA and YCS data from the first to last wave of each survey to identify whether specific groups of young people were more likely to drop out of the survey and therefore not be represented in the final wave. Thirdly, we consider the ethical issues relating to recontacting respondents from the EMA and YCS surveys. Finally, a definition of the NEET group is considered, including how this group might be separated for sampling purposes and the potential pool of young people these definitions provide to follow up for a main survey.

4.1 Revisions to the pilot sampling methodology

Defining activity groups

In the pilot survey, the four key activity groups used for sampling were: NEET, jobs without training (JWT), jobs with training and full time education (FTE). There are a number of other activities that did not fall into these groups and were excluded from the pilot. These included government supported training (GST), part time education; other education or training and 'other' activities[13]. In preparing the sample groups for a main survey, consideration would need to be given as to how these activities are allocated or grouped.

For the purposes of eliciting the EMA and YCS sample frames for the pilot study, only the final waves of address data were available for each of the selected cohorts. In the case of the EMA pilots and YCS 11 and 12 this data was based on respondents who were aged 19/20 at the time their data was recorded. For a main survey, it may be useful to have address information from other waves as well.

As explained in section 3.1, respondents were categorised into the six pilot strata groups in two different ways. For the EMA respondents (the majority of the sample) NEET status was defined as NEET in any of the survey waves, so aged 16-20. For the YCS respondents, NEET was defined as at the final wave of cohort 10 and 11, therefore NEET status was defined when the respondents were aged 18/19 and 19/20 respectively. As the specification for the study was to understand the trajectories of those who were NEET or in other activities when aged 16-18, the discussion about defining NEET further on in this chapter does not take into account the responses of the final wave participants.

4.2 Attrition of the sample

The pilot sample frame was derived using the final wave only of each of the selected cohorts, primarily because the address data provided at this point was the most up-to-date and therefore we presumed most likely to yield potential responses to a follow-up survey. However, for a main survey, a decision would need to be taken as to whether to draw only from the final wave or whether to also include respondents to earlier waves. Two key issues are firstly whether the profile of respondents changes

[13] Whilst these activities were excluded in the pilot, a main stage survey would include these groups.

across the survey series and secondly, whether it is ethically sound to contact respondents at each wave. This section presents the analysis of attrition between waves and the following section discusses the ethical considerations around recontacting respondents.

Tables 4.1 and 4.2 show that at each wave, the proportion of respondents who report that they are in full time education declines. The proportion of respondents who are NEET or in jobs with or without training varies. However, Tables 4.3 and 4.4 demonstrate that the respondents who reported their main activity as FTE at wave 1 were more likely to remain in the EMA and YCS over time and those who reported that they were NEET or JWT were less likely to remain in the survey.
For example, while those in full-time education at wave 1 of EMA comprised 74 per cent of the achieved sample, by wave 4 this group accounted for 83 per cent of the sample (for YCS the corresponding figures were 79 per cent and 88 per cent). This is in contrast to the NEET group who made up 8 per cent of the wave 1 EMA sample, reducing to 4 per cent in wave 4. This suggests that only using wave 4 as the sample frame for a main survey would reduce the potential numbers of NEET young people to follow up and may also lead to more biased estimates.

Table 4.1 Main activity of respondent by wave: EMA[14]

EMA 2

	Wave of EMA			
	Main activity Wave 1	Main activity Wave 2	Main activity Wave 3	Main activity Wave 4
Respondent Main activity	Count (%)	Count (%)	Count (%)	Count (%)
NEET	922 (8%)	492 (6%)	633 (11%)	334 (7%)
Job with training	955 (9%)	278 (4%)	1148 (20%)	674 (15%)
Job without training	823 (8%)	1049 (14%)	845 (14%)	456 (10%)
In full-time education	8022 (74%)	5273 (68%)	2616 (45%)	2069 (46%)
In part-time education	62 (1%)	23 (0%)	17 (0%)	20 (0%)
Full or part-time work not specified whether with training-same as previous wave	0 (0%)	541 (7%)	598 (10%)	982 (22%)
Other	86 (1%)	52 (1%)	16 (0%)	12 (0%)
Bases	*10,870*	*7,708*	*5,873*	*4,547*

[14] Note that the tables in this section are based on all respondents, including those who declined to be recontacted.

Table 4.2 Main activity of respondent by wave: YCS

YCS 12

	Wave of YCS			
	Main activity Wave 1	Main activity Wave 2	Main activity Wave 3	Main activity Wave 4
	Count (%)	Count (%)	Count (%)	Count (%)
Respondent Main activity				
NEET	671 (5%)	404 (5%)	632 (11%)	246 (6%)
Job with training	255 (2%)	356 (4%)	599 (10%)	433 (11%)
Job without training	678 (5%)	527 (6%)	963 (16%)	732 (18%)
In full-time education	10458 (83%)	7262 (82%)	3822 (65%)	2510 (63%)
GST	880 (7%)	571 (6%)	348 (6%)	170 (4%)
Other Education / Training	259 (2%)	184 (2%)	168 (3%)	123 (3%)
Bases	*12,530*	*8,900*	*5,900*	*3,968*

Table 4.3 Attrition by main activity of respondent at wave 1: EMA

EMA 2

	Wave of EMA			
	Wave 1 respondents - main activity at wave 1	Wave 2 respondents - main activity at wave 1	Wave 3 respondents - main activity at wave 1	Wave 4 respondents - main activity at wave 1
	Count (%)	Count (%)	Count (%)	Count (%)
Respondent Main activity				
NEET	922 (8%)	443 (6%)	260 (4%)	168 (4%)
Job with training	955 (9%)	626 (8%)	453 (8%)	323 (7%)
Job without training	823 (8%)	481 (6%)	326 (6%)	229 (5%)
In full-time education	8022 (74%)	6096 (79%)	4788 (82%)	3793 (83%)
In part-time education	62 (1%)	28 (0%)	21 (0%)	15 (0%)
Other	86 (1%)	34 (0%)	25 (0%)	19 (0%)
Bases	*10,870*	*7,708*	*5,873*	*4,547*

Table 4.4 Attrition by main activity at wave 1: YCS

YCS 12

	Wave of YCS			
	Wave 1 respondents - main activity at wave 1	Wave 2 respondents - main activity at wave 1	Wave 3 respondents - main activity at wave 1	Wave 4 respondents - main activity at wave 1
Respondent Main activity	Count (%)	Count (%)	Count (%)	Count (%)
NEET	671 (5%)	322 (4%)	177 (3%)	100 (2%)
Job with training	255 (2%)	142 (2%)	91 (1%)	56 (1%)
Job without training	678 (5%)	374 (4%)	222 (3%)	116 (3%)
In full-time education	10458 (83%)	7795 (87%)	5605 (88%)	3688 (90%)
GST	880 (7%)	508 (6%)	329 (5%)	192 (5%)
Other Education / Training	259 (2%)	163 (2%)	108 (2%)	62 (2%)
Bases	*12,530*	*8,982*	*6,355*	*4,114*

A number of characteristics are associated with young people being NEET. Tables 4.5 to 4.7 show attrition over the waves according to lone parent, truancy and disability status at wave one. Table 4.5 shows that the proportion of lone parents does not change substantially over time, however, there are very few lone parents to start with and it is evident that the number of lone parents decreases over the time series.

Table 4.5 Attrition by lone parent status at Wave 1

YCS 12

	Wave of YCS			
	Wave 1	Wave 2	Wave 3	Wave 4
Lone parent status	Count (%)	Count (%)	Count (%)	Count (%)
Not answered	151 (1.1%)	71 (0.8%)	39 (0.6%)	21 (0.5%)
Yes	58 (0.4%)	37 (0.4%)	20 (0.3%)	9 (0.2%)
No	12992 (98.4%)	9196 (98.8%)	6473 (99.1%)	4184 (99.3%)
Bases	*13,201*	*9,304*	*6,532*	*4,214*

Respondents who reported truancy at wave 1 were more likely to drop out of subsequent waves (Table 4.6).

Table 4.6 Attrition by truant status at Wave 1

YCS 12

	Wave of YCS			
	Wave 1	Wave 2	Wave 3	Wave 4
Truant status	Count (%)	Count (%)	Count (%)	Count (%)
Not answered	157 (1%)	79 (1%)	51 (1%)	25 (1%)
Persistent truancy	321 (2%)	169 (2%)	95 (1%)	51 (1%)
Occasional truancy	3459 (26%)	2206 (24%)	1466 (22%)	882 (21%)
No truancy	9264 (70%)	6850 (74%)	4920 (75%)	3256 (77%)
Bases	*13,201*	*9,304*	*6,532*	*4,214*

Table 4.7 shows that there was little change in the proportion of participants with a long-standing illness or disability over the survey series. These conclusions were similar across both surveys.

Table 4.7 Attrition by disability status at wave 1

EMA 2

	Wave of EMA			
	Wave 1	Wave 2	Wave 3	Wave 4
Disability status	Count (%)	Count (%)	Count (%)	Count (%)
Yes	1721 (16%)	1291 (17%)	999 (17%)	768 (17%)
No	9100 (84%)	6416 (83%)	4873 (83%)	3778 (83%)
Bases	*10,821*	*7,707*	*5,872*	*4,546*

Tables 4.8 to 4.11 depict the living arrangements of respondents to the EMA and YCS survey surveys, using the variables as stated at wave one. The tables show that respondents who lived in rented accommodation, who didn't live with their parents and whose parents were unemployed were more likely to drop out over the time series. Table 4.11 demonstrates that in addition, respondents who received free school meals whilst in FTE were more likely to have dropped out of the survey than those who did not. Using the final survey respondents (aged between 18 and 20 depending on the survey) as the sample frame potentially restricts the ability to track respondents who have these characteristics, some of whom may be more susceptible to being NEET.

Table 4.8 Attrition by tenure of respondent at Wave 1

YCS 12

	Wave of YCS			
	Wave 1	Wave 2	Wave 3	Wave 4
Tenure	Count (%)	Count (%)	Count (%)	Count (%)
Not answered	374 (3%)	204 (2%)	129 (2%)	76 (2%)
Owned	10719 (81%)	7829 (84%)	5620 (86%)	3672 (87%)
Council rented	1300 (10%)	783 (8%)	482 (7%)	291 (7%)
Other rented	782 (6%)	472 (5%)	289 (4%)	168 (4%)
Other	26 (0%)	16 (0%)	12 (0%)	7 (0%)
Bases	*13,201*	*9,304*	*6,532*	*4,214*

Table 4.9 Attrition by whether or not respondent lives with their parents

EMA 2

	Wave of EMA			
	Wave 1	Wave 2	Wave 3	Wave 4
Live with parents	Count (%)	Count (%)	Count (%)	Count (%)
Yes	10302 (95%)	7617 (99%)	5819 (99%)	4508 (99%)
No	474 (4%)	54 (1%)	36 (1%)	29 (1%)
No parent lives with the young person	97 (1%)	37 (0%)	18 (0%)	10 (0%)
Bases	*10,870*	*7,708*	*5,873*	*4,547*

Table 4.10 Attrition by parental employment at Wave 1

YCS 12

	Wave of YCS			
	Wave 1	Wave 2	Wave 3	Wave 4
Parental employment	Count (%)	Count (%)	Count (%)	Count (%)
Neither parent employed	1538 (12%)	937 (10%)	592 (9%)	348 (8%)
Mother only	1459 (11%)	990 (11%)	638 (10%)	390 (9%)
Father only	2114 (16%)	1499 (16%)	1072 (16%)	691 (16%)
Both parents employed	8090 (61%)	5878 (63%)	4230 (65%)	2785 (66%)
Bases	*13,201*	*9,304*	*6,532*	*4,214*

Table 4.11 Attrition by proportion of respondents who received free school meals at any point

EMA 2

	Wave of EMA			
	Wave 1	Wave 2	Wave 3	Wave 4
Received free school meals	Count (%)	Count (%)	Count (%)	Count (%)
Yes	2129 (20%)	1221 (16%)	806 (14%)	574 (13%)
No	8687 (80%)	6484 (84%)	5067 (86%)	3973 (87%)
Bases	*10,816*	*7,705*	*5,873*	*4,547*

Some of the characteristics that we know to increase the risk of a respondent being NEET were associated with ongoing participation in the YCS and EMA. The evidence is clearer when looking at deprivation proxies that may be synonymous with an increased risk of being NEET. Over the survey series, respondents who were more likely to have received free school meals, live in council rented accommodation, have both parents unemployed or be a lone parent were also more likely to have dropped out of the survey. Ethnicity, gender and region presented very little differences over time in both the YCS and EMA pilot survey data.

The analysis above suggests that ideally, respondents from all eligible waves of the YCS and EMA waves would be included in the sample frame for a main survey in order to maximise the numbers within each of the key activity groups. There are, however, some important ethical considerations which are addressed in the following section.

4.3 Ethical considerations

The analysis of attrition demonstrated the potential benefits of including respondents from all waves in the sample frame. However, not all respondents gave their permission to be recontacted and even where they did so, the question wording did not directly encompass different studies taking place in the future. The ethical basis for recontacting different groups of respondents is outlined below.

Final wave respondents

We would suggest that the recontact question asked of final wave respondents in EMA and YCS provides adequate basis for approaching them for a main survey. In EMA, respondents were asked:

'We would like to contact you again in the future, probably in about a year's time to find out what you are doing then. Would you be willing to have another interview?
Again, your replies will be treated in strict confidence'.

The recontact question for YCS respondents was:

'This is the last of the planned Pathways studies, but just in case we ever needed to contact you again please complete your details below'.

Although the questions do not specifically ask whether the respondent would be happy to be recontacted for another study, this could be made explicit in the introductory letter along with the opportunity to opt-out. For example, the letter could include text along the lines of,

"… although we recorded that you had agreed to be contacted again as part of the ['Pathways'/' EMA '] study, we are giving you the opportunity to tell us in advance if you would prefer **not** to be re-contacted for the new follow-up research study …"

This text could include the point that interviewers will be able to answer any queries that those in the sample may wish to ask before deciding whether to be interviewed.

Respondents who refused to be recontacted

Respondents who responded negatively to the recontact question at any wave would not be included in the sample frame. The numbers involved are not known at this stage as we did not have access to

this information during the feasibility study. However, on the basis of the experience across other surveys, it is likely to involve a fairly small number.

Respondents who dropped out before the final wave

The ethical situation in relation to the respondents who agreed to be recontacted but who dropped out of the surveys before the final waves is less clear. The argument against including them in the sample frame includes the fact that they made a choice not to take part and also that the recontact question was asked in relation to a particular survey at a particular point in time. However, it may be considered ethical to recontact them given the passage of time and the opportunities that would be given for opting out in the new survey. A view would need to be taken by the Department and/or organisation carrying out the survey. It would be preferable but not essential to include these individuals[15], and therefore the feasibility of the study would not rest on making a decision at this point. A similar view would need to be taken in relation to individuals who had not been asked a recontact question.

4.4 Defining NEET young people

The EMA and YCS data provide a range of variables about young peoples' activities at the time of interview, and since they left school, in addition to some background and attitudinal measures. Therefore, there are a number of different ways this data could be used to classify young people in specific activity groups to be used for sampling for a main survey. This richness also presents a number of challenges for creating a consistent definition.

As the EMA and YCS data both include details of young people's main activity at the time of interview, and some overlapping information about the main activity in certain months of the year this information has been used to create two definitions of the NEET group to examine their potential yield of young people for a follow up survey.

Tables 4.12 and 4.13 display the total number of respondents categorised by the main activity groupings discussed above. The categories are all defined by taking into account main activity status at the annual interview across waves. Therefore, if a respondent was classified as in FTE at any of the four EMA waves, and they were not NEET at any point they are included in the 'FTE or PTE' category.

Table 4.12 EMA main activity across waves

Base: 10,668 *EMA 2*

	EMA cohort 2 Waves 1-3	
Activity status	Count	%
FTE or PTE	7,504	70%
Job with training	904	8%
Job without training	600	6%
NEET as a main activity in any wave	1,660	16%
Bases	*10,668*	*100%*

[15] Note that in addition to the young people from EMA2 and YCS12, there are other cohorts to draw on for the sample frame.

Table 4.13 YCS main activity across Waves

Base: 13,201

YCS 12

Activity status	YCS cohort 12 Waves 1-3 Count	%
FTE or GST	9,272	70%
Job with training	1,107	8%
Job without training	1,362	10%
NEET as a main activity in any Wave	1,460	11%
Bases	*13,201*	*100%*

As mentioned in section 3.1, a number of main activity categories (PTE, GST, other and other education or training) were excluded from the pilot sample definition. They have been included in the tables above as indicated. 'Other' has been included in the NEET category in table 4.12 and 'Other education or training' has been included in the job with training category of table 4.13. There are some additional variables in the YCS and EMA datasets that could be interrogated to try and classify these respondents more adequately.

YCS 12 did not include any questions in the survey asking respondents to recall their monthly activity, therefore tables 4.14 to 4.15 display the EMA definition taking account of monthly NEET status only. YCS 10 and 11 did include monthly variables and therefore it is possible to apply the same methodology to these data.

The EMA wave one survey asked respondents to recall their main activity status in the nine months since they last attended school in Year 11. The subsequent waves asked respondents to recall their activities in February, May and September. Being NEET is often a transitional state, and it is possible that respondents recorded their status as NEET at one or more months during the year but did not record their main activity for the year as NEET. In order to capture the breadth of respondents who may have been NEET between the ages of 16-18 table 4.14 displays the number of respondents recorded as NEET at any wave or month. The categories other than NEET have been defined in the same way as table 4.12.

Due to seasonal variations in young people being NEET[16] it is important to distinguish between those who were recorded as NEET for a short period (such as during school holidays for example) and those who had a more sustained period of being NEET. Table 4.15 shows that 890 (37%) of those classified as NEET in table 4.14 recorded their status as NEET in more than one month or in more than one wave.

[16] NEET Statistics Quarterly Brief February 2010 (available at http://www.dcsf.gov.uk/rsgateway/DB/STR/d000913/index.shtml). Seasonal variations were also observed in authors' own analysis of YCS and EMA data.

Table 4.14 EMA main activity across waves

Base: 10,668

EMA 2

Activity status	EMA cohort 2 Waves 1-3 Count	%
FTE or PTE	6,872	64%
Job with training	828	8%
Job without training	548	5%
NEET as a main activity in any month or wave	2,420	23%
Bases	*10,668*	*100%*

Table 4.15 EMA main activity across waves

Base: 10,668

EMA 2

Activity status	EMA cohort 2 Waves 1-3 Count	%
FTE or PTE	6,872	64%
Job with training	828	8%
Job without training	548	5%
NEET as a main activity in 2/3 months or waves	*890*	*8%*
NEET as a main activity in 1 month or wave	*1,530*	*14%*
Bases	*10,668*	*100%*

A further consideration is whether to base the NEET definition on the activity status at the time of the interview or taking into account activity recorded on a monthly basis. Table 4.16 shows how the two definitions compare. A key point is that of the 9,031 respondents who did not classify themselves as NEET at the time of interview in any wave, 763 (8%) recorded their monthly activity as NEET at least once. However, since the majority (74%) only indicated that their monthly activity was NEET on one occasion, a decision would need to be taken as to whether to include them in a NEET definition.

Table 4.16 EMA Main activity as NEET annual by NEET monthly

Base: 10,668

EMA 2

NEET category - Monthly	NEET category - at main interview: Not NEET at any point Count (%)	NEET as a main activity in 1 wave only Count (%)	NEET as a main activity in any 2 waves Count (%)	NEET as a main activity in any 3 waves Count (%)	Total Count (%)
Not NEET at any point	8268 (92%)	709 (53%)	41 (17%)	0 (0%)	9018 (85%)
NEET as a monthly activity in 1 month only	571 (6%)	250 (19%)	42 (17%)	1 (1%)	864 (8%)
NEET as a monthly activity in any 2 months	112 (1%)	153 (12%)	48 (20%)	7 (10%)	320 (3%)
NEET as a monthly activity in any 3 months	80 (1%)	214 (16%)	113 (46%)	59 (88%)	466 (4%)
Bases	*9,031*	*1,326*	*244*	*67*	*10,668*

This analysis suggests that to increase the sample size, particularly for the NEET main activity category, taking account of the monthly activity status would be beneficial. However, it is important to be mindful of the different types of NEET and to consider which groups it would be useful to include in any definition of NEET used in a main survey.

4.5 Defining NEET sub-groups

It is acknowledged that NEET young people are not a homogenous group. However, segmenting the NEET group further using two different data sources is complex and has presented a number of challenges. A number of sub-group classifications have been used in previous research, these include:

- The model used by the DCSF using the Customer Case Information System (CCIS) which identified four sub-groups of 16-18 NEET young people: those doing some activity that was not counted as EET, such as voluntary work (13%), those who had an identifiable barrier to participation such as childcare responsibilities, illness or a disability (17%), those whose activity is known but who are not in the above groups (60%) and others whose activity was not known (10%).
- Using the length of time NEET in the original survey period. Young people who have been continually NEET for the most prolonged periods are likely to be the most disadvantaged and have complex needs (Payne, 2000).
- A system based on the mix of activities between the ages of 16-18, for example whether there was any spell of work and any spells of education or training.
- A classification based on personal characteristics of NEET young people, including gender, being a parent and problems to do with health or a disability.
- Segmenting based on responses to the Youth Cohort Survey, including attitudes to educational experience, post 16 experience, future activity and perceived barriers (Spielhofer et al 2009). This analysis identified three NEET sub-groups: the 'open to learning'; the 'sustained'; and the 'undecided', using Latent Class Analysis[17].

We have looked at a number of ways that NEET sub-groups could be defined with the data available. However, there are a number of challenges and constraints associated with each method and consistently applying this to participants from different surveys (and different cohorts within surveys). Key issues whilst considering different definitions include

- Can this be applied to EMA and YCS data?
- Can young people who only participated in wave 1 of the previous research be adequately/reliably categorised?
- Are the numbers of young people in each group sufficient for sampling?

Ideally, a system of sub-grouping young people based on the duration that they were NEET when aged 16-18 would help ensure that young people who were NEET for a longer time were sufficiently

[17] The young people in these groups were described by Spielhofer et al as follows.
Open to learning: more likely to re-engage in learning in the short-term, have a positive attitude to school and achieved higher attainment levels.
Sustained NEET: more likely to remain NEET, negative experiences of school including higher rates of truancy and exclusion, and lower attainment.
Undecided: dissatisfied with the opportunities available and access to activities that they want to do. Similar experiences of school and attainment levels to the open to learning group.

represented in a main survey, alongside those who experienced short spells of being NEET and distinctly from those who may have had just one spell of being NEET. However, the data available from EMA and YCS, whilst providing some data about main activity does not provide the start and end dates of activities which makes it difficult to determine the length of time a young person was NEET.

Defining NEET sub-groups using activity duration

When thinking about duration of NEET and the variables available we considered a number of options. Firstly, we investigated using the duration a young person was NEET with reference to the number of waves that a respondent had recorded their activity status as NEET. This proved to be too stringent a definition as there were few respondents in the EMA or YCS who were NEET in all waves. The definition was then expanded to take account of respondents who recorded their status as NEET in any 1, 2, or 3 waves and this proved to be a more workable solution with respect to the number of respondents who were classified in each group. The EMA results are displayed in table 4.16 and the YCS 10[18] results in table 4.17. The number of respondents who were NEET in 3 or more waves is still quite low (although it should be borne in mind that the data presented are for two cohorts only, there are a further three cohorts that could be utilised). Alternatively main activity in 2 or 3 waves could be combined to create a category of respondents who were NEET on 2 or more occasions.

Table 4.17 YCS main activity - annual NEET

Base: 12,906

YCS 10

	YCS cohort 10	
	Waves 1-3	
NEET duration	Count	%
Not NEET at any point	11,883	92%
NEET as a main activity in 1 Wave only	844	7%
NEET as a main activity in any 2 Waves	135	1%
NEET as a main activity in 3 Waves	44	0%
Bases	*12,906*	*100%*

The same methodology was then applied to the monthly data (not accounting for annual activity status), tables 4.18 and 4.19 show the results of this analysis. The results are consistent with those above, defining NEET in terms of monthly duration rather than annual activity yields a larger number of respondents by category.

[18] YCS 10 data is displayed instead of YCS 12 as table 3.9 is displayed as a comparison between the monthly and annual definition, as YCS 12 does not include monthly variables YCS 10 data was used for both tables.

Table 4.18 YCS main activity - monthly NEET

Base: 12,906 *YCS 10*

NEET duration	YCS cohort 10 Waves 1-3 Count	%
Not NEET at any point	9,709	75%
NEET as a monthly activity in 1 month only	2,198	17%
NEET as a monthly activity in any 2 months	540	4%
NEET as a monthly activity in any 3 months	459	4%
Bases	*12,906*	*100%*

Table 4.19 EMA main activity - monthly NEET

Base: 10,668 *EMA 2*

NEET duration	EMA cohort 2 Waves 1-3 Count	%
Not NEET at any point	9,018	85%
NEET as a monthly activity in 1 month only	864	8%
NEET as a monthly activity in any 2 months	320	3%
NEET as a monthly activity in any 3 months	466	4%
Bases	*10,668*	*100%*

In terms of the seasonal aspect of NEET respondents, a further consideration is the extent to which respondents reported consecutive months/waves of NEET as opposed to 1, 2 or 3 non-consecutive periods of NEET within the specified time frame. Based on EMA respondents who reported consecutive or non-consecutive waves of NEET activity, the frequencies were low for non-consecutive respondents therefore this definition was not pursued any further. If the same methodology was to be applied to the monthly data then this may yield better frequencies.

Defining NEET sub-groups using known barriers to EET

The pilot survey investigated how easy respondents who were deemed to have 'identifiable barriers' to being EET were to re-contact. This was not necessarily a strata that was intended to be continued into the main study, however, respondents who were disabled, lone parents or persistently truant whilst at school were cross tabulated with the NEET duration categories defined above. The number of respondents with barriers in one cohort limits the analysis (particularly in terms of EMA), however tables 4.20 and 4.21 lend some weight to the argument that by accounting for duration of NEET it is possible to capture the differentials between those who do and do not have perceived barriers to employment, education or training.

Table 4.20 EMA Main activity as NEET annual by barrier to employment, education or training

Base: 10,668

EMA 2

	NEET category - at main interview				
	Not NEET at any point	NEET as a main activity in 1 wave only	NEET as a main activity in any 2 waves	NEET as a main activity in any 3 waves	Total
Barrier to employment, education or training	Count (%)	Count (%)	Count (%)	Count (%)	Count (%)
No barrier to employment, education or training	8893 (98%)	1289 (97%)	232 (95%)	61 (91%)	10475 (98%)
Barrier to employment, education or training	138 (2%)	37 (3%)	12 (5%)	6 (9%)	193 (2%)
Bases	*9,031*	*1,326*	*244*	*67*	*10,668*

Table 4.21 YCS Main activity as NEET annual by barrier to employment, education or training

Base: 13,201

YCS 12

	NEET category - at main interview			
	Not NEET at any point	NEET as a main activity in 1 Wave only	NEET as a main activity in any 2 Waves	Total
Barrier to employment, education or training	Count (%)	Count (%)	Count (%)	Count (%)
No barrier to employment, education or training	11158 (95%)	1085 (87%)	162 (76%)	12405 (94%)
Barrier to employment, education or training	583 (5%)	163 (13%)	50 (24%)	796 (6%)
Bases	*11,741*	*1,248*	*212*	*13,201*

A more complex NEET sub-group definition

The NEET definitions that have been discussed so far have only considered splitting the NEET young people into sub-groups of NEET, based on the number of months (or interviews) at which they have been NEET. The segmentation analysis carried out by Spielhofer et al (2009) suggests that the divisions of NEET are more complex than the interpretation taken so far. Therefore, a possible extension would be to consider what respondents were doing when they were not NEET, with a view to creating a matrix of combined activities in order to represent all possible NEET sub-groups.

Tables 4.22 to 4.24 show the main activities of respondents across waves for respondents who were NEET in any one wave, NEET in two or more waves or not NEET in any waves.

Table 4.22 Main activity of respondent by wave - respondents who were NEET in any 1 wave

EMA 2

	Wave of EMA			
	Wave 1	Wave 2	Wave 3	Wave 4
Respondent Main activity	Count (%)	Count (%)	Count (%)	Count (%)
NEET	662 (50%)	223 (25%)	441 (66%)	107 (23%)
Job with training	79 (6%)	37 (4%)	82 (12%)	120 (26%)
Job without training	80 (6%)	201 (23%)	67 (10%)	85 (18%)
In full-time education	486 (37%)	357 (41%)	42 (6%)	86 (18%)
In part-time education	9 (1%)	5 (1%)	2 (0%)	7 (2%)
Full or part-time work not specified whether with training-same as previous wave	0 (0%)	38 (4%)	29 (4%)	58 (12%)
Other	10 (1%)	18 (2%)	1 (0%)	2 (0%)
Bases	*1,326*	*879*	*664*	*465*

Respondents who were NEET in only one wave were predominantly in full time education in wave two (41%; Table 4.22) and most likely to be NEET in wave three (age 18/19). This group were then divided more equally between NEET, FTE and jobs with or without training at wave four (age 19/20).

Table 4.23 Main activity of respondent by wave - respondents who were NEET in any 2 or more waves

EMA 2

	Wave of EMA			
	Wave 1	Wave 2	Wave 3	Wave 4
Respondent Main activity	Count (%)	Count (%)	Count (%)	Count (%)
NEET	228 (73%)	269 (86%)	192 (84%)	72 (51%)
Job with training	9 (3%)	3 (1%)	18 (8%)	29 (20%)
Job without training	14 (5%)	24 (8%)	15 (7%)	23 (16%)
In full-time education	55 (18%)	8 (3%)	2 (1%)	2 (1%)
In part-time education	2 (1%)	1 (0%)	1 (0%)	4 (3%)
Full or part-time work not specified whether with training-same as previous wave	0 (0%)	0 (0%)	0 (0%)	11 (8%)
Other	3 (1%)	6 (2%)	1 (0%)	1 (1%)
Bases	*311*	*311*	*229*	*142*

Respondents who were NEET in more than one wave were very unlikely to be in FTE post 16 with the majority of this group remaining NEET until wave three (age 18/19). Fifty-one per cent were still NEET at wave four with more respondents moving into jobs with or without training (Table 4.23).

Table 4.24 Min activity of respondent by wave - respondents who were not NEET in any waves

EMA 2

Respondent Main activity	Wave of EMA: Wave 1 Count (%)	Wave 2 Count (%)	Wave 3 Count (%)	Wave 4 Count (%)
NEET	0 (0%)	0 (0%)	0 (0%)	155 (4%)
Job with training	852 (9%)	238 (4%)	1048 (21%)	525 (13%)
Job without training	716 (8%)	824 (13%)	763 (15%)	348 (9%)
In full-time education	7370 (82%)	4908 (75%)	2572 (52%)	1981 (50%)
In part-time education	49 (1%)	17 (0%)	14 (0%)	9 (0%)
Full or part-time work not specified whether with training-same as previous wave	0 (0%)	503 (8%)	569 (11%)	913 (23%)
Other	44 (0%)	28 (0%)	14 (0%)	9 (0%)
Bases	*9,031*	*6,518*	*4,980*	*3,940*

Finally, respondents who were not NEET in any of the waves were predominately in FTE throughout the survey period (Table 4.24). This analysis suggests that using duration to define the NEET sub-groups may be sufficient to pick up the complexities in any main activity patterns.

The segmentation analysis by Spielhofer et al (2009) using the YCS drew heavily on the attitudinal data from the survey and provided a useful distinction between young people who were NEET, particularly the 'open to learning' group. The attitudinal responses available from the first waves of EMA and YCS could be used to attempt to identify similar groups to identify young people who were 'open to learning' or had characteristics associated with barriers to engagement in wave 1. However, doing so would be more subjective, and would be based on less rich data than originally used in previous research to define these groups. Therefore, it is suggested that any methods of such detailed segmentation are reserved for analysis of a main survey, rather than as a method of identifying strata for sampling.

It is of course possible to not divide the NEET group any further for sampling, and just use the main activity groups as strata. However, when considering a sampling strategy it is important to consider the breakdowns by respondent characteristics which would be desired from the main survey. It may be that it is important to show outcomes for specific groups such as people from non white ethnic groups, with a long-standing illness or disability or lone parents and to do so the main survey would need to sample a sufficient number of people from these groups to be able to yield enough interviews to enable such analysis. This is particularly important if any of these groups are less likely to have up-to-date contact details or are expected to be less likely to respond to the survey.

4.6 Sample size

Tables 4.12 and 4.13 give an indication of the available EMA 2 and YCS 12 sample size by main activity status. These numbers are based on the total number of EMA 2 and YCS 12 respondents, (although they do not take into account the number of respondents who agreed to be re-contacted). It is expected that a similar number of respondents per main activity sub-group would also be available to include in the sample frame from EMA 1, YCS 11 and YCS 10.

The research design envisages an achieved sample of between 1,500 and 2,500 divided approximately equally between the four main activity status sub-groups. Assuming that four sub-groups were proposed and the response rate is similar to the pilot study, the initial sample would be in the range 4,000 to 5,000. Tables 4.12 and 4.13 suggest that there would be ample respondents available from the EMA and YCS cohorts in order to achieve a sample within this range.

As discussed in section 4.5 there are a number of ways in which the NEET main activity category could be sub-divided in order to ensure that the variety of within NEET categories are represented. The sample size affects the power of the statistical tests and the precision of the estimates, therefore when thinking about the possible NEET sub-groups it is important to consider how large these groups would have to be to be able to detect a statistically significant difference with reference to key outcome variables.

When considering minimum sample size requirements the required level of precision of the estimate in question needs to be specified. Confidence intervals are a measure of sample precision and show the interval in which the true population value is likely to fall. A 95% confidence interval is constructed in such a way that 95 times out of 100 it captures the true population value that we are trying to estimate, and therefore the interval demonstrates the likely range of the true population measure. A narrow interval suggests a better level of precision. Table 4.25 illustrates how the level of precision changes depending on the sample size and survey estimate. For example, if the sample size were 200 for each main activity status group, the confidence interval around a survey estimate of 10% would be +/-4.2 (5.8%, 14.2%).

Table 4.25 Precision estimates

	Survey estimate			
	5%	10%	25%	50%
Sample Size				
n=200	+/-3.0	+/-4.2	+/-6.0	+/-6.9
n=400	+/-2.1	+/-2.9	+/-4.2	+/-4.9
n=800	+/-1.5	+/-2.1	+/-3.0	+/-3.5

The amount of statistical power attributed to any statistical tests that are likely to be carried out is an additional sample size consideration. Sections 4.3 and 4.4 discuss the possible options in terms of sub-dividing the target population into the differing categories of NEET to enable representation of a number of NEET trajectories. Consequently it is likely that analyses will focus on the differences between the main activity and NEET sub-groups. An example analysis might be whether or not there is a statistically significant difference between the proportions of NEET respondents with Level 2 qualifications compared with those who are not NEET.

The power of a sample is an estimate of the probability that a true difference in the population is detected by a statistical test on the survey sample (at the 0.05 significance level). The greater the power, therefore, the more likely that a true difference would be detected; as a rule of thumb, a power estimate of 0.8 is usually considered to be acceptable. Table 4.26 displays estimates of the power to detect a statistically significant difference between one main activity sub-group and another e.g. NEET / Not NEET or NEET for 2 months or less / Sustained NEET. For a sample size of 200 per sub-group, any differences would have to be larger than 15% to be detectable with an acceptable level of power (0.8 and above). However, if the sample size were 400 or more it would be possible to detect differences of 10% with adequate statistical power.

Table 4.26 Power calulations

		Difference (increase)			
	Lower Percentage	5%	7%	10%	15%
Sample Size					
n=200	10%	0.33	0.54	0.80	0.98
	15%	0.26	0.44	0.71	0.95
	25%	0.20	0.34	0.59	0.89
	50%	0.17	0.29	0.52	0.86
n=400	10%	0.57	0.83	0.98	1.00
	15%	0.46	0.72	0.94	1.00
	25%	0.35	0.59	0.87	0.99
	50%	0.29	0.51	0.81	0.99
n=800	10%	0.86	0.98	1.00	1.00
	15%	0.75	0.95	1.00	1.00
	25%	0.61	0.87	0.99	1.00
	50%	0.52	0.80	0.98	1.00

In order to calculate accurate sample size options it is recommended that the most important outcome variables and the intended analyses with reference to this particular survey are identified.

One possible indicator could be the proportion of respondents without any qualifications; this could then be compared across main activity status groups. To estimate the sample size accurately an estimate of the proportion of NEET and not NEET young people without any qualifications is required. The proportion of 18-24 year olds in 2008 without any qualifications by NEET status is available (Barnham et al 2009:26)[19]. The population estimates suggest that there is a difference of 10 percentage points between the proportion of NEET (33%) and Not NEET (23%) young people without any qualifications. Given the effect size (10 percentage points) if this were a key outcome measure then a sample of 400 per main activity sub-group would be sufficient to observe a statistically significant difference.

[19] In 2008 the EMA and YCS respondents would have been aged 18-25 so this is a reasonable proxy.

In summary, to be able to predict required sample sizes accurately it is recommended that the expected proportions or means of desired outcome variables are estimated within the chosen NEET sub-groups. Depending on the required level of precision and given the available sample we would expect that, subject to a good response rate, the data collected in a full scale survey could be categorised into groups for analysis and statistical differences on key outcome measures detected. The size of the effect that could be detected will depend on the achieved sample size and the number of sub-group divisions required.

4.7 Summary

This chapter has highlighted some key sampling issues that would need to be considered before embarking on a main survey. The pilot survey demonstrated the need for a more definitive activity coding strategy and highlighted ethical issues around re-contacting respondents from previous surveys.

Analysis of attrition between waves of the previous surveys showed that by using only the final wave of survey data from the YCS or EMA would be likely to reduce the pool of NEET respondents and sub-groups of NEET from which the main survey sample could be drawn. Analysis of continuous participants suggests that later waves of the survey were less likely to include respondents who had barriers to employment, education or training or those who were potentially susceptible to deprivation. However, given the ethical situation with regards to recontacting respondents who had dropped out of the previous surveys before the final wave, it may be necessary to accept these limitations or to account for them through weighting.

The chapter discussed how to define young people as NEET and how to divide them into subgroups of NEET. The most robust definition that yields adequate sample sizes takes account of NEET activity status across waves. An alternative or additional definition that could be utilised takes account of monthly NEET activity status; this provides a larger sample of NEET respondents. This also provides the possibility of including a more diverse group of NEET respondents in the definition. Furthermore, breaking down the NEET respondents into sub-groups of NEET duration provides a definition that seemingly accounts for the majority of sub-groups identified in the analysis.

Finally, the chapter discussed how large each of the NEET subgroups would need to be in order to detect statistically significant differences on outcome variables of interest.

5 Recommendations for main survey

This chapter draws on the findings from the pilot fieldwork and the investigation of the sample sources to present recommendations for conducting a full-scale survey.

5.1 Recommendations for sampling

- Base the sample frame on respondents from the final wave of the survey and consider the ethics of re-contacting respondents who had dropped out prior to the final wave.
- Define as NEET young people who identified themselves as NEET at *any* of the waves of data that incorporate young people aged 16-18 (waves 1-3 in both surveys).
- A number of challenges were identified in defining sub groups of NEET young people, but a definition based on NEET duration (in terms of months or years) seems to be the most viable option.
- The evidence suggested that there may be a need to over-sample the NEET young people with identifiable barriers since it proved more difficult to find follow up addresses for this group and tracing this group of respondents is likely to be more resource intensive and result in fewer interviews.

5.2 Recommendations for fieldwork

Tracing respondents

The experience of tracing respondents for the pilot survey was discussed in detail in Chapter 3. The key recommendations are outlined below.

- Make full use of any telephone numbers available for the named respondent from the sample.
- Where available, provide extra information to interviewers about which survey the named contact participated in, the mode of interview and their age at last interview to aide recall.
- Since parents were the 'gatekeepers' for up-to-date contact details for many respondents, it may help to schedule fieldwork to coincide with times of the year that the named sample are likely to return to their parental home (e.g. Christmas).

Making contact and the advance letter

The pilot fieldwork elicited a range of responses from the named respondent and other residents. In cases where making contact with the named respondent was not straightforward, the following procedures are recommended:

- Consider sending letters to 'the resident' at the address so that they are prepared when the interviewer calls and may increase their willingness to share contact details.
- On the envelope for the named respondent, indicate that the letter should be forwarded if necessary.
- It may also be helpful to supply interviewers with study specific calling cards that they could leave with residents, particularly family members who are wary about passing on new contact details due to concerns about privacy. This may help reassure residents about why interviewers are calling and help improve co-operation at the tracing stage.
- Provide spare copies of the advance letter for other residents.

Selling the survey

It was noted in Chapter 3 that this was not an easy survey to 'sell' to respondents since it lacks a clear policy focus and may be of low relevance to adults in their mid twenties. In light of this, the following recommendations are made:

- Interviewers should focus on the respondent's personal experience when talking about the survey, e.g. what their aspirations/desires were when they were 16 and how these might have changed and what they are doing now.
- Find ways to emphasise the relevance of the study for the potential respondents.
- Identify any policy implications which may be used by interviewers to convince potential respondents of the importance of the study.
- Retain a modest incentive payment. Although interviewers fed back that the incentive payment did not appear to affect the willingness of sample members to take part, the evidence from previous research demonstrates that such payments can indeed improve response (particularly among those with low income and education) and representativeness and be cost effective. On some large-scale face-to-face surveys carried out by NatCen within the last few years, a £5 incentive payment has been effective in securing a higher response rate than where there has been no incentive.

Survey mode

The original research specification for this feasibility study required the collection of detailed activity history information in order to compare the amount of time spent in employment according to activity status at 16-18. A face-to-face survey was deemed as the most appropriate survey mode and the pilot survey confirmed that this mode had the following benefits:

- Recall of activities by working through the activity diary in person
- The length of interview could be longer than other modes allowing the collection of more detailed data.
- Aided the tracing of respondents, creating a rapport with residents, following up leads.
- proved to be an appropriate mode for interviewing vulnerable people
- Helped to engage potential respondents more effectively than through a phone call or letter.

We would therefore recommend a face-to-face interviewing mode for a full scale survey as the best way in which to achieve a good response rate and detailed activity history.

However, we recognise that face-to-face interviewing is resource intensive and may be outside the scope of the available budget. We have therefore costed for alternative options: a telephone survey and a qualitative study. The following chapter describes these options in more detail, but we outline below the issues that would be pertinent in deciding on the mode and scope of a main survey.

Level of activity detail required

For a high level of detail, as collected in the pilot survey, a face-to-face interview is likely to be the most suitable mode. The presence of the interviewer helped to keep the respondent engaged with the task and they were able to prompt them to recall their activities through the use of the activity diary.

A telephone survey would be a viable alternative if it was accepted that fewer years were covered and/or a lower level of detail was collected about each activity. NatCen has recently conducted a

number of telephone surveys with a similar group of respondents where some information about current and past activities (including qualifications) has been collected successfully including the Activity and Learning Agreement Evaluations, the Activity Agreement Follow up Survey, Adult Learning Grant evaluation and the ESF Cohort Study.

Making contact

The quality and type of contact details available would also be an important consideration in deciding on the mode of interview. Relying on telephone contact, particularly for respondents who only participated in the first and second waves of the EMA survey and YCS, is likely to yield a lower number of traced respondents and interviews.

Furthermore, the opportunity to create a rapport with residents at issued addresses and stable contacts is more limited by telephone, particularly for obtaining up-to-date contact details from reluctant family members.

Vulnerable respondents

For some groups of respondents, for example those with learning difficulties or other communication barriers a telephone survey may not be practical. These groups of people are also believed to be more at risk of being NEET. Therefore, if a main survey is conducted by telephone some provision of face-to-face interviews for more vulnerable respondents should be considered to ensure that this group are not excluded.

5.3 Recommendations for questionnaire development

The following recommendations are based on the pilot survey and therefore assume a face-to-face survey. Some of the recommendations would nevertheless be relevant in the context of other modes.

- The activity diary was considered a useful tool and we recommend that it is developed according to the following points:
    - Simplifying the form - leave the columns blank and provide more space for notes
    - Fit the diary into one A4 page
    - Interviewers reported that some respondents found it easier to work forwards in time when recalling their activities and the activity diary layout could be changed accordingly. The benefits would be to obtain more accurate information and retain the engagement of the respondents in the exercise.
    - Provide interviewers with prompts which can be used to aide date recall e.g. where were you living when you started this job?
    - Consider including a list of national/international events to jog memory.

- Keep the CAPI program flexible in terms of the order activities are recorded, allowing interviewers to move forwards or backwards in time.
- Reduce the questionnaire length to help avoid respondent fatigue; this would also help interviewers to ‘sell’ the survey if the average time was shorter.
- Reduce repetition when recording activities, this particularly relates to the following sections:
    - The life events questions on living arrangements

  - Benefits questions (this section could be asked only of respondents who did not give permission to link interview data with administrative data held by DWP)
  - Employment questions on training
  - Detailed qualification details when higher qualifications have been achieved (i.e. collect summary information about the time spent studying for qualifications but only collect detailed qualification information (e.g. type, level, full/part etc) for the highest qualification).
- Find a balance between gathering accurate activity data for each month and minimising the burden on respondents to recall this.
- Make use of more showcards to reduce the number of 'read out' questions.
- Other areas of development include recording more details about long-term illnesses or disabilities, focusing detailed data collection of main spells of activity so as to not spend too much time on short-term jobs, including a 'minimum wage' option for pay questions.
- Review question wording of the data linkage question to ensure this is consistent with other DWP/DCSF studies where necessary.
- If permission to link data is given by the respondent, collect date of birth and National Insurance Number.
- Review data linkage handout to include contact details if the respondent wishes to withdraw this permission.

5.4 Summary

In this chapter, we set out the recommendations for a full survey arising from the pilot fieldwork and investigation of sample sources. Among the lessons learned for making contact with the respondents were to use a range of approaches including face-to-face contact with residents and neighbours, introductory letters and telephone contact with the resident and stable contacts. Given the low salience of the survey for many respondents and the time since the last interview, the potential to build rapport with the respondent and residents through face-to-face contact was considered valuable in many cases.

We recommend a face-to-face survey as the most suitable approach for collecting the data required to answer the research questions set out in the specification. However, we consider the alternative of a telephone survey in this chapter and also consider an alternative qualitative study in the following chapter.

6 Feasibility assessment

This section considers the overall feasibility of carrying out a main survey to assess how activity at 16-18 years impacts on future pathways and also sets out an alternative qualitative study.

6.1 Feasibility of a main survey

Findings from the pilot survey indicated that very few sample members were still resident at the address where the family had been living at the time of EMA/YCS. However, via parents, neighbours and 'stable address' contacts, interviewers were successful in gaining at least some information on the whereabouts of (approximately 4 out of 5, 76) issued sample members. As may be expected, the probability of making contact varied, and the sample members with a record of being NEET at some time in the past, and some groups of young adults identified as having barriers were both more difficult to trace and, when contact was made, were less likely to agree to be interviewed.

As discussed in the previous chapter, an important issue affecting the choice of data collection method is the value of obtaining a complete activity history. If this is felt to be critical to the study, then face-to-face interviewing will be essential. On the other hand, if the new data collection effort focuses on main activities in a recent period of 2-3 years, then it may be possible for the interview to be conducted by telephone. To some extent this is a matter of the length of interview, given that a full activity history would require too much time for a telephone interview. It is also a matter of the presence of the interviewer encouraging respondents to provide more accurate data. If the activity history data is an important aspect of the project, then face-to-face interviewing is very desirable.

There is an issue of how much the main data collection method might benefit from use of the other method of contact. In other words, supposing that telephone interviewing was going to be the main method of data collection, to what extent could the tracking of sample members be improved by having interviewers able to make inquiries 'on the ground'? Or, if the main data collection method was face to face interviewing, to what extent would it be efficient to make inquiries initially by telephone?

In the view of NatCen's research team, if a detailed activity history covering the years since leaving school is required, then a face-to-face mode would be the recommended approach. However, we consider that a worthwhile study could be carried out by telephone if the following limitations were acceptable: (1) a lower level of detail about activities and (2) uneven coverage of sample groups and potentially a higher level of bias. As well as limiting the activity history, the chief limitation of a purely telephone approach could be relatively poor coverage of the NEET and Job Without Training sample sub-groups. The pilot survey indicates that the more socially excluded young people could be least likely to be traced by a telephone survey approach. However, even in this situation, we expect that administrative data could provide the means to identify the potential (at least in broad terms) for bias in the survey findings.

One of the chief advantages of face-to-face interviewing is that it provides the opportunity for interviewers to conduct tracing 'on the ground'. The pilot survey suggests this effort can be fruitful in many cases, including cases where telephone calls would be unproductive from an early stage. This more pro-active approach could be a significant factor in locating and in motivating the previously NEET sample members to take part in the study, even in cases where a parent or other intermediary is involved. A personal caller may be more effective at gaining the support of third parties (e.g.

parents, siblings or neighbours), as compared with a telephone interviewer. However, the interviewer can also make use of telephone numbers where they have been provided and result in contact with the sample member or someone who is able to help in tracing the sample member.

An additional consideration is the impact of sample dispersion. The YCS sample is relatively dispersed and could be costly to include except where it falls within areas in which the EMA study was conducted. The members of the EMA sample have of course partly migrated to other parts of the country, and these sample members would not necessarily be accessible to one of the interviewers assigned to the project. However, a viable sample can be provided for a face-to-face survey that would largely be localised in the original EMA areas.

6.2 An alternative qualitative study

This study was commissioned to provide an assessment of the feasibility of undertaking a quantitative survey to assess the medium term impact of individuals spending a period of time not in education, employment and training (NEET) at ages 16-18, and an assessment of comparative outcomes for those in jobs without training, jobs with training, and full time education between 16 and 18.

A qualitative alternative is also possible, although it is important to be clear about the research questions it could address. It would not, for instance, be possible for a qualitative study to undertake an assessment of comparative outcomes. Qualitative work would however be well suited to exploring individuals' experiences in work, training and education, the pathways followed and the nature of their transitions into adulthood, and the factors affecting the nature of these pathways and transitions.

The Department has indicated that it would wish any qualitative study to focus attention more widely than solely on individuals defined as NEET and include those defined as in JWT, jobs with training and in FTE at ages 16 to 18. Inclusion of this wide range of sample groups would necessitate a relatively large qualitative sample and this, of course, has implications for the cost of such a study. Another option would be to focus a qualitative study only on people defined as NEET or in JWT at ages 16 to 18 and explore the experiences and pathways of sub-groups within this population as outlined above. This would likely necessitate a smaller sample and thereby reduce the cost of this option (the cost of designing and setting up the study would remain the same but there would be savings for recruitment, fieldwork and analysis).

A series of qualitative in-depth interviews, each lasting between one and one-and-a-half hours, would be appropriate and would provide the opportunity to explore in detail an individual's trajectory since the age of 16-18. This will require a high degree of collaboration and cooperation between the researcher and the participant which will be best achieved in a one-to-one encounter.

Participants could be sampled from the EMA and/or Youth Cohort surveys as for the quantitative options outlined above. From the figures presented by the quantitative feasibility study (where 35 interviews were achieved from a sample of 90), we would anticipate that the sample frame would need to be at least three times the size of the desired achieved sample. This would likely need to increase significantly if no face-to-face recruitment (telephone only) was undertaken and if further sampling criteria (in addition to categorisation as NEET or otherwise at 16-18) were to be included. We would anticipate that the information provided in these data sets would need to be supplemented with additional information, principally with up to date contact details for participants (parental or other

stable contact details may be used in the first instance) but also other relevant details about their activities since 16-18. This could be obtained via a short, telephone screening exercise during which potential participants would be asked some brief questions to determine their pathway since 16-18, their interest in participating in the study and to confirm their stable contact details. Those who were willing and matched the sample criteria would be invited to take part in an interview. Potential participants would be sent an advance letter to introduce the study and advise that a researcher may get in touch with them about the study.

We have assumed that a good proportion of survey respondents supplied parent or stable contact details, and that it will be possible to make contact with a good proportion of potential participants using the contact details for their parents or another stable contact supplied by the survey sample. Recruitment via telephone may therefore suffice, without the need for any face-to-face recruitment, however it would be sensible to revisit this assumption with the survey samples if such a study were to go ahead.

The samples from the EMA and Youth Cohort surveys are likely to be geographically dispersed. It would be advantageous to be able to cluster qualitative fieldwork and so we would recommend selecting sample to receive an advance letter in a small number of areas. The extent to which this would be possible would need to be investigated and would be dictated by the amount of sample available in either survey in any one area. Were fieldwork not able to be clustered in this way, this would impact on the cost of the study.

Whatever the focus of a qualitative sample, we would recommend that at least 10 participants from each primary sample sub-group (e.g. NEET, JWT, jobs with training, FTE at 16-18) were included in the study. This would provide sufficient scope to capture diversity within the overall sample against other criteria such as demographic characteristics, work and training history, and educational attainment. It may also be possible to purposively select participants on the basis of their pathway since the age of 16-18.

Appendix A Advance letter

P2984
February 2010

The Twenties Study

Some years ago when you were aged between 16 and 19 years, you kindly took part in a study about your experiences at school and any activities you had done since leaving school. You may remember the study being called 'Pathways' or 'Destinations for 16-19 year olds'. These studies were conducted by the National Centre for Social Research (NatCen) on behalf of the Department for Children, Schools and Families (DCSF) formerly known as the Department for Education and Skills (DfES).

Thank you for helping to make these studies a success. The findings have been used by the Government to plan the support given to young people after they leave school.

We would now like to talk to you again to find out what you have been doing since we last spoke to you. This would include any jobs, learning-related activities, caring for children or other activities. We also wish to find out about the things that make it difficult to take up jobs and learning opportunities, so if you haven't been working or studying over recent years your views are still important to us. Your participation will help us get an even clearer picture about how young people make the transition into adulthood and what support that they need.

A NatCen interviewer will contact you in the next few weeks to arrange an interview at a convenient time. Your answers will be treated in strict confidence in accordance with the Data Protection Act. No information that can identify you will be passed to anyone else without your permission.

I very much hope that you will be able to help us with this important study. If you have any questions about the study please call us on 0800 652 4574 (Monday to Friday 9:30am-5:30pm) or email us on twenties@natcen.ac.uk.

We look forward to speaking to you. Thank you in advance for your help.

Yours sincerely,

Deborah Bagheri

Appendix B Data linkage sheet

Permission to Link Data

In order to make the information we collect on this study more useful we would like to link your answers from these questions to administrative records held by the Department for Work and Pensions. These records hold information about what benefits and tax credits people are receiving and details about the periods of time when people have a job.

In order to do this we need your permission to pass your full name, sex, date of birth and address to the Department for Work and Pensions.

- The information will only be used for research and statistical purposes.
- The information will be kept confidential.
- Your personal details will **not** be passed to anyone else outside the research team.
- The information will **not** be used to work out whether anyone is claiming benefits or tax credits they should not be.
- Any current or future claims for benefits or tax credits will not be affected.
- You can withdraw your permission to link to these records at any time.

Appendix C References

Barham, C., Walling, A., Clancy, G., Hicks, S. and Conn, S. (2009) *Young people and the labour market,* Economic & Labour Market Review. ONS, Vol 3, No 4.

DCSF: Participation in Education, Training and Employment by 16-18 Year Olds in England, June 2009, http://www.dcsf.gov.uk/rsgateway/DB/SFR/s000849/SFR12_2009v2.pdf

EdComs (2007). *90% Participation Project Desk Research* (DfES Research Report RW102). London: DfES http://www.dfes.gov.uk/research/data/uploadfiles/RW102.pdf

Ofsted (March 2010) *Reducing the numbers of young people not in education, employment or training: what works and why*. Ofsted.

Payne, J. (2000) *Youth Cohort Study: Education, Training and Employment of 16-18 Year Olds in England and the Factors Associated with Non-participation,* Statistical Bulletin Number 02/2000, Nottingham: DfEE.

Perry, J., Purdon, S., D'Souza, J., Finch, S., (2009) *DCSF: Learning Agreement Pilots - Quantitative Evaluation*, DCSF-RR119.

Rennison, J. Maguire, S. Middleton, S. and Ashworth, K. (2005) *Young People not in Education, Employment or Training: Evidence from the Education Maintenance Allowance Pilots Database,* Nottingham, DfES-RR628.

Spielhofer, T., Walker, M., Gagg, K., Schagen, S., O'Donnell, S. (2007), *Raising the Participation Age in Education and Training to 18 Review of Existing Evidence of the Benefits and Challenges.* DCSF-RR012.

Spielhofer, T., Benton, T., Evans, K., Featherstone, G., Golden, S., Nelson, J. and Smith, P. (2009) *Increasing Participation: Understanding Young People who do not Participate in Education or Training at 16 and 17,* DCSF-RR072.

Tanner, E., Purdon, S., D'Souza, J., Finch, S., (2009) *Evaluation of Activity Agreement Pilots: Report of the Quantitative Evaluation,* DCSF-RR096.

Ref: DFE-RR035

ISBN: 978-1-84775-797-5

September 2010